Coarse-Grained Modeling of the Assembly and Mechanical Properties of Viruses

María Aznar Palenzuela

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Coarse-Grained Modeling of the Assembly and Mechanical Properties of Viruses

María Aznar Palenzuela

Departament de Física Fonamental
Facultat de Física
Universitat de Barcelona

Universitat de Barcelona

Facultat de Física

Departament de Física Fonamental

Coarse-Grained Modeling of the Assembly and Mechanical Properties of Viruses

Programa de doctorado: *Física*

Línea de investigación: Física de la materia condensada

Director y tutor de tesis: David Reguera López

Memoria presentada por María Aznar Palenzuela para optar al título de Doctor en Física (25 de Septiembre de 2013).

Agradecimientos

Una gran parte de esta tesis es gracias a las personas que me han apoyado y acompañado a lo largo de los últimos años. Una gran parte es mi familia, especialmente los más cercanos como mis padres, mi hermana y mi tía Valle. Que casi se han alegrado de mis éxitos más que yo, agüita!.

También me gustaría agradecer a esos amigos de la infancia (y no tan infancia) que pese a la distancia están en mi vida de una manera activa y para lo que se les necesite, como Jorge, Elia, Ulises, Carola, Emilio, Bea, Yonte y Néstor. Mucha gracias Churri por aguantarme en los momentos de llantina y los de jarana.

Gracias a ese gran grupo, enorme, que son mis compaeros de Física de ULL. MariPili, Carla, Yohana, Hector y Kike. Esos email colectivos en los que se ve que no sólo a mi se me va el baifo, porque esos sueños paranohicos de Kike o Hector, no tienen precio. Kike, gracias por esas chuletadas, moding mierda, noches del Skecht, Renes Barbier, etc ...

Muchas gracias al 7 islas, la reina de África y Libelula y especialmente a su gente, por esos buenos momentos de cervezas y esparcimiento. Siempre me han dado todo lo que he necesitado, y más.

No me puedo olvidar mencionar a esas maravillosos tangueros que me han enseñado tanto, Fernando, Gonzalo, Alex, Carmen, Yago, Pablo, Dr. Thompson, Lucia y Javier. Muchas gracias por esos buenos momentos de milonga y por sacarme a bailar cuando aún me pisaba a mi misma, ahora sólo les piso a ustedes.

Unas de las personas que más me han animado con este trabajo es Sebastián. Sin él, entre otras muchas cosas, no existiría esta maravillosa portada y probablemente mucho del trabajo de esta tesis.

En los años que llevo formándome, los primeros que me hicieron apasionar por las matemáticas y la física fueron Cutillas y Luis Balbuena, en el instituto. Estoy especialmente agradecida a Javier Hernandéz Rojas por poner en mis manos el primer paper de virus.

Gracias a mis primeros compañeros en Barcelona, Nuria Tabener, Mary, Irina, Ronald, Nuria Amigó, Jordi, Xavi, Fabiana. Con especial cariño a Ignasi que a parte de compañero y amigo, compañero de piso durante 3 años y que por tanto le ha tocado aguantarme un poco más que el resto.

Durante estos últimos 4 años he tenido muchos buenos amigos compartiendo despacho con Nidal, Nico, Martin, Adriana, Guillem e Iván. Compañeros de almuerzos siempre muy animados como Álvaro, Ricard, Noela, Cristina, Benedetta, Xavi, Andrea, Pau, Joaquin, Thomas, Paolo, Valentino, Fabio, Carlos, Marco. A Carlos y los juguitos y en especial a Paquito con el que he pasado tan buenos momentos de mezcales, pozole y las veces que casi me echan del Shangó, darle las gracias por estar ahí.

No me puedo olvidar de compañeros del departamento de ECM, Laura, Xavi, Guillermo, Carles, Jordi y porsupuesísimo Claudia. Claudia y las infinitas horas que habremos pasado tomando juguitos, terremoteando, luchando con rosas y saludando a Pepe. Sin ti todo hubiera sido mucho más aburrido.

Especialmente quisiera dar las gracias a mi director David Reguera por lo muchísimo que he aprendido con él y sobretodo por la paciencia que me ha tenido.

Gracias a todo el laboratorio de AFM de la Autonoma de Madrid, especialmente a los componentes de la mesa de los niños, Álvaro, Pablo, Gabriela, Ana, Merche. Gracias Merche por esas cañas y por la ayuda con los capítulos de AFM.

Table of Contents

Introduction

Viruses have been present in human life since immemorial times and stand out in the biological context by their efficiency and relative simplicity [1]. They have a well-defined structure, they can survive to very harsh conditions, and replicate efficiently in all kinds of hosts.

Until recently, mankind was only aware of the existence of viruses through the diseases caused by them. The first recorded evidence of a viral disease dates back to 1800 BC and is an Egyptian stele representing a person with a withered leg showing the effect of paralytic poliomyelitis. From this initial evidence to nowadays there are many examples of the devastating effects caused by viruses such as smallpox, that was endemic in China by 1000 BC, the influenza of 1918 (Spanish Flu) which took over 40 millions of lives, or more recent emergent viruses as Ebola or AIDS [2, 3].

However, although the diseases caused by viruses have been known for more than 10000 years, the nature of the infectious element (the virus) was unidentified until nearly a century ago. In fact, the first primitive vaccinations (variolation) for the smallpox were performed in China and India in the 11th century without knowing the causative effect. But there was no rationale about infectious elements, until Robert Koch and Louis Pasteur, who proposed the “germ theory” in the 1880s [2, 4]. Neverthreless, at that early time they were not able to distinguish between different infectious agents yet, for example bacteria and viruses.

The modern concept of a virus as a submicroscopic biological entity was born in 1890s. Dimitrii Ivanovsky was the first to report in 1892 that the causative agent of tobacco mosaic disease was not detained even after being passed from plant to plant through filters fine enough to trap the smaller known bacteria [4]. In 1898 Martinus Beijerinck made the same observations independently, and identified the responsible promoter for the disease as a distinctive agent. These new entities were called “ultrafilterable viruses” (from the latin word “virus” meaning poison), and eventually simplified to viruses [2]. However, the idea of human viruses was not accepted until 1909 by Karl Landsteiner and Erwin Popper, who showed that poliomyelitis was caused by a virus [5].

The study and information about viruses developed in parallel with the new technologies. In particular, the introduction of the electron microscope in the 1930s revolutionized virology, see Fig. 0.1 and 0.2. The first study about the structure of viruses was made in 1935 by Wendell Stanley, who crystallized tobacco mosaic virus (TMV) [6], aided by its structural simplicity. The study of the TMV structure, originated the first viral classification, based on their morphology [7].

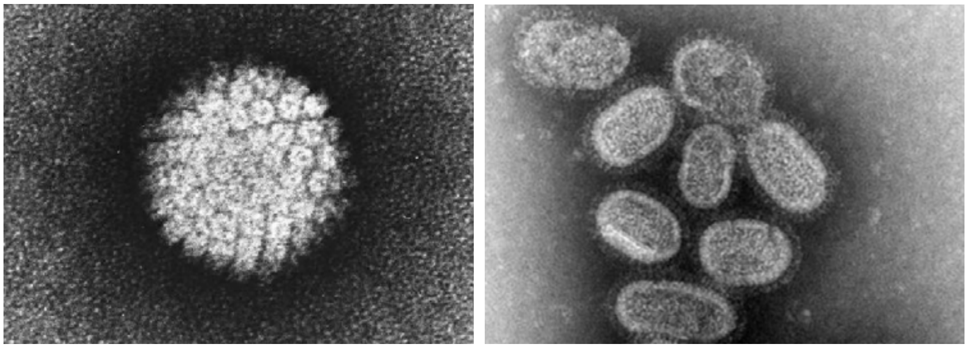

(a) (b)

Figure 0.1: Electron microscopy images of two examples of human viruses: (a) papilloma virus (taken from Ref. [8]) and (b) influenza A (taken from Ref. [9]).

The study of bacteriophages (viruses that infects bacteria), was decisive for virology and for the development of molecular biology. Studies by Max Delbrück, showed that phages are stable, and self assembled following common characteristics. Moreover, in 1952 Alfred Hershey and Martha Chase [10] showed that the two main components of viral particles, proteins and the genetic material, were used for their molecular replication in the host cell.

Nowadays we know that viruses infect humans, animals, plants, insects, bacteria, archea and, interestingly, also other viruses. These viruses which infect other viruses once these have taken control of the cell are called virophages [13]. Given the huge variety and nature of hosts, viruses show different characteristics and have developed diverse replication strategies. However, viruses can be defined in general as biological entities having the following properties, listed in Ref. [2]:

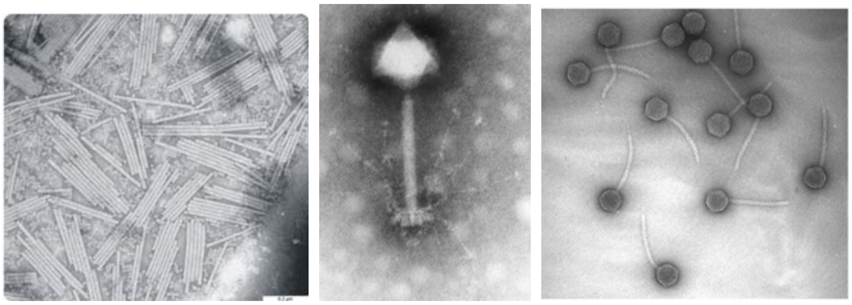
“* A virus is an infectious, obligate, intracellular parasite.

* The viral genome comprises either DNA or RNA.

* Within an appropriate host cell, the viral genome is replicated and directs the synthesis, by cellular systems, of other viral components.

* Progeny infectious virus particles, called *virions*, are formed by *de novo* assembly from newly synthesized components within the host cell.

* A progeny virion assembled during the infectious cycle is the vehicle for transmission of the viral genome to the next host cell or organism, where its disassembly leads to the beginning of the next infectious cycle.”



(a) (b) (c)

Figure 0.2: Electron microscopy images for three examples of viruses: (a) A plant virus, TMV [6], and two bacteriophages, (b) Phage Synechococcus-PM2 [11] and (c) λ [12].

There is a huge variety of viruses whose, life cycle differs depending on the type of host cell. The most simple virus cycle starts in the extra-cellular phase, where the virus recognizes the host cell surface. It then penetrates inside the cell (entirely or some of its components) and travels through the cell. In a second step the virus disassembles or releases its genetic material at some place of the cell and the expression and replication of the viral genome takes place. After that, the protein shell self-assembles (in some cases, by assembling first an empty protein shell and packing the genome, and in other cases the shell and genome co-assemble together). In many viruses, the last stage of assembly involves a maturation step needed for the virus to become infective. An inter-cellular travel and the release in some virus takes place at different moments of the assembly and maturation steps. Finally, the released viruses travel in the extra-cellular environment to find a new host, to start the cycle again. This is a general viral life cycle but it is important to bear in mind that each virus follows specific steps and strategies [14].

Despite the many different types of viral life cycles, all viruses are made by at least two essential ingredients: the genetic material and a protein shell (called *capsid*). Despite the presence of their own genetic material, viruses need to infect and use the molecular machinery of host cells to produce new viral particles under the right cellular conditions, since they cannot do it by themselves [15].

The pathway of replication in the host depends very much on the type of genome. The virus genome can be encoded in RNA or DNA, which can be double-stranded (ds) or in other cases, single-stranded (ss) [2]. The first viral classification, known as *classical viral classification*, was made in terms of the nature of the genetic material, the symmetry of the capsid and the presence or not of a membrane, as shown in Figs. 0.3 and 0.4 for plants and vertebrates, respectively [16].

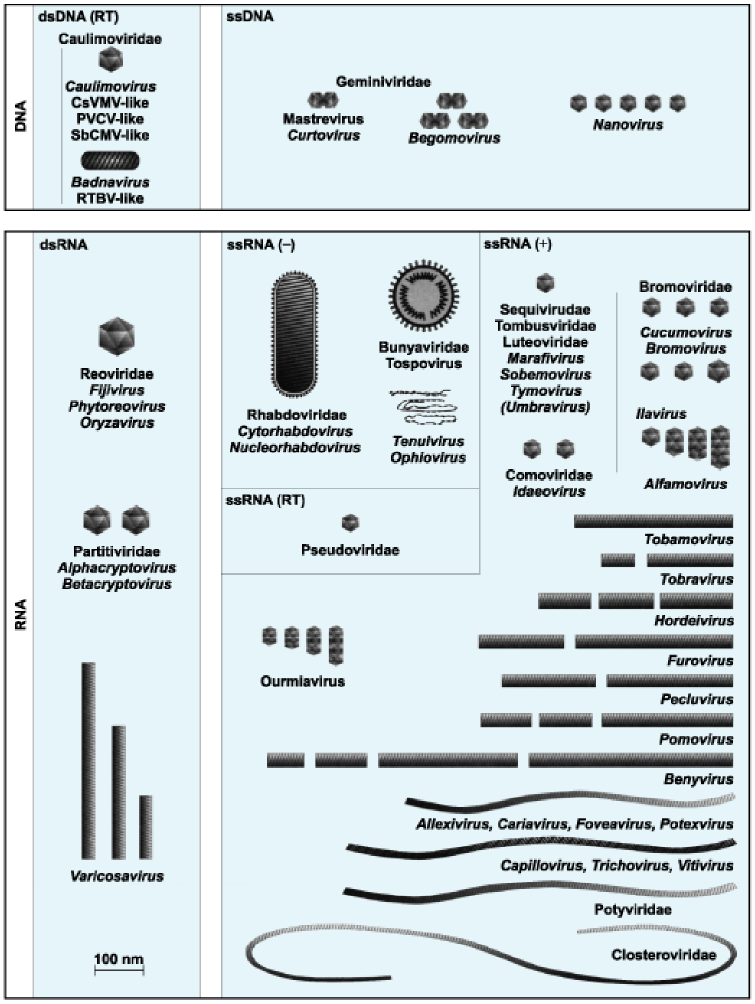


Figure 0.3: Families and genera of viruses infecting plants classified in terms of the type of genetic material (i.e ds and ss DNA and RNA genomes). Figure taken from Ref. [16].

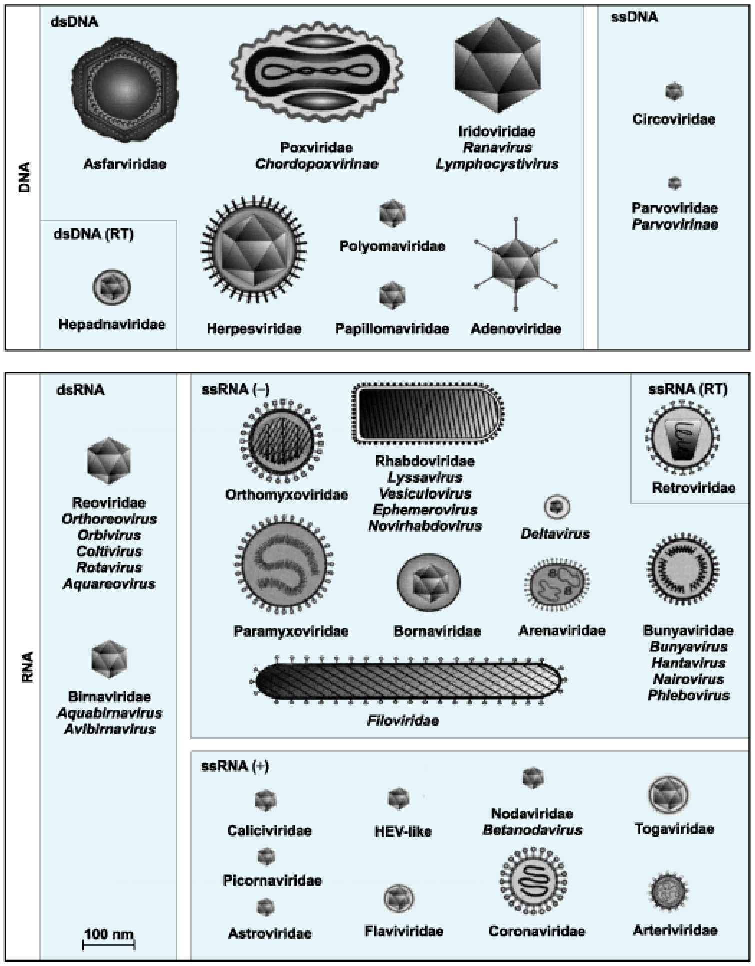


Figure 0.4: Families and genera of viruses infecting vertebrates, classified in terms of the type of genetic material (i.e ds and ss DNA and RNA genomes). Figure taken from Ref. [16].

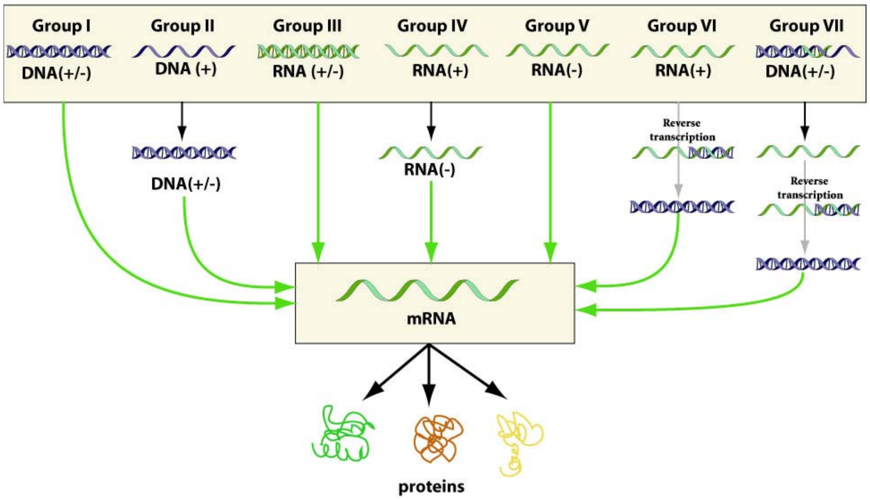
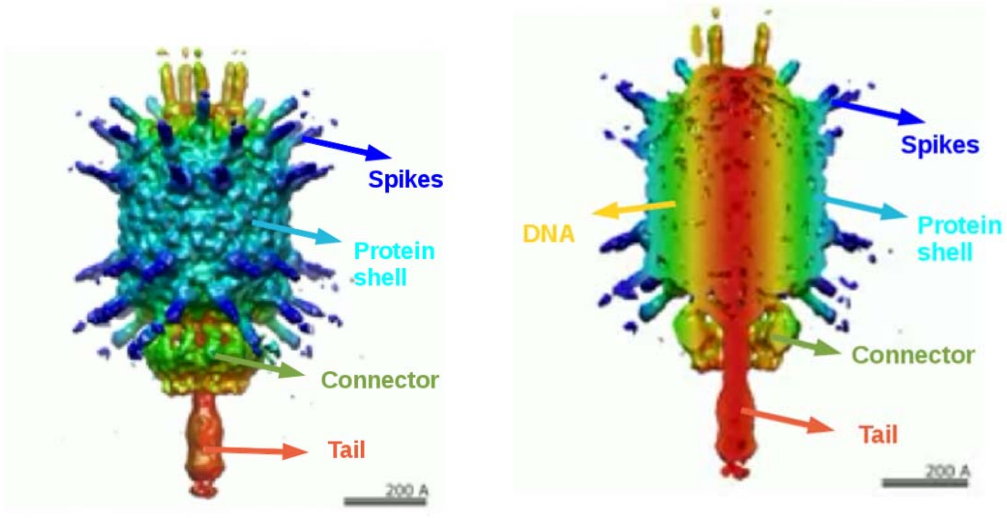


Figure 0.5: Baltimore classification of viruses, where the pathway from the viral genome to mRNA and the nature and polarity of their genome define the viral class. Figure taken from Ref. [2].

Cells typically express the genetic information using messenger RNAs (mRNAs) that are translated in the cytoplasm by associated machinery. Viruses use mRNA to produce proteins, but depending of the cell machinery of the host the synthesis steps are different. For example, herpes Virus replicates its genome using a host polymerase, whereas viruses such as picornavirus, require a viral polymerase to transcribe first the genome. The viral protein synthesis inspired the *Baltimore classification*, where different viral groups are defined depending on their type of genome and method of replication, see Fig. 0.5. Remarkably, the origin of the viral genomes types and their methods of replication are still in debate [17–19].

The second main ingredient of viruses is the protein shell or *capsid* (see Fig.0.6. Its main role is to protect the genome from environmental conditions and to facilitate its safe delivery into a new host. Thus, between the formation of the viral particle and the infection of a new host, viruses rely on the capsid to protect their genomes. The viral shells are made of multiple copies of one or a few different proteins that spontaneously self-assemble, without consumption of chemical energy [2–7, 10–21]. Viruses are built by a set of common architectures for the capsid, having between a few nanometers to hundreds of nanometers in size. The shape and symmetry of the capsid is an important element in their classification.

Typically, each native virus self-assembles *in vivo* into a unique viral capsid shape. For instance, TMV, the first virus observed using an electron microscope, has a rod-like structure with coat proteins organized in a helical sheath, opened at both ends [22], see Fig 0.2(a). The most frequently observed viral capsid shapes are spherical [23], prolate [24, 25] or conical [26]. Other less abundant viral shapes have



(a) (b)

Figure 0.6: 3D Electron Microscopy reconstructiong of bacteriophage $\phi 29$: Image (a) shows the exterior of the capsid with spikes, the connector and the tail. Image (b) shows a cross-section of the same bacteriophage full of DNA. Figures taken from Ref. [48].

also been observed, such as the bullet-shaped particles of some animal viruses like rhabdovirus vesicular stomatitis virus [27] or the amazing variety of morphologies reported for several virions that infect extremophiles archaea [28]. Nevertheless, there are many viral structures that have not been totally characterized and studied yet. Moreover, many viruses can be reconstituted *in vitro* in different non-native shapes, by playing with the assembly conditions or with the properties of the cargo.

Some, viruses may have additional components besides the coat proteins and the genetic material. For example, many bacteriophages have a tail connected to the capsid that facilitates the attachment and injection of the genetic material into the host [29]. Other viruses have spikes at specific positions on the capsid, which are used for recognition and binding to the host cell [30, 31], see Fig. 0.6. A prevalent element in many animal viruses are lipid membranes that depending on the specific virus, play an important role in the formation, maturation, entry or disassembly processes during the viral cycle life. For instances, lipid membranes help the assembly process in PRD1 virus [32] or is involved in the assembly, maturation, and infectivity of HIV [33].

There is more diversity between viruses than plants, bacteria and animal all together, which turn the study of viruses into a huge task [4]. An added difficulty is that viruses are continuously evolving, as exemplified for instance by the case

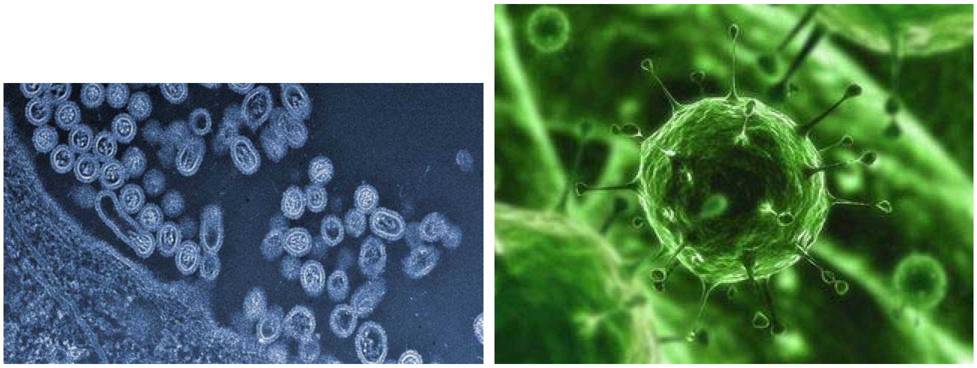

(a) (b)

Figure 0.7: Two examples of new viruses. (a) H7N9 avian virus, which killed hundreds of corral birds last year (figure taken from Ref. [34]). (b) A new coronavirus, which infects humans, detected for the first time last year (figure taken from Ref. [35]).

of avian influenza H7N9, see Fig. 0.7 (a). Moreover, new viruses emerge and originate, as in the case of the Middle East Respiratory Syndrome coronavirus first reported on September 2012, see Fig. 0.7 (b).

In the last decades, our understanding of viruses has significantly increased, but there are many questions whitout answer. As mentioned before, each virus is a world in its own, and it is full of specific details in its life cycle and infection processes. However, there are some general features and abilities common to all viruses. Moreover, viruses lack the sophisticated and active machinery of cells, and work most of the time using passive and spontaneous processes. That suggest that, during their natural cycle, they must rely on general physical and chemical principles to succeed in most of their tasks, and to resist the different possible extreme physiological conditions of the environment.

The study of the physics involved in the assembly of viral capsids, their stability, or their mechanism of infection/replication, is very important in order to understand the properties of viruses that are common to different species. The many remarkable properties have stirred the interest in the study of viruses to develop novel strategies aimed at stopping or minimizing the damage of viral infections, or to use their remarkable abilities for biotecnological applications. In fact, viruses have already been used in gene therapy [36], as an alternative to antibiotics [37] or as insecticides [38]. In nanotechnology viral capsids are envisaged as nanoreactors [39], as templates for nanowire fabrication [40] or as nanocontainers [41, 42].

A better understanding of the physical mechanisms involved in the viral lifecycle may also be useful in many biophysical applications and antivirals. Additionally, viruses are composed by proteins, phospholipid elements and genetic material which are also the essential elements of other biological systems. The physical principles governing viral capsids might thus be also important for the understanding of other biological systems with similar shapes or components such as vesicles [43, 44], vault shells [45] or red bloods cells [46], and could be helpful to produce artificial structures with similar properties [47]. But still there are many open questions about, for instance, the kinetics of viral assembly or their mechanical properties, which remain to be solved, and which will be studied in this thesis.

Scope of the thesis

The main goal of this thesis is to study the physical mechanisms implicated in the self-assembly and mechanical properties of non-enveloped viral capsids using coarse-grained models. The thesis has been divided into three main blocks discussing the physical modelling of viruses, the viral self-assembly process, and the mechanical properties of viral capsids. The thesis combines theoretical analysis, mostly developed in the first part, with different simulations, developed in the second and third parts. We have tried to discuss the results in the context of previous studies, and to compare them with experiments. In particular, the simulation of the third part on the mechanical properties of different viruses have been specifically developed to interpret experimental results obtained by atomic force microscopy (AFM) nanoindentation experiments.

The first part of this thesis (*Physical modeling of viruses*) focuses on summarizing some common properties of viruses. Using these common properties it is possible to propose physical theories in order to gain insight into general behavior of different viral processes such as the viral architecture or self-assembly. We will review important ideas from previous works and we will introduce a new coarse-grained model and the main simulation techniques that will be used in the thesis.

In the first chapter (*Viral architecture*), we will revisit the geometrical principles involved in the construction of spherical and prolate viral shells. First, we will review the Caspar and Klug construction for spherical viral capsids with icosahedral symmetry, which is the basis of their structural classification in terms of their triangulation number T. We will emphasize the organization of the T-capsids in terms of classes, which groups shells having similar geometrical characteristics, and as we will see, also common physical properties. The extension of Caspar and Klug's ideas to classify bacilliform viruses will be also described. The chapter will be ended by discussing the interesting physical origin of the prevalence of icosahedral symmetry in spherical and bacilliform capsids.

Chapter 2 (*Theoretical description of viral assembly kinetics*) introduces the general theoretical framework used to characterize the self-assembly of viral capsids. In particular, after reviewing the common features of viral assembly, Classical

Nucleation Theory (CNT) will be described as a powerful framework to characterize the kinetics of the viral assembly process. Using this theory it is possible to explain the universal traits observed experimentally in the assembly of different viruses such as the steep concentration dependence, the lag time at the first stages of assembly, and the sigmoidal kinetics of capsid production. We will also use this theory to analyze and interpret our simulations.

The third chapter (*Coarse-grained modeling of capsids*) reviews different strategies and models that have been proposed to study viral capsids. In particular, we will revise previous coarse-grained models that have been used to describe different aspects of capsid structure and assembly, highlighting their advantages and shortcomings. Building up on previous models, we will then present our coarse-grained model that will be used in the following chapters to analyze the self-assembly and mechanical properties of viruses.

Chapter 4 (*Elastic modeling of viral capsids*) introduces a simple theoretical model connecting the most relevant elastic properties of viral capsids with the microscopic interactions between coat units proposed in our model. These ideas will be used to formulate a simple continuum elastic model in terms of bending and stretching energies. We will analyze the bending-dominated, and the stretching-dominated limits, and how these two contributions compite and determine the size and mechanical properties of viral shells.

The last chapter of this part focus on *Simulation Methods*. We have implemented diverse simulation techniques for the study of different viral processes. In this chapter we will describe the three main methods used: Monte Carlo (MC) simulations, Brownian Dynamics (BD) simulations, and Finite Elements (FE) simulations. Monte Carlo simulations will be employed to study capsid stability, using our coarse-grained model. In turn, Brownian Dynamics simulations will be used to study the dynamics of capsid assembly, and to mimic nanoindentation experiments in our "Virtual AFM". Finally, Finite Elements analysis will be introduced as a tool for the study of some mechanical properties of the capsid in the continuum limit. These different techniques will be used in parts II and III of this thesis.

The second part of the thesis (*Self-assembly of viruses*) is dedicated to the analysis of the viral self-assembly and the physical ingredient that control this process and the final capsid shape.

In particular, Chapter 6 will focus on the *Physical ingredients controlling the polymorphism and stability of empty viral capsids.* Using the model proposed in Chapter 3, we will implement a MC simulation to study the stability of the smallest spherical structures and their tolerance to changes in bending stiffness and spontaneous curvature. In this study we will characterize the regions of stability, the competition between different structures, and the morphological changes suffered by the different structures. These results are discussed in connection to their potential biological implications. The aberrant structures that appear when the capsid loses its integrity will also be examined, as a guidance for assembly experiments.

The second chapter of Part III (*Simulations of the assembly kinetics of empty viral shells*) will present the results of Brownian Dynamics simulations of the assembly of empty capsids using our coarse-grained model. We will see that the optimal conditions required for assembly are not necessarily the same as the ones that warrant the stability of the capsid. We will also monitor the kinetics of capsid formation. The results are interpreted in the context of CNT, obtaining a promising agreement with the theoretical predictions. Remarkably, this study provides interesting tools to understand and control the role of different biophysical parameters in the self-assembly process.

Part III of this thesis is dedicated to the analysis of the *Mechanical properties of viral capsids.* During the virus life cycle, the mechanical properties of its capsid play an important role in several biological processes. The viral capsid is indispensable to protect the genetic material against environmental changes and aggressions. Hence, its mechanical response is crucial not only in the viral life cycle, but also for different biotecnological applications. In this block, we will discuss the remarkable mechanical properties of viral capsids combining theory, simulations, and experiments.

In Chapter 8 we will analize *The importance of the structure in the buckling and maturation of spherical viruses.* Using a simple coarse-grained model, we will compare the energetics of spherical and icosahedral shapes and show that the preferred shape of a virus is strongly influenced by the icosahedral class of the viral structure. We will also discuss the connection between this morphological change and viral maturation, and the potential mechanical and biological advantages of the resulting capsids.

Chapter 9 (*Internal pressure in bacteriophage ϕ29*) investigates the mechanical properties of bacteriophage ϕ29, a prolate virus that packs dsDNA at high densities. The study was performed in collaboration with the experimental groups of P. J. de Pablo and J. L. Carrascosa. Their experiments, using atomic force microscopy, revealed that the presence of dsDNA packaged inside the capsid was translated in an increase of the effective stiffness of the virus. By using theoretical models and finite element simulations, we were able to interpret the results of the AFM experiments in terms of the effective pressure built-up by the confined dsDNA. We will describe how the combination of experiments and simulations provides a direct way to determine the genome pressure in bacteriophage ϕ29. This pressure is crucial for the infectivity of the virus.

In Chapter 10 (*Virtual Atomic Force Microscopy*) we will implement a "virtual AFM" using our coarse-grained model to mimic typical nanoindentation experiments. Using these simulations, we will be able to explain the anisotropic response in the different axis of symmetry found for bacteriophage T7 in AFM experiments. The influence of the different physical parameters on the mechanical response and the mechanisms of failure under large indentations will also be discussed.

Finally, in the last chapter (*Conclusions and Perspectives*) we will summarize the main conclusions of the thesis, emphasizing the biological implications of the work, and propose new questions and future works.

References

[1] M. Madigan and J. Martinko, Brock Biology of Microorganisms. Prentice Hall, London, *11th ed.*, 2005.

[2] S. J. Flint, L. W. Enquist, V. R. Racaniello, and A. M. Skalka, Principles of virology. ASM Press cop., Washington D.C., 2004.

[3] D. H. Crawford, The invisible enemy: a natural history of viruses. Oxford University Press, Oxford, 2000.

[4] A.J. Cann, Principles of Molecular Virology. Elsevier, Oxford, 2005.

[5] L. Sompayrac, How pathogenic viruses work, Jones and Bartlett Publishers, Toronto, 2002.

[6] W. Stanley, Isolation of a crystalline protein possessing the properties of tobaccomosaic virus, *Science*, **81**, 644-645, 1935.

[7] T. Shors, Understanding Viruses, Jones and Bartlett Publishers, Toronto, 2009.

[8] *http://commons.wikimedia.org/wiki/PapillomaVirus/HPV/EM.jpg.*

[9] RG. Webster, WJ. Bean, OT. Gorman, TM. Chambers and Y. Kawaoka. Evolution and ecology of influenza A viruses, *Microbiological Reviews*, **56**, 1, 152-179, 1992.

[10] A. Hershey and M. Chase, Independent functions of viral protein and nucleic acid in growth of bacteriophage, *J. Gen. Physiol.*, **36**, 39-56, 1952.

[11] NH. Mann, The Third Age of Phage, *PLoS Biol*, **3**, 5, 2005.

[12] *http://www.biochem.wisc.edu/faculty/inman/empics/virus.htm.*

[13] B. La Scola, C. Desnues, I. Pagnier, C. Robert, L. Barrassi, G. Fournous, M. Merchat, M. Suzan-Monti, P. Forterre, E. Koonin, and D. Raoult, The virophage as a unique parasite of the giant mimivirus, *Nature*, **455**, 100-104, 2008.

[14] MG. Mateu, Structure and Physics of Viruses, Springer, London, 2013.

[15] N.H. Acheson, Fundamentals of Molecular Virology, Wiley, Washington, D.C, 2011.

[16] M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, p. 30, 2000.

[17] M. G. Fischer and C. A. Suttle, A Virophage at the Origin of Large DNA Transposons, *Science*, **231**, 2011.

[18] P. Forterre, Defining life: the virus viewpoint, *Orig. Life Evol. Biosph.*, **40**, 151-160, 2010.

[19] D. Moreira and P. López-García, Ten reasons to exclude viruses from the tree of life, *Nat. Rev. Microbiol.*, **7**, 306-311, 2009.

[20] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Viral self-assembly as a thermodynamic process, *Phys. Rev. Lett.*, **90**, 248101, 2003.

[21] R. Zandi, P. van der Schoot, D. Reguera, W. Kegel, and H. Reiss, Classical nucleation theory of virus capsids, *Biophys. J.*, **90**, 1939-1948, 2006.

[22] A. Klug, The tobacco mosaic virus particle: structure and assembly, *Philos. Trans. R. Soc. London. Biol.*, **354**, 531-535, 1999.

[23] F. H. C. Crick and J. D. Watson, Structure of small viruses, *Nature*, **10**, 473-475, 1956.

[24] C. A. Tidona and G. Darai, eds., The Springer index of viruses. Springer-Verlag Berlin Heidelberg, 2002.

[25] M. F. Moody, The shape of the T-even bacteriophage head, *Virology*, **26**, 567-576, 1965.

[26] B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist, Assembly and analysis of conical models for the HIV-1 core, *Science*, **283**, 80-83, 1999.

[27] P. Ge, J. Tsao, S. Schein, TJ Green, M. Luo and ZH. Zhou, Cryo-EM Model of the Bullet-Shaped Vesicular Stomatitis Virus, *Science*, **327**, 5966, 689-693, 2010.

[28] D. Prangishvili, P. Forterre, and R. A. Garrett, Viruses of the Archaea: a unifying view, *Nat. Rev. Microbiol.*, **4**, 837-848, 2006.

[29] W. M. Gelbart and C. M. Knobler, Pressurized viruses, *Science*, **323**, 1682-1683, 2009.

[30] H. Lortat-Jacob, E. Chouin and S. Cusack *et al.*, Kinetic analysis of adenovirus fiber binding to its receptor reveals an avidity mechanism for trimeric receptor-ligand interactions, *J. Biol. Chem.*, **276**, 12, 9009-9015, 2001.

[31] S. Belouzard, JK. Millet, BN. Licitra *et al.*, Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein, *Viruses-Basel*, **4**, 6, 1011-1033, 2012.

[32] NGA. Abrescia, JJB. Cockburn and JM. Grimes, *et al.*, Insights into assembly from structural analysis of bacteriophage PRD1, *Nature*, **432**, 7013, 68-74, 2004.

[33] B. K. Ganser-Pornillos, M. Yeager, and W. I. Sundquist, The structural biology of HIV assembly, *Curr. Opin. Struc. Biol.*, **18**, 203-17, 2008.

[34] $http://sociedad.elpais.com/sociedad/2013/08/07/actualidad/1375886598-033585.html$

[35] $http://blogs.elpais.com/la-doctora-shora/2013/05/las-incognitas-del-coronavirus.html$

[36] RC. Mulligan, The basic science of gene therapy, *Science*, **260**, 926-932, 1993.

[37] S. Matsuzaki, M. Rashel and J. Uchiyama, Bacteriophage therapy: a revitalized therapy against bacterial infectious diseases, *J. Infect. Chemother* **11(5)** 211-9, 2005.

[38] LK. Miller, AJ. Lingg and A. Bulla, Bacterial, viral, and fungal insecticides, *Science*, **219** , 715-721, 1983.

[39] M. Comellas-Aragonés, *et al.*, A virus-based single-enzyme nanoreactor, *Nature Nanotechnology*, **2** 635-639, 2007.

[40] Lee SW, Mao CB, Flynn CE and Belcher AM., Ordering of quantum dots using genetically engineered viruses, *Science*, **296**, 892-895, 2002.

[41] Douglas T, Young M 1998. Host-guest encapsulation of materials by assembled virus protein cages, *Nature* **393** 152-155.

[42] SK. Dixit *et al.*, Quantum dot encapsulation in viral capsids, *Nano letters*, **6**, 1993-1999, 2006.

[43] A. E. Walsby, Gas vesicles, *Microbiol. Rev.*, **58**, 94-144, 1994.

[44] M. Sxener, J. Struumpfer, J. A. Timney, A. Freiberg, C. N. Hunter, and K. Schulten, Photosynthetic vesicle architecture and constraints on efficient energy harvesting, *Biophys. J.*, **99**, 67-75, 2010.

[45] J. Querol-Audi, A. Casanas, I. Uson, D. Luque, J. R. Caston, I. Fita, and N. Verdaguer, The mechanism of vault opening from the high resolution structure of the N-terminal repeats of MVP, *EMBO J.*, **28**, 3450-3457, 2009.

[46] CD. Eggleton and AS. Popel, Large deformation of red blood cell ghosts in a simple shear flow, *Physics of fluids*, **10**, 8, 1834-1845, 1998.

[47] J. Zhang, D. Li, G. Liu, K. J. Glover, and T. Liu, Lag periods during the self-assembly of macroions: connection to the virus capsid formation process, *J. Am. Chem. Soc.*, **131**, 15152-15159, 2009.

[48] Electron Microscopy Data Bank (EMDB) $http://www.ebi.ac.uk/pdbe/emdb/$.

Part I

Physical modeling of viruses

CHAPTER 1

Viral architecture

1.1 Introduction

An amazing feature of viruses is the fact that viruses of many different families, which infect widely different hosts (i.e. animals, plants, bacteria, etc) share a set of common architectures. Essentially, viruses can be classified in four different groups, according to their capsid shape (see Fig. 1.1): rod-like (like Tobacco Mosaic Virus [1, 2]), conical (e.g. Lentivirus like HIV [3]), bacilliform (e.g. Rice Tungro virus [4] or bacteriophage $\phi29$ [5]) and quasi-spherical (e.g. Herpes Simplex Virus [6] or Hepatitis B [7]). There are also a few viruses, most of them having a membrane envelope, that are pleimorphic and do not have well-defined shapes (like Halorubrum pleomorphic virus 1 [8] or smallpox [9]).

The simplest shape of these four groups is the rod-like shape. The coat proteins in these capsids are organized in a compact helical structure, with open ends and a small internal channel occupied by the genome. The length of these viruses is commonly controlled by the number of base pairs of the genetic material. The typical example, and the first virus discovered and observed using an electron microscope (EM) [10], is the Tobacco Mosaic Virus, see Fig. 1.1(a).

The conical shape is less common, but very important because it is the typically shape of the human immunodeficiency virus (HIV), see Fig. 1.1(b). One of the characteristics that difficult the study of these as well as of other related retroviruses is their facility to adopt different conical structures and irregular shapes [3, 11, 12]. This fact justify why it has not been possible to have a complete reconstruction of the capsid with atomic resolution yet [13, 14].

The bacilliform, prolate or elongated shells are characterized by a central cylindrical body closed by hemispherical caps, see Fig. 1.1(c). There are many viruses of this class showing also a great variety of sizes [15, 16].

The most common viral shape, adopted by more than 50% of all viruses infecting animals, plants, bacteria, and humans, are spherical viruses, see Fig. 1.1(d). These viruses are characterized by the presence of icosahedral symmetry in the arrangement of their coat proteins. There is a big variety of these viruses with sizes in a range between a few tens of *nm* to 600 *nm*. These viruses with icosahedral symmetry have been studied since the mid 1950s, when they were first proposed [17–19], but their origin and biological implications are still a subject of intense investigations [15, 20, 22–24].

The goal of this chapter is to review the main structural and geometrical properties of quasi-spherical and bacilliform viruses. These geometrical principles are the basis of the structural classification of viruses and will be very important for the remaining chapters. In addition, the well-defined size and shape of viral capsids make them ideal candidates for many promising applications as nanoreactors,

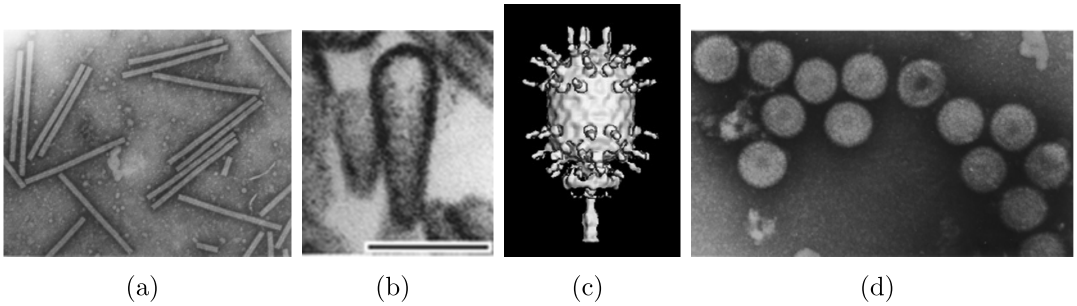

(a) (b) (c) (d)

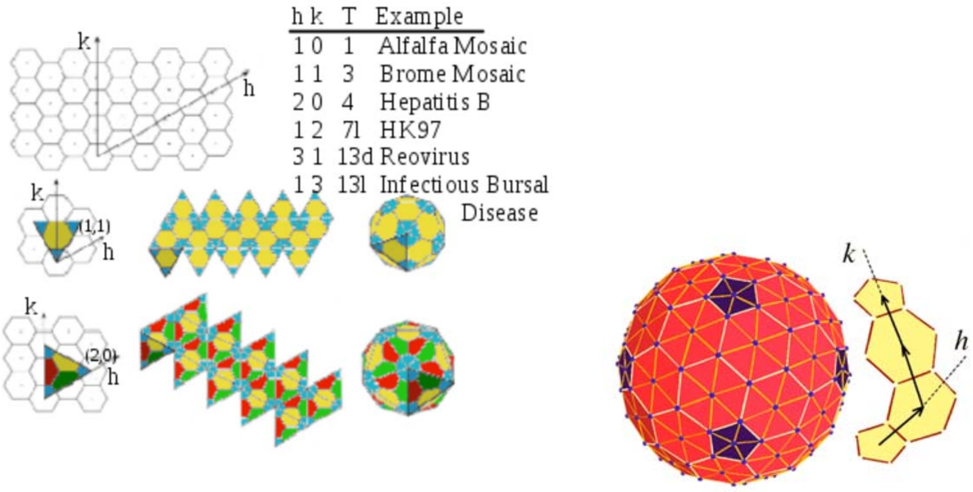
Figure 1.1: Electron microscopy images of a representative example of the different viral capsid shapes (a) rod-like shape: Tobacco Mosaic virus [2], (b) conical: HIV [3], (c) bacilliform: ϕ29 [5] and (d) spherical: Herpes Simplex [6].

nanocontainers or templates in areas such as medicine, nanosciences or pharmaceutical. Thus the importance of this geometrical description is that it provides the catalogue and properties of the structures that can be built for each shape in terms of a discrete number of values of length, radius and number the proteins. Understanding these selection rules is the first step towards controlling the dimensions and architecture of viral capsids.

This chapter will be distributed as follows. In Section 1.2 we review the Caspar and Klug construction for icosahedral viral capsids. Then, in Section 1.3, we highlight the concept of classes, introduced also by Caspar and Klug, to group triangulation numbers associated to similar subunits distribution. In Section 1.4 the recent generalization of Caspar and Klug ideas to describe bacilliform viruses will be described. After that, we will discuss the interesting physical origin of the prevalence of icosahedral symmetry in the capsid construction. Finally, the last two sections will be devoted to discuss some interesting exceptions (Section 1.5) and to summarize (Section 1.6) the main conclusions.

1.2 The Caspar and Klug construction and the T-number

Crick and Watson in 1956 were the first to suggest that small spherical viruses should adopt a regular capsid made of multiple copies of the same protein for reasons of genetic economy [17]. These identical subunits will tend to be arranged forming a regular structure, where all of them will have an equivalent environment. In fact, the two spherical viruses crystallized at that time, bushy stunt virus (BSV) [25, 26] and turnip yellow mosaic virus (TYMV) [27], had a diffraction patten characterized by cubic point group symmetry. The structures with this simmetry are related to the regular polyhedra [18] known as Platonic solids: tetrahedron, cube, octahedron, dodecahedron, and icosahedron. Of these, the icosahedron is the one that has an optimal surface/volume relation and can accommodate the largest number of proteins in equivalent positions, namely 60 proteins.



(a) (b)

Figure 1.2: Caspar and Klug model to build a spherical structure with icosahedral symmetry. (a) CK model starts with a hexagonal lattice to get the 20 equilateral triangles. (b) The vector (h, k) gives the number of steps along the principal directions on the hexagonal lattice between two adjacent pentamers of the viral capsid. Image taken from Ref. [28].

Few years later, in 1962, Caspar and Klug proposed a way to construct quasi-spherical structures with icosahedral symmetry [19]. But there are viruses with more than 60 proteins and for these larger capsids, the distribution coat proteins cannot be strictly regular, so it was necessary to introduce quasi-equivalent environments.

Our modern understanding of the geometry and the distribution of proteins in quasi-spherical viral shells is based on the Caspar and Klug (CK) model [19]. This model describes the geometrical rules of construction of a spherical virus with icosahedral symmetry.

The way to make the different quasi-spherical viral capsids with the Caspar and Klug model is to use a flat hexagonal lattice of protein subunits as starting point, see Fig. 1.2(a). To make a closed shell is then necessary to introduce at least twelve defects in the hexagonal lattice according to Euler's theorem [29]. In the CK construction, the twelve defects are regularly spaced at the vertices of the resulting structure which has icosahedral symmetry. In the final structure, there are two different morphological units or capsomers, called *pentamers* and *hexamers*, corresponding to clusters of five or six proteins, respectively (see Fig.1.3). The different possible structures are characterized by the vector $C_T = h\overrightarrow{a_1} + k\overrightarrow{a_2} \equiv (h, k)$,

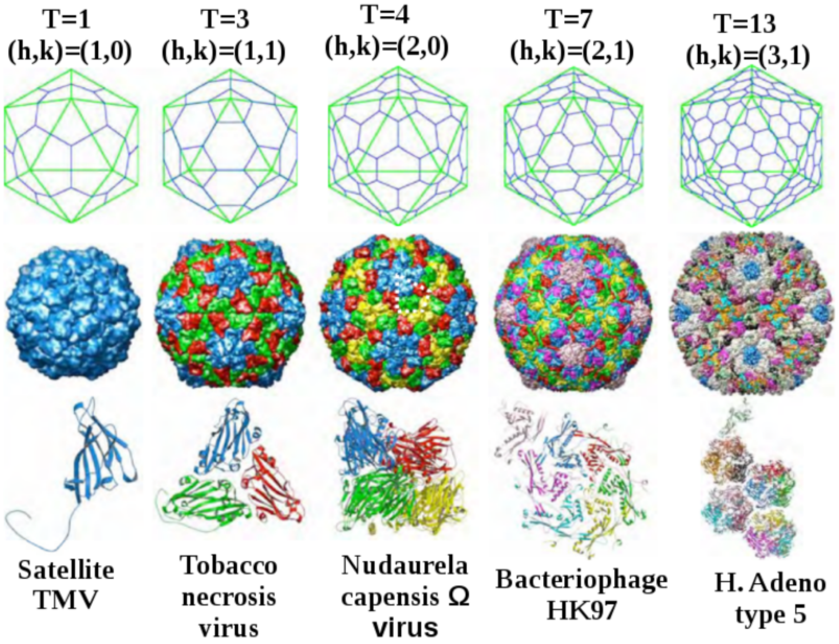


Figure 1.3: Caspar and Klug model to build some examples of quasi-spherical viral capsids for the smallest T-numbers. The top row shows the geometry of the shell, the middle row is a detailed structural image of the capsid and the botton row represents the structure of the coat basic subunits . Figure taken from Ref. [21].

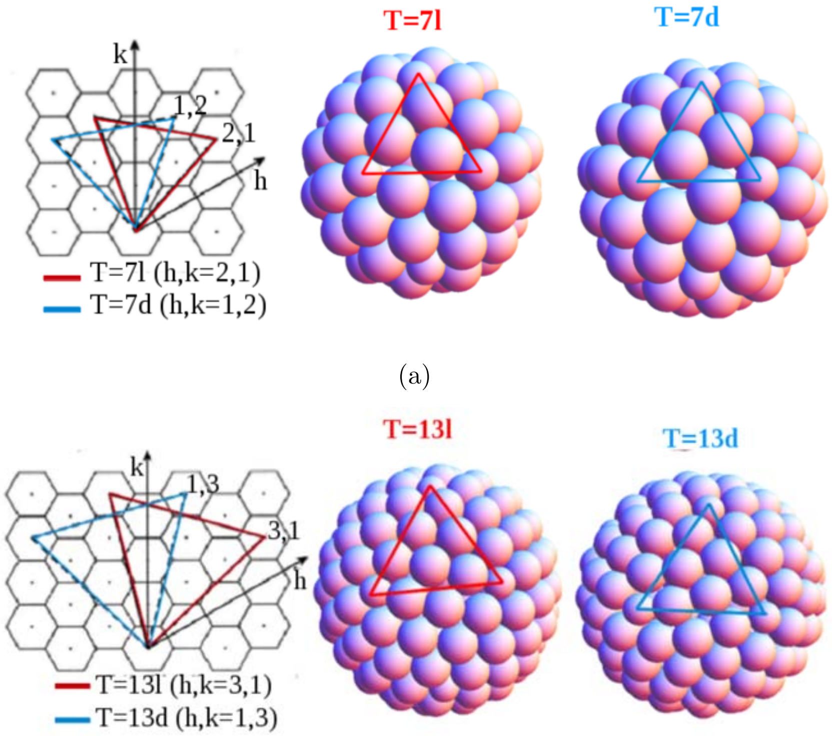
where (h, k) are non-negative integers that give the number of steps to connect two nearest pentamers along the principal directions $\overrightarrow{a_1}$ and $\overrightarrow{a_2}$ of the hexagonal lattice, see Fig. 1.2(b).

The vector $\overrightarrow{C_T}$ defines an equivalent triangle in the hexagonal lattice whose area is given by the triangulation number that characterizes the different structures.

$$T = h^2 + hk + k^2 \tag{1.1}$$

Then, the final closed shell is constructed using 20 of these equilateral triangles forming an icosahedron characterized by the presence of twelve pentamers at its vertices. Since h and k are integer numbers, there is a limited number of ways to create a closed shell with icosahedral symmetry corresponding to the discrete sequence of magic numbers $T = 1, 4, 7, 9, 12, ...$, see Fig. 1.3.

The resulting T-number capsid has invariably 12 pentamers, $N_H = 10(T - 1)$ hexamers, adding up to a total of $N = 10T + 2$ capsomers and $N_{sub} = 60T$ proteins. Moreover, the final capsid has 15 2-fold, 10 3-fold and 6 5-fold rotation axes as an icosahedron, which are the defining traits of icosahedral symmetry.



(a)

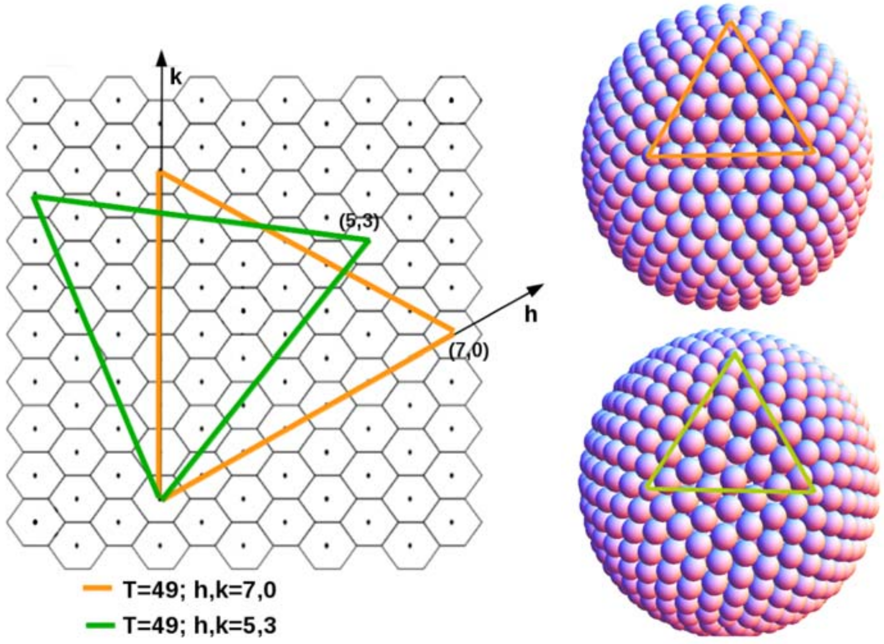
(b)

Figure 1.4: Construction of the equilateral triangle in the laevo direction (red one) and for the dextro direction (blue triangle), and the resulting laevo and dextro structures for $T = 7$ (top row) and $T = 13$ (bottom row).

The CK construction is the basis to understand and classify the structure of spherical viruses [22], but in some cases the triangulation number does not describe univocally the structure of the shell. For instance, (h, k) and (k, h) build equilateral triangles of the same size, but are different vectors where the orientation of the triangle over the lattice and the distribution of subunits are different. For example the vectors $(h, k) = (3, 1)$ and $(h, k) = (1, 3)$ correspond to the same $T = 13$ number (see Fig. 1.4). When $h \neq h$, it is possible to distinguish between two different structures having the same T-numbers, called *laevo* (when $h > k$) and *dextro* (when $h < k$). In addition, some T-number correspond to more than one structure, such as $T = 49$ that can be obtained from $(h, k) = (7, 0)$ and $(h, k) = (5, 3)$, see Fig. 1.5.

1.3 The P-class

In the previous section, we have introduced the CK construction to describe the geometry of quasi-spherical viral capsids using a T-number or univocally using (h, k). In this section we will emphasize a reorganization of the T-number in



(a) (b)

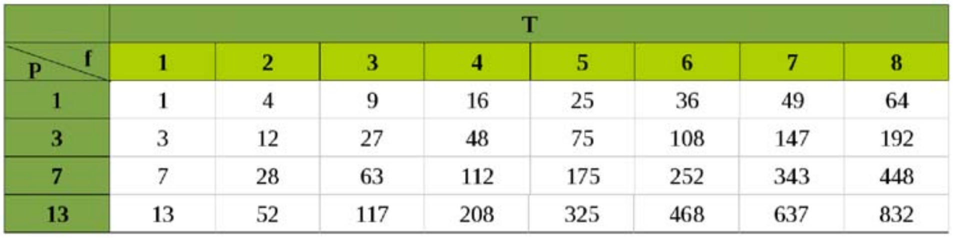
Figure 1.5: (a) Construction of the basic equilateral triangle for $T = 49$, with $(h, k) = (7, 0)$ (orange triangle) and $(h, k) = (5, 3)$ (green triangle). (b) Image of the resulting structures.

classes, also proposed by Caspar and Klug [19]. These classes are described by a second number, the P-number, that groups structures of the CK model having a similar arrangement of subunits. The letter P used to label each class has not to be confused with the pseudotriangulation number used to describe quasi-spherical capsids whose subunits are chemically different.

The P-class is defined in terms of the T-number (see eq. 1.1) as

$$T = h_0^2 f^2 + h_0 k_0 f^2 + k_0^2 f^2 = P f^2 \tag{1.2}$$

where $P = h_0^2 + h_0 k_0 + k_0^2$ and f is the greatest common divisor of h and k. The first P-classes, as well as the smallest T-numbers in the class, are listed in Table1.6. Structures from the class $P = 1$ correspond to shells with $(h_0, 0)$ or $(0, k_0)$, and are characterized by having a straight line of hexamers at the edges of the triangles connecting two consecutive pentamers, see Fig. 1.7. On the other hand, $P = 3$ structures have the pentamers connected by a zig-zag line of hexamers (see Fig. 1.7). All classes with $P > 3$ correspond to skewed structures. The importance of the reorganization of T-numbers into classes is that structures belonging to the same class are expected to have similar mechanical and physical properties, given their common arrangement of capsomers.



	T							
P \ f	1	2	3	4	5	6	7	8
1	1	4	9	16	25	36	49	64
3	3	12	27	48	75	108	147	192
7	7	28	63	112	175	252	343	448
13	13	52	117	208	325	468	637	832

Table 1.1 T-numbers for the first four P classes (rows) classified by the value of f (columns), see Eq. 1.2, up to $T_{max} = 117$ [32].

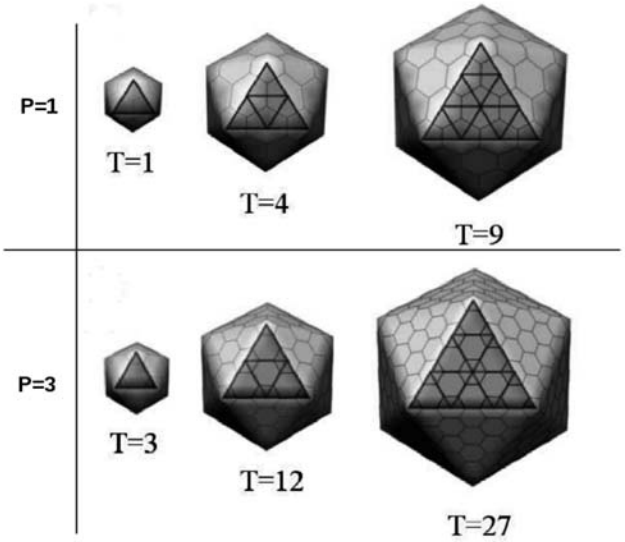


Figure 1.7: Smaller T-numbers for the $P = 1$ and $P = 3$ classes. Figure taken from Ref. [30].

This reorganization of the T-number in classes provides useful information about the geometrical distribution of subunits in the capsid. But as with the T-number, the classification of viral structures using only the class P is not unique, since different combinations of h_0 and k_0 can yield the same value of P, e.g., $P = 91$ can be obtained by two different pairs of (h_0, k_0), $(6, 5)$ and $(9, 1)$.

1.4 Generalization of the CK construction to classify prolate viruses

As mentioned at the introduction, there are many viruses with an elongated closed shell. In this section we will describe the geometrical way to characterize the elongated structures using an extension of the Caspar and Klug model [30, 31]. These elongated or prolate structures are assumed to have a hexagonally ordered cylindrical body closed by two hemispherical caps with icosahedral symmetry. Hence, they can be considered as the result of elongating a CK spherical capsid along one of its possible axis of symmetry (5-fold, 3-fold or 2-fold). This idea was first proposed by Moody in the sixties to describe elongated viruses with 5-fold axial symmetry [33, 34].

Moody's model for elongated structures with 5-fold axial symmetry is based on defining two types of triangles (see Fig1.8(a)). There are 5 identical triangular faces characterized by $T_{end}(h,k) = h^2 + hk + k^2 = f^2P$ in each cap, as in the CK model, plus 10 triangular faces in the cylindrical body, characterized by a second triangulation number [33]

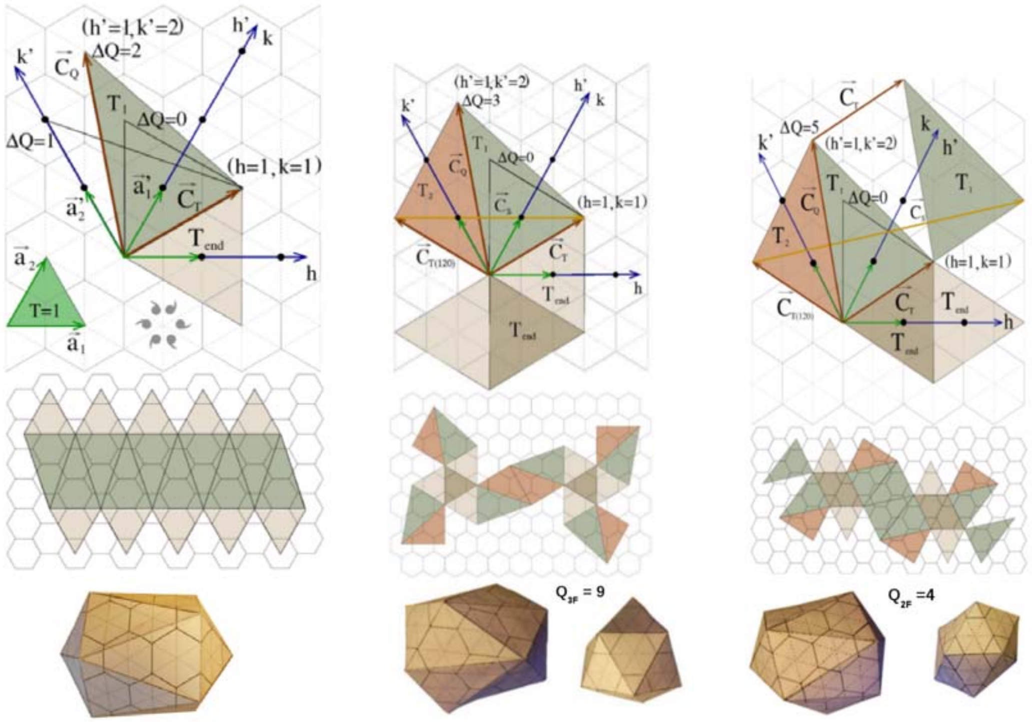
$$\begin{aligned} T_{mid} &= hh' + hk' + kk' = fQ \\ Q &= h_0h' + h_0k' + k_0k' \end{aligned} \tag{1.3}$$

This second triangulation number is defined by the integers (h,k) and (h',k') and controls the length of the cylindrical body. The few elongated viruses whose structures has been reconstructed by cryo-EM so far follow this constructed and correspond to $T_{end} = 3$ and $Q = 5$ for bacteriophage ϕ 29 [35] and $T_{end} = 13l$ and $Q = 20$ for $T4$ [36]. However, there are other prolate viruses such as Alfalfa Mosaic Virus (3-fold axis) [37–39] or Cocoa Swollen Shoot Virus (2-fold axis) [40], which 5-fold is not the axial symmetry and Moody's model cannot apply. For these viruses it is possible to extend the CK model getting general rules for the construction of elongated viruses with 2-fold, 3-fold and 5-fold axial symmetries [31, 41].

The case with 3-fold axial symmetry is more complicated than the 5-fold, because one needs two different middle triangles for the cylindrical body, instead of just one. Each cap is made by 4 equilateral triangles characterized by (h,k) and T_{end}, whereas the body is defined by a second set of integers (h',k'), (see green and red triangles in Fig. 1.8(b)).

Similarly, to build the prolate capsid with 2-fold axial symmetry, the caps are made by two equilateral triangles defined by the vector (h,k). The body is determined by three non-equivalent triangles defined by (h',k') as it is illustrated in Fig. 1.8(c). The details of the geometrical construction are given in Ref. [30].

But the important result is that in general the number of subunits N_{sub}, the number of capsomer N, the radius R, and length L of the prolate structures are discretized and determined by the icosahedral cap symmetry. The discretization rule for the minimum length step $\triangle L^{min}$, the minimum step in Q, $\triangle Q^{min}$, and the minimum step in number of capsomers $\triangle N^{min}$, and the possible values of N_{sub}, N, and R are summarized below for each symmetry [31, 41]:



(a) (b) (c)

Figure 1.8: Basic elements required to build a prolate capsid with the three different axial symmetries for an elongated virus based on a $T = 3$ with a) 5-fold symmetry, as $\phi 29$, b) 3-fold symmetry, and c) 2-fold axial symmetry. The triangles of the cap are defined by $(h, k) = 1, 1$ (brown region), and the triangles for the body by $(h', k') = 1, 2$ (green region). The top image shows the basic triangles; the middle image, the complete design and the bottom image, the resulting shell. Images adapted from Ref. [30].

$$5-fold \begin{cases} N = 5(T_{end} + T_{mid}) + 2 \\ N_{sub} = 30(T_{end} + T_{mid}) \\ Q = h_0(h' + k') + k_0 k' \\ \triangle N^{min} = 5f \\ R/a = \frac{5}{2\pi}\sqrt{T_{end}} \\ \triangle L^{min}/a = \frac{\sqrt{3}}{2\sqrt{P}} \\ \triangle Q^{min} = 1 \end{cases} \qquad 3-fold \begin{cases} N = 4T_{end} + 3Q + 2 \\ N_{sub} = 6(4T_{end} + 3Q) \\ Q = h_0(2h' + k') + k_0(2k' + h') \\ \triangle N^{min} = 3f \triangle Q^{min} \\ R/a = \frac{3}{2\pi}\sqrt{3T_{end}} \\ \triangle L^{min}/a = \frac{\sqrt{\triangle Q^{min}}}{2\sqrt{P}} \\ \triangle Q^{min} = \begin{cases} 3 & if|h_0 - k_0| \propto 3 \\ 1 & otherwise \end{cases} \end{cases}$$

$$2-fold\begin{cases} N = 4T_{end} + 2Q + 2 \\ N_{sub} = 12(2T_{end} + Q) \\ Q = h_0(3h' + 2k') + k_0(3k' + h') \\ \triangle N^{min} = 2f \triangle Q^{min} \\ R/a = \frac{2}{2\pi}\sqrt{7T_{end}} \\ \triangle L^{min}/a = \frac{\sqrt{21}}{14}\frac{\triangle Q^{min}}{\sqrt{P}} \\ \triangle Q^{min} = \begin{cases} 7 & if|h_0 - 2k_0| \propto 7 \\ 1 & otherwise \end{cases} \end{cases} \tag{1.4}$$

Interestingly, in all cases the possible values of these viral capsid properties are determined by the class $P = h_0^2 + h_0k_0 + k_0^2$ through $T_{end} = f^2P$ or, alternatively, by (h_0, k_0).

1.5 The origin of icosahedral symmetry in viral capsids

The ultimate reason for the election and predominance of icosahedral symmetry in quasi-spherical and bacilliform structures, is due to a combination of genetic, geometrical, and physical arguments.

Viral genomes are relativity short, with lengths between 2 and 50 *kbp* confined at high densities inside the capsids [42]. Given the limited coding space of viral genome, the best strategy to make the capsid is to use multiple copies of a limited number of small proteins rather than encoding for a large coat protein. This principle of genome economy implies that viral shells have been build only with one or a few different types of proteins [17]. This identical coat proteins would prefer to assemble into a structure where all of them see an equivalent environment. The best way to maximize the packing and the number of contacts using a few number of different proteins is to organize them into a hexagonal lattice. Using as stating point this hexagonal lattice, it is then possible to build all different capsids shapes by simple geometrical transformations.

By Euler's theorem [29], it is necessary to introduce 12 defects on the original hexagonal lattice to construct closed shells [29]. In the simplest case of spherical capsids, the 12 defects are evenly distributed forming a quasi-spherical (or polyhedral) shape with icosahedral symmetry. The advantage of this shape is that it provides the maximum volume to enclose the genome. Another option is wrapping the lattice into an helical tube with either open ends (like in rod-like viruses) or by closing each of its ends with a semi-spherical cap, containing six defects, to make a bacilliform capsid [18, 40]. The conical shape corresponds to the case where the twelve defects are not located homogeneously in the caps, e.g. four in one cap and eight in the other [3].

The ultimate reasons for the prevalence of icosahedral symmetry in quasi-spherical and bacilliform structures are physical. The beautiful and regular structures adopted by viral capsids are just a consequence of minimization of the free energy of the interactions between the different subunits. This has been shown to be indeed the origin of icosahedral symmetry in spherical [23] and prolate [41] viruses. But in these previously studies, structures with icosahedral symmetry where found to be the optimal arrangement of subunits when a template or constraint on the shape was imposed.

In this thesis we will go beyond that constrained minimization of the energy, and we will develop simple coarse-grained models to study the self-assembly process. We will try to find what physical ingredients play an important role in the selection of a particular structure and what are the mechanisms that regulate the kinetics of the assembly.

1.6 Discussion

In this section we will make some clarifications regarding the structure of viruses.

Most spherical viruses comply with the CK model, but there are some notable exceptions [43]. For example, some viruses like polyoma and papilloma viruses self-assemble. Using exclusively pentamers, rather than pentamers and hexamers, as in the CK construction. The resulting native structures also have icosahedral symmetry and correspond to a $T = 7$ shell made only by pentamers. Interestingly, this arrangement has been found to be an energy minimum for model viruses made by a single type of structural units, and it has been also justified in geometric terms by using a tiling approach [44, 45, 49]. Remarkably, these viruses can also self-assemble *in vitro* into a spherical structure without icosahedral symmetry, known as snub cube [47, 48].

Another important observation is that, although pentamers and hexamers are clear morphological structures in the CK model, they are not always the elementary units of assembly. Some viruses are built from the assembly of individual proteins, dimers, or trimers. In fact, some dsRNA viruses like the Bluetongue virus (BTV) core [46] assemble into a shell with 120 proteins (which would correspond to a forbidden $T = 2$ capsid), which is equivalent to a $T = 1$ shell having dimers instead of individual proteins. In this context, it is worth emphasizing that the CK model determines the group of symmetry, not the clustering of the proteins. Thus capsid built by $20T$ trimers, $30T$ dimers, or $60T$ monomers, or other combinations are completely valid.

In this thesis, we will focus on the assembly of capsids made by capsomers (either all-pentamers or by pentamers and hexamers) as assembly units.

1.7 Conclusions

In this chapter, we have reviewed the geometrical construction and classification of spherical and elongated viral shells. Using the ideas of CK [19], Moody [33, 34] or its recent generalization in Ref. [31, 41] it is possible to describe geometrically quasi-spherical and bacilliform viruses with icosahedral symmetry. For both shapes, the possible sizes and number of capsomers of the viral shell have a discrete number of values characterized by the T-number or the P-class. In the case of the prolate shape, the rules for the possible discrete values of length and number of subunits are also determined by the axial symmetry and the T-number of the cap.

Interestingly, a few viruses, like Hepatitis B [7], are polymorphic and are present *in vivo* forming two different capsids shapes. Moreover, prolate and spherical viruses can be reconstructed and assembled *in vitro* into sizes and shapes different from the native one, as in the cases of Polyoma [47, 48] or CCMV [50]. Accordingly, in principle, it should be possible to control the selection of prolate and spherical structures by using the proper assembly conditions and concentrations. Hence, a best understanding of the geometrical and physical principles of the viral capsids could be potentially very helpful to control the outcome of *in vitro* experiments for promising applications in nanosciences.

For that reason, in the following chapters we will study the role of different physical ingredients in the self-assembly of quasi-spherical and elongated capsids, using a simple model of interaction between capsomers.

References

[1] H. Fraenkel-Conrat and R. C. Williams, Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components, *Proc. Natl. Acad. Sci. USA.*, **41**, pp. 690-698, 1955.

[2] W. K. Kegel and P. van der Schoot, Physical regulation of the self-assembly of tobacco mosaic virus coat protein, *Biophys. J.*, **91**, 1501-12, 2006.

[3] B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist, Assembly and analysis of conical models for the HIV-1 core, *Science*, **283**, 80-3, 1999.

[4] MC, Jones, K. Gough and I. Dasgupta, Rice Tungro disease is caused by an RNA and a DNA virus, *J. Gen. Virol.*, **72**, 757-761, 1991.

[5] WR. Wikoff and JE. Johnson, Virus assembly: Imaging a molecular machine, *Current Biology*, **9**, 8, R296R300, 1999.

[6] K. Grunewald, P. Desai and DC Winkler, *et al.* Three-dimensional structure of herpes simplex virus from cryo-electron tomography, *Science*, **302**, 5649, 1396-1398, 2003.

[7] RA. Crowther, NA. Kiselev, B. Bottcher, JA. Berriman, GP. Borisova, V. Ose and P. Pumpens, 3-Dimensional structure of Hepatitis-B virus core particles determined by electron cryomicroscopy, *Cell*, **77**, 6, 943-950, 1994.

[8] MK. Pietila, S. Laurinavicius and J. Sund, The Single-Stranded DNA Genome of Novel Archaeal Virus Halorubrum Pleomorphic Virus 1 Is Enclosed in the Envelope Decorated with Glycoprotein Spikes, *J. Virol.*, **84**, 2, 788-798, 2010.

[9] GW. Long, J. Noble, Jr. FA. Murphy, KL. Herrmann and Bernard Lourie, Experience with Electron Microscopy in the Differential Diagnosis of Smallpox, *Appl Microbiol.*, **20**, 3, 497-504, 1970.

[10] S. J. Flint, L.W. Enquist, V. R. Racaniello, and A. M. Skalka, Principles of virology. ASM Press cop., Washington, D.C., 2004.

[11] J. B. Heymann, C. Butan, D. C.Winkler, R. C. Craven, and A. C. Steven, Irregular and semi-regular polyhedral models for Rous sarcoma virus cores, *Comput. Math. Methods. Med.*, **9**, 197-210, 2008.

[12] C. Butan, D. C. Winkler, J. B. Heymann, R. C. Craven, and A. C. Steven, RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis, *J. Mol. Biol.*, **376**, 1168-1181, 2008.

[13] B. K. Ganser-Pornillos, M. Yeager, and W. I. Sundquist, The structural biology of HIV assembly, *Curr. Opin. Struc. Biol.*, **18**, 203-17, 2008.

[14] O. Pornillos, B. K. Ganser-Pornillos, and M. Yeager, Atomic-level modelling of the HIV capsid, *Nature*, **469**, 424-427, 2011.

[15] ICTVdB, The Universal Virus Database, version 4. http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/index.htm.

[16] C. A. Tidona and G. Darai, eds., The Springer index of viruses. Springer-Verlag Berlin Heidelberg, 2002.

[17] F. H. C. Crick and J. D. Watson, Structure of small viruses, *Nature*, **10**, 473-475, 1956.

[18] R. W. Horne and P. Wildy, Symmetry in virus architecture, *Virology*, **15**, 348-73, 1961.

[19] D. L. D. Caspar and A. Klug, Physical principles in the construction of regular viruses, in Cold Spring Harbor Symp. Quant. Biol., **27**, 1-24, Cold Spring Harbor Laboratory Press, 1962.

[20] J. E. Johnson and J. A. Speir, Quasi-equivalent viruses: a paradigm for protein assemblies, *J. Mol. Biol.*, **269**, 665-75, 1997.

[21] Viral Particle ExploreR2 *http://viperdb.scripps.edu/index.php.*

[22] T. S. Baker, N. H. Olson, and S. D. Fuller, Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs, *Microbiol. Mol. Biol. Rev.*, **63**, 862-922, 1999.

[23] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, Origin of icosahedral symmetry in viruses, *Proc. Natl. Acad. Sci. USA.*, **101**, 15556-15560, 2004.

[24] D. H. Bamford, J. M. Grimes, and D. I. Stuart, What does structure tell us about virus evolution?, *Curr. Opin. Struc. Biol.*, **15**, 655-63, 2005.

[25] J. D. Bernal, I. Fankuchen, and D. P. Riley, Structure of the crystals of tomato bushy stunt virus preparations, *Nature*, **142**, p. 1075, 1938.

[26] C. H. Carlisle and K. Dornberger, Some x-ray measurements on single crystals of tomato bushy-stunt virus, *Acta Crystallogr.*, **1**, 194-196, 1948.

[27] J. D. Bernal and C. H. Carlisle, Unit cell measurements of wet and dry crystalline turnip yellow mosaic virus, *Nature*, **162**, 139-140, 1948.

[28] A. Siber, Buckling transition in icosahedral shells subjected to volume conservation constraint and pressure: Relations to virus maturation, *Phys. Rev. E*, **73**, p. 061915, 2006.

[29] H. S. M. Coxeter, Introduction to geometry. Wiley, 2nd ed., 1989.

[30] Luque A., Thesis. Structure, Mechanical Properties, and Self-Assembly of Viral Capsids, 2011.

[31] A. Luque and D. Reguera, The structure of elongated viral capsids, *Biophys. J.*, **98**, 2993-3003, 2010.

[32] M. Aznar, A. Luque, and D. Reguera, Relevance of capsid structure in the buckling and maturation of spherical viruses, *Phys. Biol.*, **9**, 1478-3975, 2012.

[33] M. F. Moody, Geometry of phage head construction, *J. Mol. Biol.*, **293**, 401-33, 1999.

[34] M. F. Moody, The shape of the T-even bacteriophage head, *Virology*, **26**, 567-76, 1965.

[35] Y. Tao, N. H. Olson, W. Xu, D. L. Anderson, M. G. Rossmann, and T. S. Baker, Assembly of a tailed bacterial virus and its genome release studied in three dimensions, *Cell*, **95**, 431-7, 1998.

[36] A. Fokine, P. R. Chipman, P. G. Leiman, V. V. Mesyanzhinov, V. B. Rao, and M. G. Rossmann, Molecular architecture of the prolate head of bacteriophage T4, *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6003-8, 2004.

[37] R. A. Heijtink, C. J. Houwing, and E. M. Jaspars, Molecular weights of particles and RNAs of alfalfa mosaic virus. Number of subunits in protein capsids, *Biochemistry*, **16**, 4684-93, 1977.

[38] S. Cusack, G. T. Oostergetel, R. C. J. Krijgsman, and J. E. Mellema, Structure of the topat component of alfalfa mosaic virus (a non-icosahedral virion), *J. Mol. Biol.*, **171**, 139-155, 1983.

[39] A. Kumar, V. S. Reddy, V. Yusibov, P. R. Chipman, Y. Hata, I. Fita, K. Fukuyama, M. G. Rossmann, L. S. Loesch-Fries, T. S. Baker, and J. E. Johnson, The structure of alfalfa mosaic virus capsid protein assembled as a $T = 1$ icosahedral particle at 4.0-A resolution, *J. Virol.*, **71**, 7911-7916, 1997.

[40] R. Hull, The structure of tubular viruses, *Adv. Virus. Res.*, **20**, 1-32,1976.

[41] A. Luque, R. Zandi, and D. Reguera, Optimal architectures of elongated viruses, *Proc. Nat. Acad. Sci. USA* **107**, 53235328, 2010.

[42] W.C. Earnshaw, S.C. Harrison, DNA arrangement in isometric phage heads, *Nature*, **268**, 598, 1977.

[43] E. L. Altschuler and A. Pérez-Garrido, Do all spherical viruses have icosahedral symmetry?, *arXiv*, **0902**, 3566v, 1-4, 2009.

[44] R. Twarock, A tiling approach to virus capsid assembly explaining a structural puzzle in virology, *J. Theor. Bio.*, **226**, 477-82, 2004.

[45] R. Twarock, Mathematical models for tubular structures in the family of papovaviridae, *Bull. Math. Biol.*, **67**, 973-987, 2005.

[46] J. M. Grimes, J. N. Burroughs, P. Gouet, J. M. Diprose, R. Malby, S. Zientara, P. P. Mertens, and D. I. Stuart, The atomic structure of the bluetongue virus core, *Nature*, **395**, 470-8, 1998.

[47] A. Klug and J.T Finch, Structure of viruses of the papilloma-polyoma type. *J. Mol. Biol.*, **11**, 403423, 1965.

[48] N. A. Kiselev and A. Klug, The structure of viruses of papilloma-polyoma type (V. tubular variants built of pentamers), *J. Mol. Biol.*, **40**, 155-171, 1969.

[49] T. Keef, A. Taormina, and R. Twarock, Classification of capped tubular viral particles in the family of Papovaviridae, *J. Phys.: Cond. Matt.*, **18**, S375-S387, 2006.

[50] J.B. Bancroft, G.J. Hills and R. Markham, A study of the self-assembly process in a small spherical virus: formation of organized structures from protein subunits in vitro, *Virology*, **31**, 354-379. 1967.

CHAPTER 2

Theoretical description of viral assembly kinetics

2.1 Introduction

The self-assembly of the capsid is one of the most important steps in the viral life-cycle. After the entrance of the genetic material into the host cell, the replication starts and the virus takes advantage of the cell machinery to produce the required components for a new virus, i.e., coat proteins, genetic material, lipids and other auxiliary proteins. These components then self-assemble forming a new viral particle.

Contrarily to the architecture of the final structures, described in the previous chapter, the microscopic mechanisms and the kinetics of the process leading to the final capsid are not well known. Understanding better the role of the different steps during viral assembly is an open question, indispensable for the control of virus infection [1, 2].

As mentioned in the introduction, viruses infect a huge variety of cells and the assembly process could be totally different for diverse viruses. In some cases the capsid is successfully assembled from the coat proteins alone [3–7], in combination with scaffolding proteins [8, 9], or in combination with the genetic material [10], see Fig.2.1. The proteins of most single-stranded viruses are expected to co-assemble in the presence of genetic material (Fig.2.1 (C)), whereas double stranded viruses typically first assemble an empty capsid that is later on packaged with the genetic material [11, 12] (Fig.2.1 (A)). In some cases, condensing agents and/or auxiliary proteins (including scaffolding proteins) are required for a successful assembly *in vivo* of an infective virus (Fig.2.1 (B)).

The pathways during the assembly process are also virus specific. For example, CCMV uses as building blocks stable dimers present in solution, that are assumed to form a pentamer of dimers structure as the nucleus for the assembly [34]. HBV uses dimers as building blocks too, but this virus forms a trimer of dimers as intermediate in the assembly, in most cases, into $T = 4$ capsids and in a small number of cases forming smaller $T = 3$ particles, [28]. Other interesting cases are Polyomavirus that use exclusively pentamers as assembly units [5].

However, despite the specificity of the assembly of different viruses, there are some common characteristics which open the door to a theoretical description of capsid assembly kinetics. The most important one is that the assembly is a spontaneous process that does not require the addition of energy (for example from ATP hydrolysis) [13, 14]. This fact indicates that the assembly of viruses must be controlled by general and basic physical principles [13–17].

Remarkably, the proteins of many viruses have the capability of assemble spontaneously *in vitro* even in the absence of their genetic material, forming empty capsids whose size and structure can be tuned by the assembly conditions. Thus in

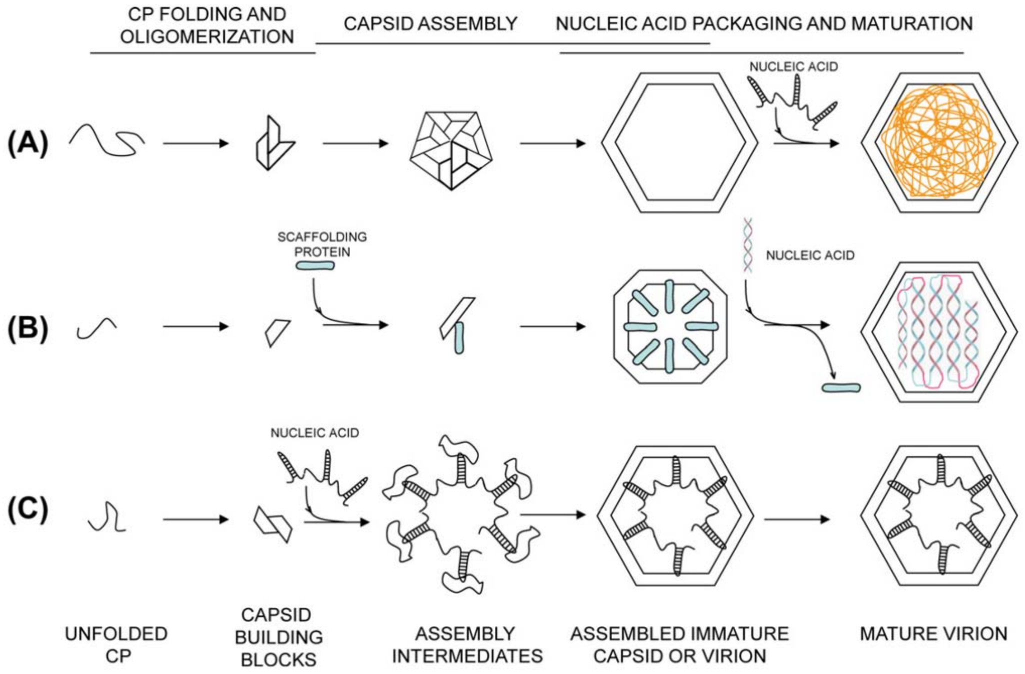


Figure 2.1: Schemes of three strategies for assembly of viral capsids. (A) Unassisted self-assembly followed by the packaging of the nucleic acid. (B) Scaffolding protein-assisted assembly. (C) Viral nucleic acid-assisted assembly. Figure by M.A. Fuertes taken from Ref. [18].

many cases, by changing the protein concentrations, the pH, or the ionic strength conditions, the same protein is able to self-assemble into different capsid structures [3, 5, 19–22]. These experiments show the important role of electrostatics, kinetics, and thermodynamic in the assembly of viral capsids and open the door to the control of size and shapes of viral shells for nanosciences [23].

In this chapter we will review the main ideas and theoretical approaches used to describe the kinetics of viral assembly. These ideas will provide us with a theoretical framework to interpret the self-assembly simulations described in Chapter 7. For simplicity, we will focus on the kinetics of assembly of empty capsids made exclusively by coat proteins (i.e. without genetic material and auxiliary proteins).

In this theoretical description, we will use in general the term *subunits*, to refer to the elementary building blocks in the assembly. It is important to note that the theoretical description and general results are valid for different kinds of subunits. For instance, the elementary building blocks could be single proteins, as in the case of penicilium chrysogenum virus (PcV) [24], dimers for CCMV [3], pentamers in polyomavirus [5] or pentamers and hexamers in solution for bacteriophage HK97 [25].

This chapter is divided as follows. In Section 2.2 we summarize the common features in the assembly of different viruses that support this theoretical description of the assembly. Then, in section 2.3, a master equation description for the kinetics of viral assembly will be introduced. An alternative description, based on classical nucleation theory (CNT), will be presented in Section 2.4. Finally, Section 2.5 summarizes the most remarkable conclusions.

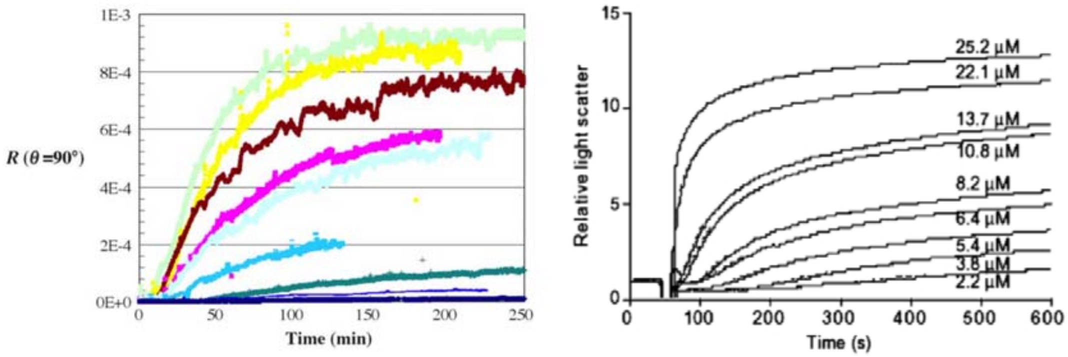
2.2 Common features of viral assembly

In the past decades, many experiments have shown that fully infective viruses can be reconstructed *in vitro*. In recent years, it has also been possible to monitor in *in vitro* experiments the kinetics of formation of empty capsids using coat proteins from different viruses like HPV [7, 26, 27], HBV [6, 28], CCMV [3], or BMV [4, 29, 30].

In these assembly experiments, one typically starts from coat proteins in solution and use light scattering techniques, turbidity, size-exclusion chromatography (SEC), transmission electron microscopy (TEM) or even mass spectrometry [31], to monitor the formation of complete capsids. Typically, the amount of capsids formed is monitored by the previously mentioned techniques versus the time for different initial proteins concentrations. For instance, Fig.2.2(a) shows the intensity of the light scattering signal as a function of time for the assembly of HPV starting at different protein concentrations. From the intensity of the signal it is possible to evaluate the total number of capsids formed and their rate of formation, as a function of temperature, pH, salt conditions and initial protein concentration. As another example, Fig.2.2(b)shows the mass fraction of proteins in capsids as a function of the initial capsid protein concentration for the *in vitro* assembly of HBV.

The results of these experiments show common characteristics in the viral assembly. In all viral assembly processes there is a strong dependence of capsid production on protein concentration. In general, if the protein concentration is too low, the assembly is not possible. Hence, there exists a minimum critical concentration required to trigger the assembly. On the other hand, if the protein concentration becomes too large, many structures are trapped into intermediate steps and the self-assembly process does not yield complete capsids, but rather aberrant structures. Another general observation is the poor number of intermediate structures observed in the experiments. There are either free subunits in solution or fully formed capsids, but not a significant concentration of partial structures.

Interestingly, from the kinetic point of view, there are also common trends



(a) (b)

Figure 2.2: Intensity of the light scattering signal of assembly reactions versus time with various initial protein concentrations for (a) Papillomavirus between 20 and 220 $\mu g/ml$ [26] and (b) for Hepatitis B from $2.2\mu M$ to $25.2\mu M$. Images taken from Refs. [2] and [26].

in the assembly of different viruses. The assembly dynamics is characterized by a sigmoidal kinetics and the existence of a lag time before the formation of the first shells, as shown in Fig.2.2. Both characteristic are typical in nucleation processes. Furthermore, there is a hysteresis effect between the assembly and the disassembly processes. This pronounced hysteresis is essential to resist the changes in environmental conditions and to prevent disassembly in the viral journey from host to host.

The steep dependence on concentration, the scarce of intermediates, the sigmoidal kinetics, the existence of a time lag, and the hysteresis between assembly and dissassembly are characteristic of a first order phase transition. Accordingly, different theoretical descriptions have been developed to describe this process. The main ones will be discussed in the following two sections.

2.3 Master equation approach to viral assembly kinetics

One of the first approaches, and the most widely used, to describe the kinetics of empty capsid formation was presented by Zlotnick and collaborators [32–36]. The basic idea is that the viral assembly can be described by a reaction cascade model as supported by *in vitro* experiments on Hepatitis B virus [6, 28] and Papilloma viruses [7] in combination with their assembly models [15, 32, 33] and simulations [34, 35].

In their approach, the assembly is conceived as a sort of chemical reaction and described in terms of the population of clusters of different sizes. The viral self-assembly is described as the step-by-step sequential addition of identical individual subunits to intermediate structures characterized by their number of subunits. In principle, there is a large number of assembly pathways, coming from different possible combinations to add a subunit to a partial capsid. The most probable pathway involves the intermediate steps that have the lowest free energy for one particular adding subunit [35, 36].

The time-evolution of viral self-assembly can be modeled exactly by a finite number of states at any given time, where switching between states is treated probabilistically. Usually, the equations for the system are a set of differential equations for the variation over time of the probabilities that the system occupies each different state. In the determination of the assembly kinetics of viral shells, the intermediate steps of these reactions are very important. In principle, the viral self-assembly could grow by different mechanisms, adding a single subunit, multiple independent subunits simultaneously, or even by merging different partial capsids. However, theoretical models and simulations show that in most of cases only one subunit is added at a time [37, 38]. With this assumption, the assembly can be described by a series of unimolecular reactions, involving the attachment or detachment of only one unit in each time step, namely

$$c_n(t) \rightleftarrows c_{n+1}(t) \tag{2.1}$$

where $c_n(t)$ is the concentration of partial capsids made of n subunits at time t. Considering that $\beta_n(t)$ and $\alpha_n(t)$ are the associated binding and unbinding rates, respectively, the change in time of the population, or concentration, of partial capsids of size n is controlled by the following master equation:

$$\frac{\partial c_n(t)}{\partial t} = \beta_{n-1}(t)c_{n-1}(t) - \alpha_n(t)c_n(t) - \beta_n(t)c_n(t) + \alpha_{n+1}(t)c_{n+1}(t) \tag{2.2}$$

The combination of intermediates of size n thus increases by partial capsids of size $n-1$, which grow by one subunit (the first term of eq.2.2) and by intermediates of size $n+1$, which loss one subunit (the last term of eq.2.2); and decrease by partial capsids of size n, which loss (the second term of eq.2.2) or gain (the third term of eq.2.2) one subunit.

The kinetic model by Zlotnick and collaborators describes two step in the assembly process: a nucleation step, which accounts for the formation of initial nuclei embryos, and an elongation process, when the nuclei growth by the sequential addition of subunits. This nucleus is considered as a special intermediate which cannot disassemble and the time required for its formation is the initial lag time of viral assembly. The population of intermediates structures are given by a set of master equations, where the rates of association β and dissociation α of subunits are calculated using the association constant for a single intersubunit contact and statistical factors, which account for the degeneracy of the reactions. This reaction cascade assembly model has been successfully used to describe the experimental results for the *in vitro* assembly of HBV [2, 6, 28] and papilloma [7].

However solving the complete set of the master-equations for all intermediates in the initial assembly could be a difficult task in some cases, since it involves many undetermined constants and requires several approximations. Moreover, in the cases that can be solved, the solution is more descriptive than predictive and does not provide any clear insight of the influence of the parameters relevant in the kinetics. To study the kinetics of viral self-assembly it is often more convenient to use an alternative description as described in the next section.

2.4 Classical Nucleation Theory applied to viral assembly

In recent years, a description of viral assembly kinetics based on Classical Nucleation Theory (CNT) [16, 40] has been developed with the aim of overcoming the limitations of the master equation approach, trying to provide quantitative predictions which could guide experiments. In this section, we will summarize the main formalism of CNT that will be later used in Chapter 7 to interpret our simulations.

2.4.1 Viral assembly as a phase transition

The formation of a capsid can be though as a transition between two phases of the subunits: free in solution and in a fully-formed capsid. When the protein concentration is low, the most favorable state for the protein is remaining in solution. As we increase concentration, the energetic gain of being in a capsid compensates the entropic penalty of leaving the solution, and the stable state becomes the fully-formed capsid.

The energetics and the conditions at which each phase becomes stable can be described in simple thermodynamic terms by the total Gibbs free energy of the system:

$$G(N_1, N_2 ... N_q, p, T) = \sum_{n=1}^{q} \mu_n N_n \tag{2.3}$$

where n is the number of subunits in a (partial) capsid that also acts as the label for each species, N_n is the number of aggregates of size n and $\mu_n = \left(\frac{\partial G}{\partial N_n}\right)_{N_j \neq n, pT}$ is the corresponding chemical potential, which is just the variation of free energy when a particle of a certain species is added to the system.

The chemical potential for relatively dilute solutions, as it is often the case in viral self-assembly experiments, can be written as:

$$\mu_n = \mu_n^0 + k_B T \ln(c_n/c_s) \tag{2.4}$$

where k_B is Boltzmann's constant, T is the temperature, and c_n is the concentration of the partial structure of size n. It is worth mentioning that the standard part $\mu_n^0(p,T)$ implicitly depends on the reference state given by c_s. By minimizing eq.2.3 it is possible to obtain the equilibrium distribution of aggregates, given by the familiar condition of equal chemical potentials, i.e: $\mu_1^{eq} = \frac{\mu_n^{eq}}{n}$. Using eq.2.4 for the chemical potential, the result is the *law of mass action*

$$\frac{c_n^{eq}}{(c_1^{eq})^n} = \frac{e^{-\Delta W(n)/k_B T}}{(c_s)^{n-1}} \equiv K_n^{eq} \tag{2.5}$$

where c_n^{eq} and c_1^{eq} are the equilibrium concentration of the partial structure of size n and the concentration of subunits in solution, respectively, K_n^{eq} is the equilibrium constant, and $\Delta W(n) \equiv \mu_n^0 - n\mu_1^0$ is the standard free energy difference required to form a cluster of n subunits.

The law of mass action can be rewritten as a standard equilibrium Boltzmann's equation as

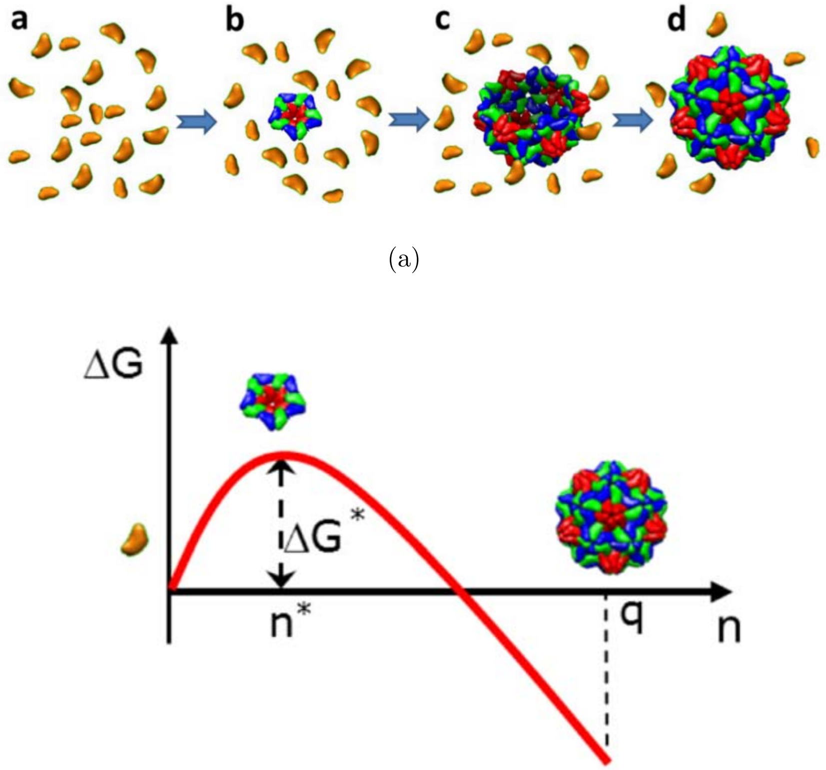
$$c_n^{eq} = c_s e^{-\Delta G(n)/k_B T} \tag{2.6}$$

where $\Delta G(n) = \Delta W(n) - nk_B T \ln(c_1/c_s)$ is the full free energy of capsid formation [40]. For a complete capsid of size q, $\Delta W(q) = \mu_q^0 - q\mu_1^0 \equiv q\Delta g$, where Δg is the effective binding energy per subunit. But there is an energy penalty (an energy barrier) to overcome, associated to the formation of partial capsids, which is described in the following section.

2.4.2 Free energy barrier for capsid formation

The assembly of the capsid is driven by the gain in free energy associated to Δg, the effective binding energy per subunit. But during the process of capsid formation there is also an energy barrier coming from the fact that the subunits at the rim of a partial capsid miss some of the contacts present in the final structure, see Fig. 2.3. In the context of CNT, the simplest way to model this energetic cost is by introducing a rim energy penalty associated to a line tension λ. Accordingly, the free energy of formation of a partial capsid containing n subunits can be modeled as:

$$\triangle G(n) = n \triangle g + \lambda l(n) - nk_B T \ln(c_1/c_s) \tag{2.7}$$



(a)

(b)

Figure 2.3: Cartoon of capsid assembly, illustrating the different steps in the process. The top line shows the assembly starts from free subunits in solution a. Thermal fluctuations generate a large enough critical cluster b that then grows forming a partial capsid c that eventually gets completed into a full capsid. The process is characterized by a free energy of formation schematically despited in the botton line. Fig. taken from Ref. [16].

The first term represents the contribution from the binding energy $\triangle g$ per subunit. This binding energy will be taken as constant, even though subunits in a capsid can occupy different local environments, or be engaged in different interactions. Surprisingly, this crude approximation works nicely [40, 41]. The second term is the total line energy for a partial capsid, given by the product of the line tension λ (i.e. the energetic cost per unit length) times $l(n)$, the length of the rim, that for a spherical capsid can be approximated as $l(n) = b\sqrt{n(q-n)}$, where $b = 4\pi\frac{R}{q}$, see Fig. 2.7. Here q is the total number of subunits in a complete formed capsid and R is its radius. This line tension term is always positive, has its maximum when the capsid is half-formed, corresponding to $n = q/2$, and it is zero for a complete capsid. The third term is the entropic penalty for removing the n free subunits from the solution into the partial capsid, which depends on the concentration of free subunits c_1 and the reference state, c_s. This free energy can be more conveniently rewritten as:

$$\triangle G(n) = n\Delta\mu + a\sqrt{n(q-n)} \tag{2.8}$$

where $\Delta\mu = -k_B T ln(c_1/c^*)$ and $a = 4\pi\lambda\frac{R}{q}$. The term c^* represents the saturation concentration at which a subunit will have the the same free energy in a capsid or free in solution.

Note that the free energy of capsid formation $\triangle G(n)$, eq.2.8, has a term proportional to n, that favors the assembly, and an energetic cost associated to the line tension proportional to $\sqrt{n(q-n)}$. The competition between these contributions bring about an energy barrier, which is decisive in the kinetic of viral shell formation.

The height of the barrier ($\triangle G^*$) and critical size (n^*) corresponding to the maximum of the barrier, illustrated in Fig.2.3 are calculated from the condition $\left.\frac{\partial G(n)}{\partial n}\right|_{n^*} = 0$, yielding:

$$\begin{aligned} \triangle G^* &= \frac{q}{2}\left(\sqrt{\triangle\mu^2 + a^2} + \triangle\mu\right) \\ n^* &= \frac{q}{2}\left(1 + \frac{\triangle\mu}{\sqrt{\triangle\mu^2 + a^2}}\right) \end{aligned} \tag{2.9}$$

For a hexagonally ordered structure the line tension can be estimated as $\lambda = \frac{|\triangle g|}{3\sqrt{3}\sigma_0}$, as described in Ref. [41], where σ_0 is the radius of the subunit. Using this approximation,

$$a = \frac{4\pi R|\triangle g|}{3\sqrt{3}q\sigma_0} \tag{2.10}$$

It is important to emphasize that both ΔG^* and n^* depend on the concentration of subunits c_1 through $\Delta\mu$. Fig. 2.5 shows the shape of $\Delta G(n)$ for a $T = 7$ capsid with $q = 72$ subunits and for different values of $\Delta\mu$ taking $\Delta g = -15k_B T$. When $\Delta\mu > 0$, the formation of a capsid is unfavorable. When $\Delta\mu = 0$, there is a significant barrier to overcome for the formation of a capsid. In this particular case, the critical size becomes $n^* = q/2$, corresponding to half a capsid, and the nucleation barrier would be $\Delta G^* = q/2a = 2\pi R\lambda$. On the other hand, if c_1 is large enough, $\Delta\mu < 0$ and the formation of a capsid becomes feasible.

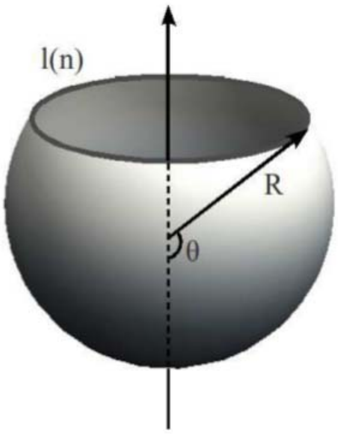


Figure 2.4: Continuum representation of a partial capsid of radius R. The angle θ characterizes its degree of completion. The circular rim has a length $l(n)$ that depends on the number of subunits n in the partial shell, and it has an associated line tension λ. Figure taken from Ref. [16].

Even in this case, if $n < n^*$ the free energy of formation grows as we add a subunit, so intermediates of these sizes will tend to disassemble back into free subunits. And when $n > n^*$, the capsid reduces their energy upon the addition of another subunit so it will tend to grow spontaneously until closing the shell.

2.4.3 Kinetics of viral assembly and disassembly

In the context of CNT, the steady state rate of formation of capsids J, which is the number of capsids formed per unit volume and time, is given by [40]:

$$J(t) = c_s \beta^* Z e^{-\Delta G^*/k_B T} \tag{2.11}$$

where β^* is the rate of attachment of subunits to the critical nucleus n^*, and $Z = \sqrt{\frac{|\triangle G''(n^*)|}{2\pi k_B T}}$ is the Zeldovich factor, which is the local curvature at the top of the barrier. The formation of capsids is thus exponentially controlled by $\triangle G^*$. In the case where the binding energy is fixed, as occurs in our simulation model, the main factor to lower the barrier is the increase in protein concentration, see eqs. 2.3 and 2.8. For a fixed subunit concentration, we will use later on this CNT model to estimate the energies and critical sizes required to form a complete capsid as well as the rate of its formation.

The thermodynamic and kinetic theory introduced in this section is also able to account for capsid disassembly, interpreting the dissociation of a shell as the inverse process of capsid formation. The dissociation of capsids requires to jump the free energy barrier in the opposite direction of assembly, and the formalism to describe the disassembly is analogous. For the specific case of $\triangle\mu = 0$, the energy

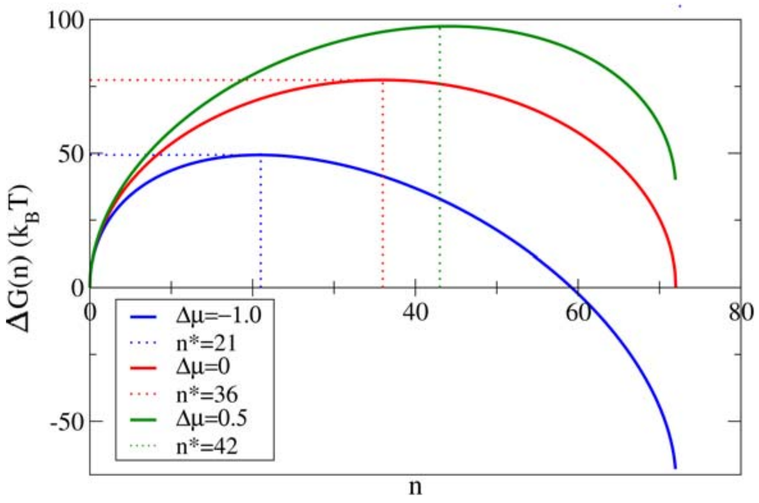


Figure 2.5: Free energy of capsid formation $\triangle G(n)$, eq. 2.8, versus the number of subunits n for different values of $\triangle\mu$. The values of the barrier height ($\triangle G^*$) and the critical size n^* are calculated using eq. 2.10.

barrier is the same for viral assembly and disassembly, but in general, these two process are different, and occur at different protein concentration, which is the ultimate reason for the hysteresis found in experiments.

2.4.4 Nucleation theorem and the size of the critical nucleus

In the context of nucleation, there is a very useful relation between the size of the critical cluster, the nucleation barrier, and the nucleation rate, known as the "nucleation theorem". This nucleation theorem was originally proposed in 1962 by Hill in the context of small systems thermodynamics [42, 43], and then applied to experiments by Kaschiev [44]. The nucleation theorem establishes that:

$$\left(\frac{\partial \triangle G^*}{\partial \triangle \mu}\right)_{T,p} = n^* \tag{2.12}$$

In view of eq. 2.12 and eq. 2.11, this theorem can be translated into a more useful relation

$$\frac{\partial \ln J}{\partial \ln c_1} = n^* + 1, \tag{2.13}$$

that establishes a direct route to extract the critical size n^* directly from the experimental measures of the capsid formation rate at different concentrations. Being a general thermodynamic result its importance lies on the fact that it establishes a relation between the nucleation rate and the critial cluster size that is independent of the specific model.

This relation will be helpful to check kinetic and thermodynamic theories against simulation and experiments. Also, this could help to understand better the role of physical and biological parameters in the construction of viral shells.

2.5 Conclusions

In this chapter we have presented the main theoretical descriptions of the kinetics of viral capsid assembled. The process of formation of a capsid can be considered as a sort of chemical reaction or phase transition between disassembly and fully assembled state of subunits, and described accordingly using a set of master equations or the standard formalism of CNT.

Both theoretical descriptions, which have been illustrated in this chapter, are connected and related, and constitute a useful framework to interpret and guide experiments and simulations [45–52]. Experimentally the theory could be helpful to design experiments at conditions that will ensure the efficient production of capsids at reasonable time scales. In particular, CNT expressions can also be used to predict, at least qualitatively, the thermodynamics and kinetics of the assembly of specific viruses.

Computer simulations constitute a useful alternative to study the general kinetics and thermodynamics involved in viral assembly. With simulations it is possible to study in a simple way the intermediate steps of the process, and they could provide the required information about the mechanisms of assembly, and the main characteristics of the critical cluster.

In Chapter 7 we will describe Brownian Dynamics simulations of a simple coarse-grained model designed to mimic viral self-assembly experiments, starting from free subunits in solution. In that chapter we will use the predictions of CNT to study the conditions required to obtain fully-formed capsids of different sizes and shapes. With the theoretical predictions we can also estimate the nucleation barrier associated to the formation of a partial shell, and how it depends on the different physical parameter of the simulation.

By estimating these free energy barriers it is possible to infer the optimal conditions for assembly for the different T-number empty shells. But it is important to emphasize that the theoretical results of CNT, could also be used in other simulations and experiments and this is one of our future works.

References

[1] P. E. Prevelige, Inhibiting virus-capsid assembly by altering the polymerisation pathway, *Trends Biotechnol.*, **16**, 61-65, 1998.

[2] A. Zlotnick and S. J. Stray, How does your virus grow? Understanding and interfering with virus assembly, *Trends Biotechnol.*, **21**, 536-542, 2003.

[3] J. B. Bancroft, C. E. Bracker, and G. W. Wagner, Structures derived from cowpea chlorotic mottle and brome mosaic virus protein, *Virology*, **38**, 324-335, 1969.

[4] C. Chen, C. C. Kao, and B. Dragnea, Self-assembly of brome mosaic virus capsids: insights from shorter time-scale experiments, *J. Phys. Chem. A*, **112**, 9405-12, 2008.

[5] D. M. Salunke, D. L. D. Caspar, and R. L. Garcea, Polymorphism in the assembly of polyomavirus capsid protein VP1, *Biophys. J.*, **56**, 887-900, 1989.

[6] P. Ceres and A. Zlotnick, Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids, *Bioch.*, **41**, 11525-11531, 2002.

[7] S. Mukherjee, M. V. Thorsteinsson, L. B. Johnston, P. A. DePhillips, and A. Zlotnick, A quantitative description of in vitro assembly of human papillomavirus 16 virus-like particles, *J. Mol. Biol.*, **381**, 229-37, 2008.

[8] P. E. Prevelige, D. Thomas, and J. King, Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells, *Biophys. J.*, **64**, 824-35, 1993.

[9] 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**, 1097-106, 2006.

[10] H. Fraenkel-Conrat and RC. Williams, Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components, *PNAS*, **41**, 690-694, 1955.

[11] S. J. Flint, L. W. Enquist, V. R. Racaniello, and A. M. Skalka, Principles of virology. Washington, ASM Press cop., D.C., 2004. Chapter 13.

[12] E. Kutter, Bacteriophage: molecular biology and applications. CRC Press, Boca Raton, 2005.

[13] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Viral self-assembly as a thermodynamic process, *Phys. Rev. Lett.*, **90**, 248101, 2003.

[14] S. Katen and A. Zlotnick, The thermodynamics of virus capsid assembly, **455**, ch. 14, 395-417. Elsevier Inc., 1 ed., 2009.

[15] D. Endres and A. Zlotnick, Model-based analysis of assembly kinetics for virus capsids or other spherical polymers, *Biophys. J.*, **83**, 1217-30, 2002.

[16] R. Zandi, P. van der Schoot, D. Reguera, W. Kegel, and H. Reiss, Classical nucleation theory of virus capsids, *Biophys. J.*, **90**, 1939-1948, 2006.

[17] M. F. Hagan and O. M. Elrad, Understanding the concentration dependence of viral capsid assembly kinetics the origin of the lag time and identifying the critical nucleus size, *Biophys. J.*, **98**, 1065-74, 2010.

[18] MG. Mateu, Structure and Physics of Viruses, Springer, London, 2013.

[19] K.W. Adolph and P.J.G. Butler, Studies on the assembly of a spherical plant virus (I. States of aggregation of the isolated protein), *J. Mol. Biol.*, **88**, 327-338, 1974.

[20] B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist, Assembly and analysis of conical models for the HIV-1 core, *Science*, **283**, 80-3, 1999.

[21] L. Lavelle, M. Gingery, M. Phillips, W. M. Gelbart, C. M. Knobler, R. D. CadenaNava, J. R. Vega-Acosta, L. A. Pinedo-Torres, and J. Ruiz-Garcia, Phase diagram of self-assembled viral capsid protein polymorphs, *J. Phys. Chem. B*, 2009.

[22] S. Kanesashi, K. Ishizu, M. Kawano, S. Han, S. Tomita, H. Watanabe, K. Kataoka, and H. Handa, Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro, *J. Gen. Virol.*, **84**, 1899-1905, 2003.

[23] T. Douglas and M. Young, Viruses: making friends with old foes, *Science*, **312**, 873-875, 2006.

[24] J. R. Castón, S. A. Ghabrial, D. Jiang, G. Rivas, C. Alfonso, R. Roca, D. Luque, and J. L. Carrascosa, Three-dimensional structure of penicillium chrysogenum virus: A double-stranded RNA virus with a genuine $T = 1$ capsid, *J. Mol. Biol.*, **331**, 417-431, 2003.

[25] Z. Xie and R. W. Hendrix, Assembly in vitro of bacteriophage HK97 proheads, *J. Mol. Biol.*, **253**, 74-85, 1995.

[26] G. L. Casini, D. Graham, D. Heine, R. L. Garcea, and D. T. Wu, In vitro papillomavirus capsid assembly analyzed by light scattering, *Virology*, **325**, 320-7, 2004.

[27] M. S. Kumar and R. Schwartz, A parameter estimation technique for stochastic self-assembly systems and its application to human papillomavirus self-assembly, *Phys. Biol.*, **7**, 045005, 2010.

[28] S. Singh and A. Zlotnick, Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly, *J. Biol. Chem.*, **278**, 18249 -18255, 2003.

[29] M. Cuillel, C. Berthet-Colominas, B. Krop, A. Tardieu, P. Vachette, and B. Jacrot, Self-assembly of brome mosaic virus capsids. Kinetic study using neutron and X-ray solution scattering, *J. Mol. Biol.*, **164**, 645-50, 1983.

[30] C. Berthet-Colominas, M. Cuillel, M. H. J. Koch, P. Vachette, and B. Jacrot, Kinetic study of the self-assembly of brome mosaic virus capsid, *Eur. Biophys. J.*, **15**, 159-168, 1987.

[31] YH. Tseng, C. Uetrecht, AJR. Heck, and WP. Peng, Interpreting the Charge State Assignment in Electrospray Mass Spectra of Bioparticles, *Anal. Chem.*, **83**, 6, 1960-1968, 2011.

[32] A. Zlotnick, J. M. Johnson, P. W. Wingfield, S. J. Stahl, and D. Endres, A theoretical model successfully identifies features of hepatitis B virus capsid assembly, *Biochem.*, **38**, 14644-14652, 1999.

[33] A. Zlotnick, Theoretical aspects of virus capsid assembly, *J. of Mol. Recogn.*, **18**, 479-490, 2005.

[34] 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 science*, **14**, 1518-1525, 2005.

[35] P. Moisant, H. Neeman, and A. Zlotnick, Exploring the paths of (virus) assembly, *Biophys. J.*, **99**, 1350-7, 2010.

[36] A. Zlotnick, To build a virus capsid an equilibrium-model of the self-assembly of polyhedral protein complexes, *J. Mol. Bio.*, **241**, 1, 59-67, 1994.

[37] D. C. Rapaport, Modeling capsid self-assembly: design and analysis, *Phys. Biol.*, **7**, 045001, 2010.

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

[39] T. Keef *et al.*, An equilibrium assembly model applied to Murine Polyomavirus, *J. Th. Med.*, **6**, 2, 91-93, 2005.

[40] A. Luque, Thesis. Structure, Mechanical Properties, and Self-Assembly of Viral Capsids, 2011.

[41] A. Luque, D. Reguera, A. Morozov, J. Rudnick and R. Bruinsma, Physics of shell assembly: Line tension, hole implosion, and closure catastrophe, *J. Chem. Phys.*, **136**, 18, 184507, 2012.

[42] T. L. Hill, Thermodynamics of small systems, *J. Chem. Phys.*, **36**, 3182-3197, 1962.

[43] T. L. Hill, Thermodynamics of small systems: Parts I and II. W. A. Benjamin Inc., New York, 1963 and 1964.

[44] D. Kashchiev, On the relation between nucleation work, nucleus size, and nucleation rate, *J. Chem. Phys.*, **76**, 5098-5102, 1982.

[45] D. C. Rapaport, J. E. Johnson, and J. Skolnick, Supramolecular self-assembly: molecular dynamics modeling of polyhedral shell formation, *Comput. Phys. Commun.*, **121**, 231, 1999.

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

[47] D. Rapaport, Role of reversibility in viral capsid growth: A paradigm for self-assembly, *Phys. Rev. Lett.*, **101**, 1-4, 2008.

[48] R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger, Local rules simulation of the kinetics of virus capsid self-assembly, *Biophys. J.*, **75**, 2626-36, 1998.

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

[50] M. F. Hagan, Controlling viral capsid assembly with templating, *Phys. Rev. E*, **77**, 051904, 2008.

[51] O. M. Elrad and M. F. Hagan, Encapsulation of a polymer by an icosahedral virus, *Phys. Biol.*, **7**, 045003, 2010.

[52] 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.

CHAPTER 3

Coarse-grained modeling of capsids

3.1 Introduction

Viral self-assembly is a complicated and virus-specific process, which however shows some universal features. The most remarkable one is that, despite the differences in shape, size, sequence, and conformation of coat proteins among different viruses, they end up adopting a common set of architectures. That suggest the existence of common ingredients that underlay the assembly [1] and open the door to the formation of generic models to describe it.

Modelization of capsid structures and assembly can be done at different resolution levels. In principle, one could study the coat proteins of viral shells with atomic detail, where every atom is explicitly represented at electronic level, and the Schrödinger equation solved and integrated in time. These atomic level models are widely used in the description of small molecules, but become inappropriate and prohibitive for complex system as viruses, where thousands to millions of complex molecules in an aqueous environment are involved. For such large systems, quantum effects are expected to be not so important, so an alternative will be to use atomistic models with classical potentials. Although recently it has been possible to perform the first all-atom simulations of small viruses [2], realistic simulations of the complete assembly process are still out of reach, since only the smallest viruses and at very short time scales (on the order of ps to ns) can be studied. Moreover, atomistic details are often unnecessary to get the relevant mesoscopic and macroscopic properties for most systems [3].

An alternative to atomistic models are coarse-grained models. The philosophy of coarse-grained models is to disregard the non-relevant details of the structure and the interaction by grouping atoms and molecules into new effective structural units that lack atomistic detail. This approximation describes often very well the relevant macroscopic behavior of the system under study, using a minimum set of essential ingredients and degrees of freedom. Thus, these models are simple, computationally inexpensive and provide invaluable information about the basic physical mechanisms often inaccessible by other means.

That is the reason why most studies on viral assembly use coarse-grained models with different levels of resolution that range from a purely geometric description in terms of tiles to a sistematic coarse-graining based on the atomic structures obtained by X-ray diffraction or cryo-electron microscopy. The lowest-resolution models typically use capsomers as subunits (see Fig. 3.1) whose interactions are described by local rules, or simple pair-potentials. More refined models use rigid assemblies of spheres and point like interactions building effective triangles, trapezoidal or truncated piramidal objects, that resemble the proteins and dimers that are the essential assembly subunits of many capsids.

In most of these models, the interactions have multiple sites whose location,

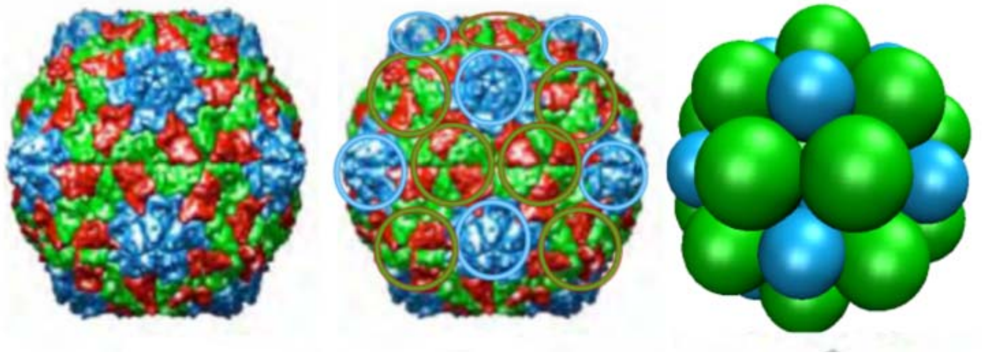


Figure 3.1: Schematic cartoon of a coarse-grained approximation for Tobacco Necrosis Virus. Clusters of 5 (pentamers) or of 6 (hexamers) proteins in the atomic structure (left) are coarse-grained and replaced by effective spheres (right image).

strenght, and directionality have to be hand-crafted in a delicate way to reproduce just one specific target structure. For this reason, it is no clear what are the essential mechanisms and general ingredients of the interaction which control the selection of a particular size or structure.

In our work, rather than resorting to complicated or elaborated models, we have rather opted for a sort of top-down strategy, where the goal is to look for the essential ingredients and simplest models capable of describing the relevant mechanisms, and add progressively the details required to achieve a more realistic description.

As seen in Chapter 2, in many viruses, capsid proteins aggregate in structural units, called capsomers, which for some viruses are indeed the elementary units of the assembly. In line with this minimalist spirit, we will focus on self-assembly models for viruses using capsomers as assembly subunits, whose interactions are expected to be simpler and more generic than protein-protein interactions.

In this chapter, we will describe the main coarse-grained models used for the thesis. We will start, in Section 3.2, by briefly summarizing some previous models, that were successful in describing equilibrium aspects of viral structure, and that will be the basis of the new model developed in this thesis and described in Section 3.3. Finally, Section 3.4 highlights the most relevant aspects of this new model, which will be used in the next chapters.

3.2 Previous coarse-grained models

Theoretical and computational models are becoming very helpful for the study of the pathways and mechanisms of assembly. Many models have been developed in recent years to study viral structure and assembly. These models range from simple continuum models to sophisticated coarse-grained strategies based on the atomic structures (see Fig. 3.2).

As a first approximation, a viral capsid can be modeled as a homogeneous spherical shell, and use accordingly continuum elasticity theory [4] to describe the energy required in its formation, see Fig. 3.2(a). This continuum approximation, which might seem very crude at first sight, is inded reasonable because viruses are composed by thousands of atoms. For most viruses, the radius of the shell is much larger than its thickness. Hence capsids are commonly described using a simplified limit of the elastic theory known as thin shell theory [5].

As we described in Chapter 1, the role of geometry in viral shell is very important. In fact some authors have used these ideas to propose models purely based on geometry [6, 7, 10–12]. For instance, some models use capsomers as subunits whose interactions are described by local rules [6]. Alternately, viral capsids can be described as a spherical or cylindrical surface covered using a regular set of tiles of different shapes as a generalization of the triangular tiles of the CK model [7] (see Chapter 1, and Figs. 3.2(b) and (c)).

Adding a bit more details, the next level comprises coarse-grained models. Coarse-grained models usually represent the capsid using highly simplified structural units that interact through simple potentials with two major contributions: an excluded volume term to prevent the overlapping of subunits and some kind of attraction to drive the assembly. Capsids with the right icosahedral symmetry are recovered using suitable simulation techniques as the structures that minimize the free energy. Various coarse-grained models have been proposed in the last years [1, 8–14]. There are many different strategies to model the basic structural units of the capsids in coarse-grained models. In the simplest models, the coarse-grained is performed at the level of capsomers, which are represented as disks [1] (Fig. 3.2(d)), balls [15] (Fig. 3.2(e)), hexagonal and pentagonal pyramids [16] (Fig. 3.2(f)), ellipsoidal capsomers [17] (see Fig. 3.2(g)) or patchy particles [14].

Specially inspiring for this thesis has been the patchy particles model of Wilber *et. al* [14]. In it, the subunits are modeled as simple spheres with excluded volume interactions and patches with sort-ranged and directional specific interactions. This type of patchy particles models get the right lowest energy configuration, corresponding to a particular target structure, by properly designing the position of the patches. In this way it is possible to study the self-assembly of small structures ($T = 1$ or $T = 3$), as a function of subunit concentration, subunit-subunit interaction strength, and the orientation specificity of subunit interactions [14]. The main inconvenient of this type of models is that the interactions are designed to recover just a particular structure and thus cannot reproduce the possibility of obtaining different structures, as often occurs in many *in vitro* assembly experiments.

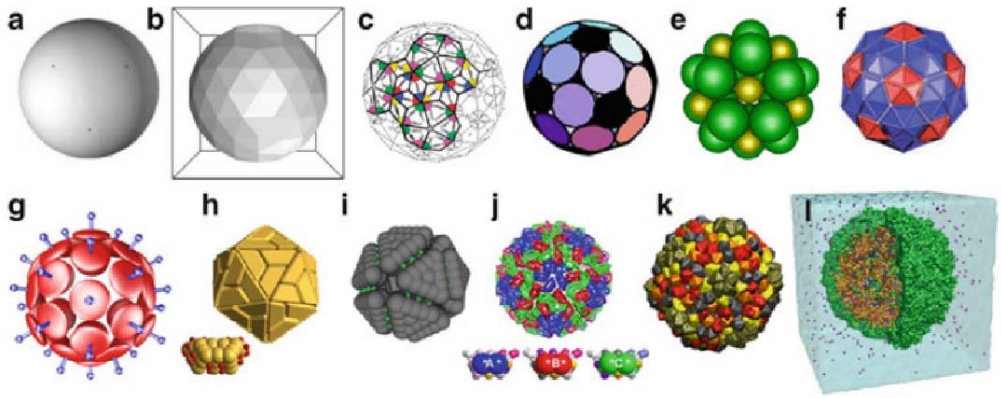


Figure 3.2: Different models of viral capsids. (a) Continuum elastic model of a spherical capsid. (b) A viral shell as a triangulated network [5]. (c) Tiling model of polyomavirus using kite and rhombic tiles [7]. (d-g). Coarse-grained models where capsomers and proteins are represented as (d) disks [1], (e) balls [15] over the surface of a sphere, (f) hexagonal and pentagonal pyramids [16], and (g) ellipsoidal capsomers with additional attractive and repulsive sites [17]. (h-j) Patchy-particles models with proteins described as: (h) multiple bead trapezoidal objects with many interaction sites [8], (i) truncated pyramids [18], (j) trapezoidal multi-beads of three different types [9]. (k) Shape based coarse-grained model of phage ΦX174 [19]. (l) Full atomic model of satellite tobacco mosaic virus, including water and ions [2]. Figure taken from Ref. [20]

In the next level of refinement, the basic units of the capsid, which can be proteins, dimers or trimers, are modeled using a collection of multiple beads properly placed to represent more accurately the shape of the protein. In this context, capsid proteins have been modeled using trapezoidal objects to represent the beta-barrel structure of some viral proteins [8, 9, 24] (Fig. 3.2(h) and (j)) or as triangular subunits which form icosahedral shells [18, 26, 27]. In most of these models, the interactions have to be arbitrarily tuned in a delicate way to reproduce just one single particular structure. That is why we opted for a simpler model, based on the minimal model of Ref. [1], described below.

3.2.1 Minimal model for capsid architecture

The coarse-grained model that will be developed and used in this thesis is partially based on the minimal model introduced by Zandi and co-authors in Ref. [1], (see Fig. 3.2(d)). In that model, the coarse-grained is performed at the level of capsomers which were originally modeled as discs of different sizes, representing pentamers and hexamers, that were forced to reside on the surface of a sphere. The ratio of sizes of the pentamers and hexamers was taken to be ≈ 0.85 and a Lennard-

Jones-like interaction with equal strength was assumed between them. Using MC simulations it was found that the Caspar and Klug structures of quasi-spherical viruses corresponded to minima of the free energy when 2 different structural units (i.e. pentamers and hexamers) were considered. The model was also used to study the minimal energy structures build up by identical capsomers, obtaining also structures with lower symmetry than icosahedral [37] that have been found for all-pentamers virus *in vitro*. Later on, this model was slightly modified replacing the capsomer discs by spheres [29] and used to find the optimal structures for elongated viruses [15]. In that study, the spherical template was replaced by a spherocylindrical surface as a constrained shape characteristic of prolate virus. In Chapter 8 of this thesis, we will use this model with a spherical and an icosahedral template to compare the energetics of spherical versus polyhedral capsids and to analyze the influence of the T-number on the tendency of buckling.

But, although this minimal model has been very insightful and successful to describe the optimal structures of both spherical and elongated viruses, the constraint of having the capsomers on top of a spherical or spherocylindrical template, makes it unrealistic to study the mechanisms of assembly and its kinetics. Thus, we wanted to extend this model and we needed to remove the template, the isotropic nature of the potential, and to incorporate a minimal set of ingredients to facilitate the assembly of empty shells directly from the solution, without any constraint. Our new model is described in the next section.

3.3 Model proposed in this thesis

As discussed previously, viruses use different strategies and elementary building blocks to build the capsid. For many viruses, individual proteins or dimers are the assembly units. However, there are some viruses such as human papillomavirus [31, 32], geminivirus [33], simian virus 40 [34], HK97 prohead [35] and bovine enterovirus [36], that form their capsids directly from hexamers and/or pentamers that have been preassembled in solution. For simplicity, we will focus on the assembly of viruses having preformed capsomers as assembly units, since capsomer-capsomer interactions are expected to be simpler and more isotropic than the complex interactions between individual proteins. In addition, since coarse-grained models are designed to be as efficient as possible, and the solvent is expensive in computational time, as it represents a substantial fraction of the system, we will use an implicit solvent model. In this way we eliminate the explicit representation of the solvent replacing it by effective interactions between the remaining solute particles.

Three are the minimal essential ingredients of the interactions required for modeling successfully the assembly. First, a short range repulsion is needed to mimic steric effects and prevent the overlapping of capsomers. Second, some kind of attraction is required to drive the assembly. A simple Lennard-Jones-like potential fullfills these requirements, but an additional ingredient is required to form shells of a particular size rather than compact clusters. In the model of the previous section, this ingredient was the fact that capsomers were constrained to lie on the surface of either a sphere [1, 37], a spherocylinder [15], or an icosahedron [30]. This simple model was enough to justify the origin of icosahedral symmetry in viruses, but cannot describe realistically the kinetics of the assembly process or the selection of a particular structure.

In this section we present an extension of that model, including two new ingredients that make the interaction more realistic and the direct assembly feasible without requiring artificial constraints, templates or local rules. Specifically, we now incorporate bending and twisting contributions to the capsomer-capsomer interaction. Thus, this potential is not isotropic but has a strong orientational dependence. The bending contribution selects a specific curvature ultimately related to the prefered angle of protein-protein interactions. So it is crucial for size selection, and facilitates the formation of shell-like structures rather than compact clusters. In fact, it was the mechanism suggested by CK in their pioneer paper [38] as the one selecting a specific T-number for a virus. The torsion contribution accounts for the differences between the internal and external surface of the coat proteins, and penalizes the formation of aberrant structures, making the assembly of closed shells more efficient. With these two new ingredients, capsomers do not need an artificial constraint or template and can self-assemble spontaneously, as it will be shown in Chapter 7.

In our coarse-grained model we will consider in general two types of capsomers: hexamers and pentamers, that are modeled as spheres of different size. The ratio between the diameter of the hexamer, σ_h, and of the pentamer, σ_p was determined by inscribing their equatorial circles, respectively, into a pentagon or a hexagon of the same edge, as depicted in Fig. 3.3. The ratio between hexamer and pentamer diameters thus

$$\frac{\sigma_p}{\sigma_h} = \frac{\tan(\pi/6)}{\tan(\pi/5)} \tag{3.1}$$

The capsomer-capsomer interaction is modeled by a pair potential that has a short-range repulsion and an attractive term that drives capsomer aggregation. The repulsion part is modeled by an isotropic Lennard-Jones potential (V_{LJ}), and the attraction is modulated by angular bending $V_{ang}(\hat{\mathbf{r}}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)$ and torsion $V_{tor}(\overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)$ terms. Specifically, the potential is given by:

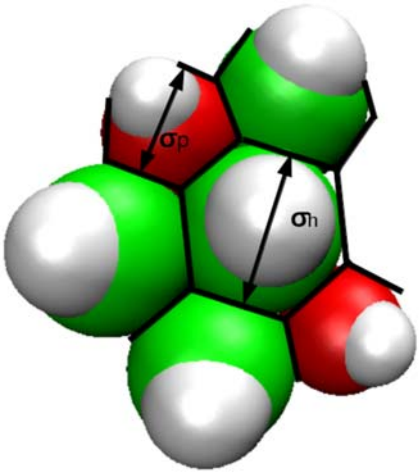


Figure 3.3: Hexagonal lattice for pentamers and hexamers of different diameters, σ_p and σ_h respectively.

$$V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) = \begin{cases} V_{LJ}(r) & r \leq 2^{-1/6}\sigma_{ij} \\ \\ V_{LJ}(r)V_{ang}(\hat{\mathbf{r}}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)V_{tor}(\overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) & r > 2^{-1/6}\sigma_{ij} \end{cases} \tag{3.2}$$

The expression for the Lennard-Jones-like part of the potential is:

$$V_{LJ}(r) = \varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r}\right)^{6}\right] \tag{3.3}$$

where ε_{ij} is the binding energy between capsomers, σ_{ij} is the equilibrium distance corresponding to the minimum of the LJ potential, r is the separation between capsomers centers, and $2^{-1/6}\sigma_{ij}$ is the distance at which the Lennard-Jones contribution vanishes (i.e., $V_{LJ}(r = 2^{-1/6}\sigma_{ij}) = 0$). For simplicity, we will assume that the strength of the binding interactions between hexamers, pentamers and between hexamers and pentamers, are all the same, $\epsilon_{ij} = \epsilon_0$. Following previous estimates of typical binding energies [1, 39], we take ε_0 to be $20k_BT$, with k_B being the Boltzmann constant and T the absolute temperature. Since we have 2 types of capsomers, the equilibrium distance σ_{ij} between them in the Lennard-Jones potential depends on the type of capsomers which are interacting and is given by

$$\begin{aligned} \sigma_{hh} &= \sigma_h \equiv \sigma \\ \sigma_{pp} &= \sigma_p = \frac{\tan(\pi/6)}{\tan(\pi/5)}\sigma \approx 0.8\sigma \\ \sigma_{ph} &= (\sigma_p + \sigma_h)/2 \approx 0.9\sigma \end{aligned} \tag{3.4}$$

For computational efficiency, the Lennard-Jones potential is truncated at 5σ. The angular dependency of the potential is given by:

$$V_{ang}(\hat{\mathbf{r}}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) = e^{\dfrac{-(\theta_{ij}-\nu)^2}{2\alpha^2}} e^{\dfrac{-(\theta_{ji}-\nu)^2}{2\alpha^2}} \tag{3.5}$$

where the vector $\overrightarrow{\Omega}_i$ defines the orientation of capsomer i and θ_{ij} is the angle between $\overrightarrow{\Omega}_i$ and the unit vector $\hat{\mathbf{r}}_{ij}$ connecting campsomers i and j (see Fig. 3.4). The parameter ν is the preferred angle, related to the spontaneous curvature, and α controls the local bending stiffness. For large values of α, the isotropic Lennard-Jones potential is recovered, whereas small values of α are indicative of a strong bending penalty.

The interaction potential also includes a torsion contribution given by:

$$V_{tor}(\overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) = e^{\dfrac{-k_t(1-cos\xi)}{2}} \tag{3.6}$$

where k_t is the torsion constant and ξ is the angle between the planes defined by $\hat{\mathbf{r}}_{ij}$ and $\overrightarrow{\Omega}_i$, and by $\hat{\mathbf{r}}_{ij}$ and $\overrightarrow{\Omega}_j$ (see Fig. 3.4). Specifically, the torsion angle ξ is given by:

$$cos(\xi) = (\widehat{r}_{ij} \times \overrightarrow{\Omega}_i) \cdot (\overrightarrow{\Omega}_j \times \widehat{r}_{ji}) = cos(\theta_{ij})cos(\theta_{ji}) + \overrightarrow{\Omega}_i \overrightarrow{\Omega}_j \tag{3.7}$$

Without torsion, two capsomers with the correct ν angle but opposite $\overrightarrow{\Omega}_i$ orientations would have the same energy than the right configuration. This torsion contribution penalizes this incorrect behavior, so capsomers form closed shells in the end, instead of many connected surfaces with different concavity. We have verified that the value of this torsion constant is not so decisive. Accordingly, in our simulations, we have fixed the torsion constant at $k_t = 1.5$, a value that warranties the succesful assembly of the smallest T-number structure in our assembly simulations.

One of the advantages of this simple potential is that, for bound capsomers, the parameters can be easily identified with simple bending and stretching contributions. Expanding the potential eq. 3.2 around the equilibrium distance σ and orientation ν, we obtain:

$$V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) \sim -\epsilon_0 + \frac{1}{2}\frac{72\epsilon_0}{\sigma^2}(r_{ij} - 0.5\sigma)^2 + \frac{1}{2}\frac{2\epsilon_0}{\alpha^2}(\theta_{ij} - \nu)^2 \tag{3.8}$$

where we have assumed that $\theta_{ij} = \theta_{ji}$ and ignored the torsional contribution.

The previous equation clearly shows that, for bound capsomers, $k_s = \frac{72\epsilon_0}{\sigma^2}$ is the effective spring constant, $k_b = \frac{\epsilon_0}{\alpha^2}$ plays the role of the local bending stiffness,

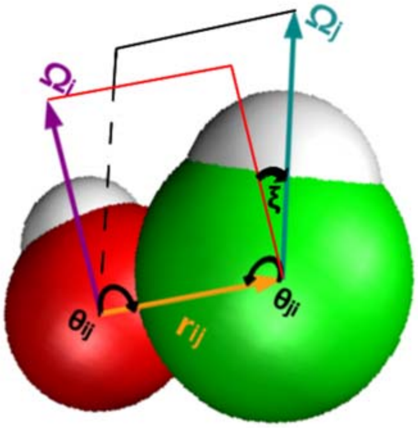


Figure 3.4: Representation of the distance $\mathbf{r}_{ij}$ (orange), the orientation vectors $\overrightarrow{\Omega_i}$ (violet) and $\overrightarrow{\Omega_j}$ (blue) and the angles used in the interaction potential in this case for an interacting hexamer (green) and pentamer (red). The white sphere is overlapped on each capsomer to indicate its orientation.

Structure	ν^{hh}_{opt}	ν^{hp}_{opt}
T=1	2.124	
Snub cube	1.954	1.912
T=3	1.936	1.897
T=4	1.885	1.847
T=7	1.797	1.780

Table 3.1: Optimal values of the preferred angle ν_{opt} predicted for the lowest T-number structures and the snub cube, calculated between two hexamers (or 2 pentamers in the $T = 1$ case) and between a hexamer and a pentamer.

and ν is the preferred angle of the interaction between capsomers (ultimately connected to the preferred angle of the interaction between coat proteins).

As we will see, ν is the main parameter in the selection of a particular structure for a viral capsid. In fact, we can estimate the theoretical values of the preferred angle that will be optimal for each T-number structure, that will be denoted as ν_{opt}. To that end, we have calculated the angle between neighboring capsomers from the coordinates of the icosahedral spherical code corresponding to the solution of the maximum volume problem listed in Ref. [40], as described in Appendix A. The resulting ν_{opt} for each T are listed in Table 3.1.

The preferred angle ν is also connected to the preferred radius R and spontaneous curvature $C = \frac{2}{R}$ of a closed spherical shell, as show in Fig. 3.5, by the

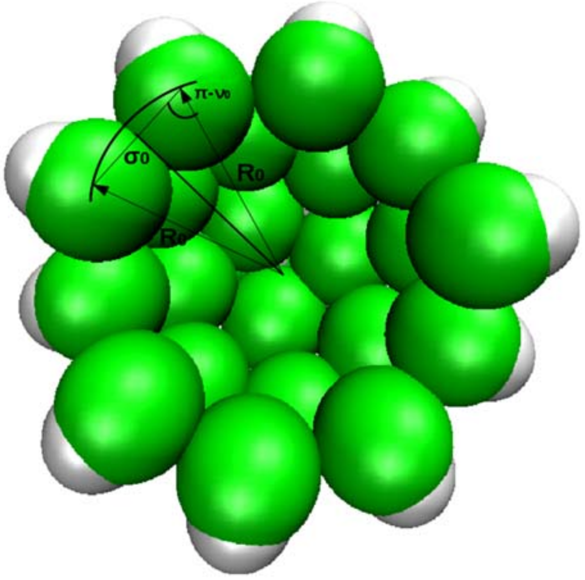


Figure 3.5: Transverse section of $T = 3$ shell of radius R_0 and preferred angle ν_0.

relation:

$$R = \frac{-\sigma}{2cos(\nu)} \tag{3.9}$$

$$C = \frac{-4cos(\nu)}{\sigma} \tag{3.10}$$

Eq. 3.9 shows that the preferred radius is very sensitive to small changes in the preferred angle of the interaction ν. It is worth emphasizing that this preferred radius set by ν may in general be different from the equilibrium (optimal) radius of a given shells, since this is ultimately dictated by the competition between bending and stretching energies (as shown in Chapter 4).

Another nice feature of the potential is that, for a closed shell, the model can be straightforwardly connected to the discretized version of the elastic energy of a triangulated sphere. From this comparison, we can quantify the effective elastic properties of a closed shell in the continuum limit. As shown in Appendix B and in Refs. [5, 41], in the continuum limit the model is equivalent to an elastic sphere with Young modulus

$$Y = \frac{2}{\sqrt{3}} k_s = \frac{144}{\sqrt{3}} \frac{\epsilon_0}{\sigma^2}, \tag{3.11}$$

Poisson's ration

$$\mu = \frac{1}{3}, \tag{3.12}$$

and bending rigidity

$$\kappa = \frac{3\sqrt{3}}{8}\frac{\epsilon_0}{\alpha^2}. \tag{3.13}$$

This correspondence will be useful to interpreted our simulation results.

The two parameters ν and α are the more relevant parameters, and we will see that they control the outcome of the self-assembly process (Chapter 6 and 7) and the mechanical response of the capsid (Chapter 10).

3.4 Conclusions

In this chapter we have seen that there are different ways to model the structure of viral capsid and to describe capsid assembly. Capsids can be described at different resolution levels, from a full atomistic description to a continuum level. All models have their limitations and the selection of a particular one depends on the kind of insight that we want. Atomistic models are adequate to understand the finer structural details of a specific virus, but cannot be used to study complex phenomena as viral assembly. On the other hand, coarse-grained models are specially useful to investigate the general physical mechanisms, but lack realistic details.

In this chapter, we have presented a coarse-grained, physically inspired, model that will be used to investigate the physical mechanisms underlying viral self-assembly. The model was inspired by the ideas of Refs. [1, 14], and incorporates two new parameters ν and α, which are related directly to spontaneous curvature and bending stiffness. One of the main advantages of the model is that, contrarily to previous models, it does not require any constraint, template or local rules for a successful assembly. In addition the parameters of the interaction are not laboriously and arbitrarily tuned, but rather have a clear physical meaning and can be linked to experimentally measurable biophysical quantities. Thus the predictions of the model could in principle be connected and useful in real experiments.

We will see that, despite its simplicity, the model can predict many interesting behavior observed in real viruses. With this model we will be able to study the stability and selection of different shapes of viral capsids, the kinetics of capsid formation, and its mechanical response upon indentation, as described in the following chapters. This type of studies can be extremely useful in *in vitro* experiments for nano-technologies and health area applications.

The approximations involved in the severe coarse-graining of the model imply a loss of potentially relevant structural details and has of course some limitations. On one hand, the possible cooperative hydrophobic effects promoted by the solvent during assembly, and the finer structural details of the interactions cannot be captured. However, more realistic details, like the shape of coat subunits and their intrinsic anisotropic interactions can be incorporated in the model, but are out of the scope of this thesis. They will be addressed in future investigations.

Appendix A

Calculation of the expected value of ν for different structures

We have calculated the theoretical values of the preferred angle that one expected to be optimal for each T-number structure. These values, denoted as ν_{opt}, are calculated as the angle between neighboring capsomers obtained from the coordinates of the icosahedral spherical code listed in Ref. [40] corresponding to the solution of the maximum volume problem. For a spherical capsid of radius R centered at the origin of coordinates, the optimal angle is the angle formed between the vector $\overrightarrow{\Omega}_1 = \frac{1}{R}(x_1, y_1, z_1)$ defining the orientation of a capsomer and the vector $\mathbf{r}_{ij} = \frac{1}{r_{ij}}(x_2 - x_1, y_2 - y_1, z_2 - z_1)$ connecting 2 neighbor capsomers. The coordinates in Ref. [40] are given for a unit sphere of radius R. As shown in Fig. 3.5, by simple geometry:

$$R \sin(\nu - \pi/2) = \frac{1}{2} r_{ij} \tag{3.14}$$

where $r_{ij} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$ and 1, and 2 refer to two nearest neighbors. Accordingly:

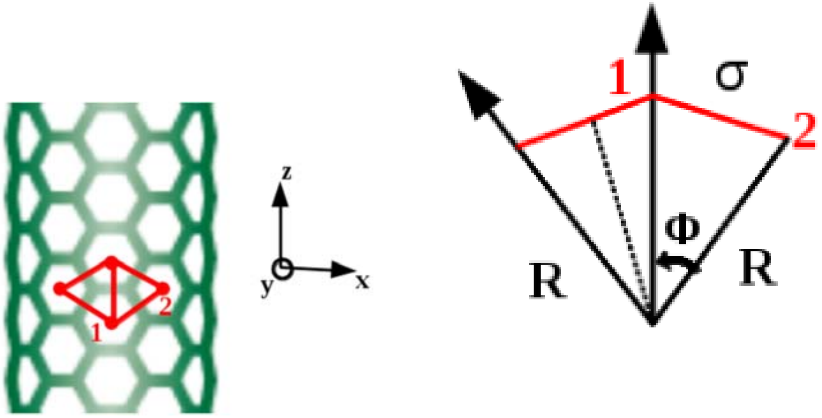
$$\nu_{opt} = \frac{\pi}{2} + \frac{arccos\,(r_{ij}/R)}{2} \tag{3.15}$$

The value of ν_{opt} can be calculated using two adjacent hexamers or a pentamer and a nearest neighbor hexamer. Table 3.1 shows the theoretical ν_{opt} for smaller T-number structures plus snub cube, calculated by eq. 3.15, in both cases.

Appendix B

Local and global bending stiffness

Comparing the global bending energy in the continuum limit with that obtained from our model, we can calculate the effective bending stiffness κ in terms of the local parameters of the interaction potential. For simplicity, we will do it by calculating the bending energy of a cylinder in our coarse-grained model.



(a) (b)

Figure 3.6: (a) Hexagonal lattice of capsomers placed over a cylinder of radius R and length L and (b) cross-section of the positions of capsomers 1 and 2 over the cylinder.

Let us assume a hexagonally-ordered network of our coarse-grained capsomers on the surface of a cylinder of length L and radius R, see Fig. 3.6 (a). According to this figure, the positions of capsomer 1 and 2 are given by:

$$\begin{aligned} &x_1 = 0; y_1 = R; z_1 = 0 \\ &x_2 = R\sin(\phi); y_2 = R\cos(\phi); z_2 = \frac{\sigma}{2} \end{aligned} \tag{3.16}$$

where the origin of z has been taken at the location of capsomer 1. The distance between these two capsomers is $r_{12}^2 = R^2\sin^2(\phi) + R^2(\cos(\phi)-1)^2 + \frac{\sigma^2}{4}$, that at equilibrium will coincide with the equilibrium distance σ, thus $r_{12}^2 = \sigma^2$. By regrouping this equation we obtain $2R^2(1-\cos(\phi)) = \frac{3}{4}\sigma^2$, that using the approximation for small angles $\cos(\phi) \approx 1 - \frac{\phi^2}{2}$, finally yields

$$\phi = \frac{\sqrt{3}}{2}\frac{\sigma}{R} \tag{3.17}$$

According to eq. 4.3 the bending energy between capsomer 1 and 2 around the equilibrium distance and orientations is given by

$$E_b^{12} = 2\frac{\epsilon_0}{2\alpha^2}(\theta_{12} - \frac{\pi}{2})^2 \tag{3.18}$$

where $\cos\theta_{ij} = \widehat{r_{ij}} \cdot \overrightarrow{\Omega_i}$. Using the positions of capsomers 1 and 2 from eq. 3.16, we obtain

$$\begin{aligned}\cos\theta_{12} &= \frac{1}{\sigma}(R\sin(\phi), R(\cos(\phi) - 1), \frac{\sigma}{2}) \cdot (0, 1, 0) = \\ &= \frac{R}{\sigma}(\cos(\phi) - 1) \approx \frac{-R\phi^2}{2\sigma}\end{aligned} \tag{3.19}$$

For small angles $\arccos(x) \approx \frac{\pi}{2} - x$ and the angle between capsomer 1 and 2 then becomes

$$\theta_{12} = \frac{\pi}{2} + \frac{R\phi^2}{2\sigma} \tag{3.20}$$

Using eq. 3.19 in eq. 3.18 the total contribution between 1 and 2 then becomes

$$E_b = \frac{9}{64}\frac{\epsilon_0}{\alpha^2}\frac{\sigma^2}{R^2} \tag{3.21}$$

Each capsomer interacts with 4 neighbors. Thus, the bending energy per capsomer becomes $E_b^c = \frac{4}{2}E_b$, where we have divided by 2 to avoid double counting. Now if we multiply it by the total number of capsomers in the cylinder, given by $N_c = \frac{2\pi RL}{\frac{\sqrt{3}}{2}\sigma^2}$, (where $A_H = \frac{\sqrt{3}}{2}\sigma^2$ is the effective area of a capsomer) the total bending energy is:

$$E_b^{tot} = \frac{4}{2}N_c E_b = \frac{3\sqrt{3}}{8}\pi\frac{\epsilon_0}{\alpha^2}\frac{L}{R} \tag{3.22}$$

Comparing eq. 3.22 with the standard expression of the total bending energy of a cylinder $E^{cyl} = \pi\kappa\frac{L}{R}$ [5], we finally get the effective bending stiffness κ.

$$\kappa = \frac{3\sqrt{3}}{8}\frac{\epsilon_0}{\alpha^2} \tag{3.23}$$

References

[1] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, Origin of icosahedral symmetry in viruses, *Proc. Natl. Acad. Sci. USA.*, **101**, 15556-15560, 2004.

[2] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, and K. Schulten, Molecular dynamics simulations of the complete satellite tobacco mosaic virus, *Structure*, **14**, 437-49, 2006.

[3] M. Deserno. Mesoscopic Membrane Physics: Concepts, Simulations, and Selected Applications. *Macro. Rap. Commun.*, **30**, 752, 2009.

[4] LD. Landau, EM. Lifshitz, Theory of Elasticity. Pergamon Press, London, 1975.

[5] J. Lidmar, L. Mirny, and D. Nelson, Virus shapes and buckling transitions in spherical shells, Phys. Rev. E, **68**, p. 051910, 2003.

[6] R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger. Local rules simulation of the kinetics of virus capsid self-assembly. *Biophy. J.*, **75**, 2626, 1998.

[7] R. Twarock, A tiling approach to virus capsid assembly explaining a structural puzzle in virology, J. Theor. Biol., **226**, 477-82, 2004.

[8] D. C. Rapaport. Modeling capsid self-assembly: design and analysis. *Phys. Biol.*, **7**, 045001, 2010.

[9] H. D. Nguyen, V. S. Reddy, and C. L. Brooks. Invariant polymorphism in virus capsid assembly. *J. Ame. Chem. Soc.*, **131**, 2606, 2009.

[10] M. Goldberg, Viruses and a mathematical problem, *J. Mol. Biol.*, **24**, 337-338, 1967.

[11] B. W. Clare and D. L. Kepert, The closest packing of equal circles on a sphere, *Proc. R. Soc. London Ser. A*, **405**, 329-344, 1986.

[12] D. A. Kottwitz, The densest packing of equal circles on a sphere, *Acta Cryst.* A, **47**, 158-165, 1991.

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

[14] 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**, 085106, 2007.

[15] A. Luque, R. Zandi, and D. Reguera, Optimal architectures of elongated viruses, *Proc. Nat. Acad. Sci. USA* **107**, 53235328, 2010.

[16] S. N. Fejer, T. R. James, J. Hernandez-Rojas, and D. J. Wales, Energy landscapes for shells assembled from pentagonal and hexagonal pyramids, *Phys. Chem.*, **11**, 2098-2104, 2009.

[17] S. Fejer, D. Chakrabarti and D. Wales, Emergent complexity from simple anisotropic building blocks: shells, tubes, and spirals. *ACS Nano*, **4**, 219-228, 2010.

[18] O. M. Elrad and M. F. Hagan. Encapsulation of a polymer by an icosahedral virus. *Phys. Biol.*, **7**, 045003, 2010.

[19] A. Arkhipov, PL. Freddolino and K. Schulten, Stability and dynamics of virus capsids described by coarse-grained modeling, *Structure*, **14**, 1767-1777, 2006.

[20] MG. Mateu, Structure and Physics of Viruses, London, Springer, 2013.

[21] P. E. Prevelige, D. Thomas, and J. King, Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells, *Biophys. J.*, **64**, 824-35, 1993.

[22] P. A. Thuman-Commike, B. Greene, J. A. 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**, 559-568, 1998.

[23] 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**, 1097-106, 2006.

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

[25] M. F. Hagan, O. M. Elrad, and R. L. Jack. Mechanisms of kinetic trapping in self-assembly and phase transformation. *J. Chem. Phys.*, **135**, 104115, 2011.

[26] H. D. Nguyen, V. S. Reddy, and C. L. Brooks. Deciphering the kinetic mechanism of spontaneous self-assembly of icosahedral capsids. *Nano letters*, **7**, 338, 2007.

[27] D. Rapaport, J. Johnson, and J. Skolnick. Supramolecular self-assembly: molecular dynamics modeling of polyhedral shell formation. *Comp. Phys. Comm.*, **231**, 121-122, 1999.

[28] H. Fraenkel-Conrat and R. C. Williams, Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components, *Proc. Natl. Acad. Sci. USA.*, **41**, 690-698, 1955.

[29] A. Luque, Thesis. Structure, Mechanical Properties, and Self-Assembly of Viral Capsids, 2011.

[30] M. Aznar, A. Luque and D. Reguera, Relevance of capsid structure in the buckling and maturation of spherical viruses, *Phys Biol.* **9** 036003, 2012.

[31] S. Mukherjee, M. V. Thorsteinsson, L. B. Johnston, P. A. DePhillips, and A. Zlotnick, A quantitative description of in vitro assembly of human papillomavirus 16 virus-like particles, *J. Mol. Biol.*, **381**, 229-37, 2008.

[32] G. L. Casini, D. Graham, D. Heine, R. L. Garcea, and D. T. Wu, In vitro papillomavirus capsid assembly analyzed by light scattering, *Virology*, **325**, 320-7, 2004.

[33] W. Zhang, NH. Olson, TS. Baker, L. Faulkner, M. Agbandje-McKenna, MI. Boulton, JW. Davies, and R. McKenna, Structure of the maize streak virus geminate particle, *Virology*, **279**, 471-7, 2001.

[34] S. Kanesashi, K. Ishizu, M. Kawano, S. Han, S. Tomita, H. Watanabe, K. Kataoka, and H. Handa, Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro, *J. Gen. Virol.*, **84**, 1899-1905, 2003.

[35] Z. Xie and R. W. Hendrix, Assembly in vitro of bacteriophage HK97 proheads, *J. Mol. Biol.*, **253**, 74-85, 1995.

[36] CL. Li, JCY. Wang, MW. Taylor and A. Zlotnick, In Vitro Assembly of an Empty Picornavirus Capsid follows a Dodecahedral Path, *J. Virology*, **86**, 23, 2012.

[37] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Viral self-assembly as a thermodynamic process, *Phys. Rev. Lett.*, **90**, p. 248101, 2003.

[38] D.L.D. Caspar and A. Klug, Physical principles in the construction of regular viruses, in Cold Spring Harbor Symp. Quant. Biol., **27**, 1-24, Cold Spring Harbor Laboratory Press, 1962.

[39] V. S. Reddy, H. A. Giesing, R. T. Morton, A. Kumar, C. B. Post, C. L. Brooks, and J. E. Johnson, Energetics of quasiequivalence: computational analysis of protein-protein interactions in icosahedral viruses, *Biophys. J.*, **74**, 546-558, 1998.

[40] RH. Hardin, NJA. Sloane and WD. Smith, Tables of spherical codes with icosahedral symmetry, published electronically at http://www.research.att.com/∼njas/icosahedral.codes/.

[41] T. Nguyen, RF. Bruinsman and WM. Gelbart, Elasticity theory and shape transitions of viral shells. *Phys. Rev. E.* **72** 051923, 2005.

[42] W. K. Kegel and P. van der Schoot, Competing hydrophobic and screened-coulomb interactions in hepatitis B virus capsid assembly, *Biophys. J.*, **86**, 3905-13, 2004.

CHAPTER 4

Elastic modeling of viral capsids

4.1 Introduction

As mentioned in the previous chapter, one of the simplest way to describe the shape and size of viral capsids is by using continuum elasticity theory [1–10]. In this framework, viruses are modeled as nanoscopic thin shells, since in most cases the thickness of the shell is much smaller than the radius of the capsid. Using this approximation, the elastic energy of viral shells can be characterized by two simple contributions: a stretching term describing the energetic cost of modifying the in-plane area of the capsid

$$E_s = \frac{1}{2}\int dS\left(2\mu u_{ij}^2 + u_{kk}^2\right) \tag{4.1}$$

and a bending contribution associated to out of plane deformations that change the local curvature H from the spontaneous or preferred curvature C_0

$$E_b = \frac{1}{2}\int dS\left(\kappa(H - C_0)^2 + 2\kappa_G K\right) \tag{4.2}$$

Thus, in this approximation, the energy of the capsids is characterized by a few phenomenological constants: μ and λ, the Lamé coefficients which are related to the 2D Young modulus Y and Poisson ratio μ; the bending stiffness κ, the Gaussian rigidity κ_G, and the spontaneous curvature C_0 [1–5]. The optimal shape of the capsid is then obtained by minimizing the total elastic energy of the shell $E_b + E_s$, a procedure that is commonly done numerically by using the discretized versions of eqs 4.1 and 4.2 on a triangular lattice [2, 3, 6–8, 10].

The elastic energy of a shell depends ultimately only on a single non-dimensional parameter, the Foppl-von Karman (FvK) number $\gamma = \frac{YR^2}{k}$, which is the ratio of the stretching and bending energies [2–4, 7, 8, 10, 11]. Using this strategy, it has been possible to study the optimal shape (i.e. spherical, elongated, or conical) of a viral shell as a function of the FvK number and the spontaneous curvature [7] (see Fig. 4.1), and interesting phenomena like the occurrence of a buckling transition, explained in more detail in Chapter 8. Thus, despite the crudeness of this description, that replaces the complex nature of the interaction between proteins by an effective continuous material, elastic models produce useful information on the shape and stability of a virus.

In this chapter we will present a simple elastic model obtained as a continuous approximation to the coarse-grained discrete model of the previous chapter. The model captures in a simplified way the competition between bending and stretching energies for a spherical capsid. We will use this simple model to analyze what is the optimal radius of our model capsids and how it depends on the different parameters of the model.

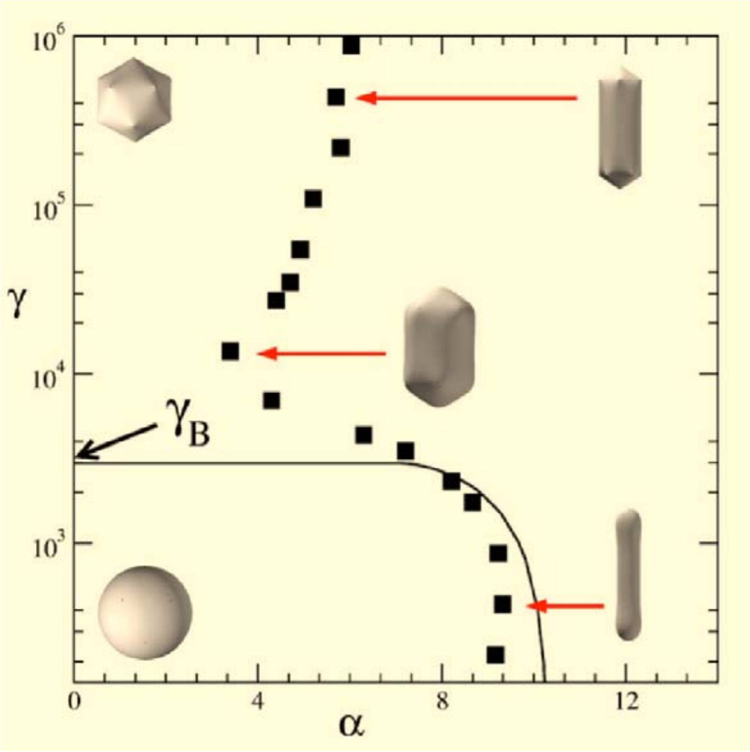


Figure 4.1: Phase diagram of viral shapes as a function of the FvK number γ and the dimensionless spontaneous curvature $\alpha = C_0 S^{1/2}$, where S is the area of the shell. The phase diagram shows the buckling transition, which separates spherical from polyhedral shells, and a first-order transition between spherical and spherocylindrical shells. Figure taken from Ref. [7].

The chapter is distributed as follows. In Section 4.2 we will present and describe our simple bending plus stretching model. The results of the model in the bending-dominated and stretching-dominated limits will be described in subsections 4.2.1 and 4.2.2. Subsection 4.2.3 will be devoted to analyze the intermediate region and its implications on the optimal size of the virus. Finally, the main results and conclusions are summarized in Section 4.3.

4.2 Bending and stretching model

We propose a simple elastic model for the study of the interplay between the bending and the stretching contributions in determining the optimal radius for a spherical capsid of radius R. The model is derived from a continuous approximation of the coarse-grained model introduced in Chapter 3. We will apply the model to viruses, but this model could also be used for other systems objects as membranes [13], colloidosomes [14], or even black holes [15].

In Section 3.3, we showed that in our model the pair potential between two bounded capsomers around the equilibrium distance and optimal orientation was given by

$$V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) \sim -\varepsilon_0 + \frac{1}{2}\frac{72\epsilon_0}{\sigma^2}(r-\sigma)^2 + \frac{1}{2}\frac{2\epsilon_0}{\alpha^2}(\theta_{ij}-\nu)^2 \tag{4.3}$$

where ε_0 is the binding energy between capsomers, σ is the equilibrium distance corresponding to the minimum of the LJ potential, r is the separation between capsomer centers, θ_{ij} is the angle between the orientation of capsomer i, $\overrightarrow{\Omega}_i$, and the unit vector $\hat{\mathbf{r}}_{ij}$ connecting campsomers i and j and the parameter ν is the preferred angle. Let us now consider that this pair of bound capsomers are part of a complete spherical capsid of radius R_0, where all capsomers are at the equilibrium distance σ and oriented radially outwards. The value of this radius R_0 can be approximately obtained by a simple full covering model, as described in Appendix A of Chapter 8.

If we now change the radius, and assume a uniform deformation, the local δr and global δR deformations are related by

$$\frac{\delta r}{\sigma} = \frac{\delta R}{R_0} \tag{4.4}$$

Accordingly, for a uniform radial deformation, the total energy of the capsid can be approximated as:

$$E_t = \frac{Nz}{2}(-\epsilon_0 + \frac{1}{2}\frac{72\epsilon_0}{\sigma^2}\delta r^2 + \frac{1}{2}\frac{2\epsilon_0}{\alpha^2}\delta\nu^2) \tag{4.5}$$

where N is the total number of capsomers, z is the average coordination number (that would be 6 for a hexagonal lattice) and $\delta\nu = \theta_{ij} - \nu$. For this closed capsid, $\theta_{ij} = \arccos\left(\frac{-\sigma}{2R}\right)$ that can be approximated as:

$$\theta_{ij} = \arccos\left(\frac{-\sigma}{2R}\right) \approx \frac{\pi}{2} + \frac{\sigma}{2R} - \bigcirc(\frac{\sigma^3}{R^3}) \tag{4.6}$$

Using eqs. 4.4 and 4.6 in eq. 4.5, the total elastic energy becomes:

$$E_t = \frac{zN}{2}\left(-\epsilon_0 + \frac{1}{2}\frac{72\epsilon_0}{R_0^2}(R-R_0)^2 + \frac{1}{2}\frac{2\epsilon_0}{\alpha^2}\left(\frac{\sigma}{2R} - (\nu - \frac{\pi}{2})\right)\right)^2 \tag{4.7}$$

This expression can be further simplified into:

$$\Delta E_t = \frac{1}{2}(R-R_0)^2 + \frac{b}{2}\left(\frac{1}{2R} - \left(\nu - \frac{\pi}{2}\right)\right)^2 \tag{4.8}$$

where $\Delta E_t = \frac{E_t - E_0}{18\epsilon_0 z N} R_0^2$ is the energy change with respect to the global optimal energy, E_0, $b = \frac{R_0^2}{36\alpha^2}$, is related to the inverse of the FvK number, and R_0 is the radius that would minimize the stretching energy. From now on, and for the sake of simplicity all radii will be expressed in units of σ.

The resulting simplified expression for the energy ΔE_t depends on the radius R of the capsid and has two clear contributions. The first term represents the stretching energy of the capsid, whereas the second term accounts for the bending cost. The relative importance of these two terms is controlled by the parameter b. Let's first discuss the two limiting regimes determined by the value of b.

4.2.1 Stretching-dominated limit

In the case of $b \ll 1$, stretching is more important than bending. Thus the first term of the energy dominates and the second term of the energy can be neglected yielding

$$\Delta E_t \approx \frac{1}{2}(R - R_0)^2 \tag{4.9}$$

In this case, the optimal radius is trivially given by:

$$\left.\frac{d\Delta E_t}{dR}\right|_{R^*} = 0 \Rightarrow R^* = R_0 \tag{4.10}$$

Therefore, for negligible bending, the optimal value of the radius is just R_0, and the energy as a function of the radius has a parabolic behavior around it, as shown in Fig.4.2. For $R < R_0$ the capsid is compressed, whereas for $R > R_0$ it is stretched.

4.2.2 Bending-dominated limit

For $b >> 1$ bending becomes the dominant contribution, and stretching can be neglected, yielding for the energy

$$\Delta E_t \approx \frac{b}{2}\left(\frac{1}{2R} - (\nu - \frac{\pi}{2})\right)^2 \tag{4.11}$$

In this case, the optimal radius R^* is dictated by the preferred angle ν:

$$\left.\frac{d\Delta E_t}{dR}\right|_{R^*} = 0 \Rightarrow R^* \equiv \frac{1}{2(\nu - \pi/2)} = R_b \tag{4.12}$$

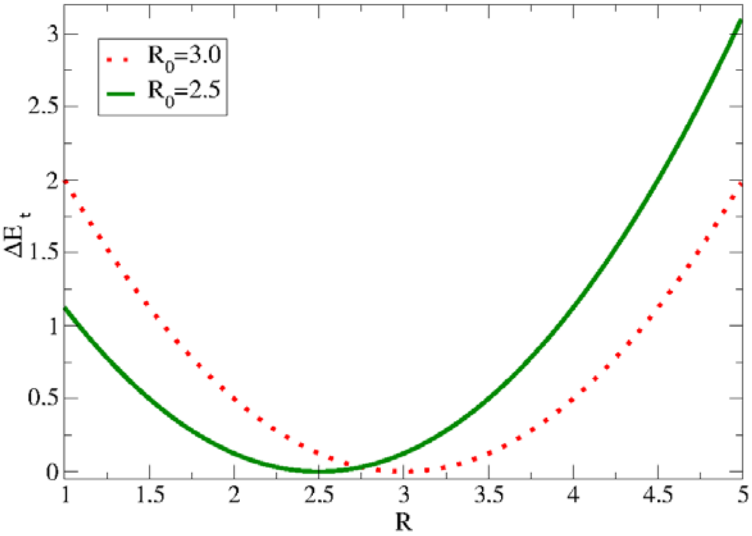


Figure 4.2: Energy ΔE_t versus the radius R in the stretching-dominated regime corresponding to $b = 0$ and for two different values of R_0. The energy is parabolic with a minimum localized at R_0.

The energy is not parabolic around R_b, as illustrated in Fig.4.3.

Thus, in general, we have two different preferred radius for the capsid. On one hand, there is R_0, which is fixed by the number of capsomers in the shell, and it is the preferred radius to minimize the stretching energy. On the other hand, we have $R_b = \frac{1}{2(\nu - \pi/2)}$ which is the radius favored by the spontaneous curvature. Hence, when these two values are different, we have a interesting competence between bending and stretching, which is explored in the next section.

4.2.3 General case

In the general case, we cannot ignore any of the two components, and the optimal radius will be given by:

$$\begin{aligned} \frac{d\Delta E_t}{dR}\Big|_{R*} &= 0 \\ R^{*4} - R_0 R^{*3} + \frac{b}{2}(\nu - \frac{\pi}{2})R^* - \frac{b}{4} &= 0 \end{aligned} \tag{4.13}$$

Thus, to obtain the optimal value R^* one has to solve a quartic equation. This type of equation can be solved analytically and has in general four different solutions, whose exact (but cumbersome) expression is listed in Appendix A. The nature of the solutions is dictated by the discriminant Δ:

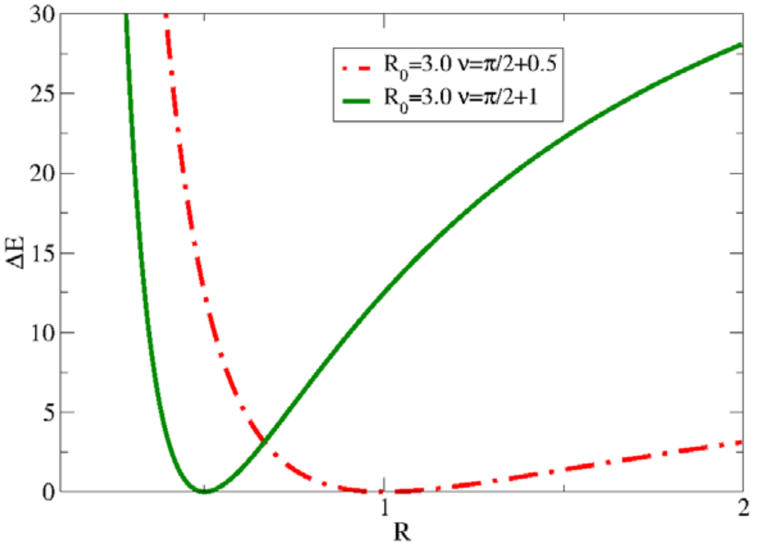


Figure 4.3: Energy ΔE_t versus the radius R in the bending dominated limit for $b = 100$ and different values of ν. The minimum value of the energy is localized in each case at $R_b = \frac{1}{2(\left(\nu_0-\frac{\pi}{2}\right))}$.

$$
\begin{aligned}
\Delta \;=\;& \frac{b^2}{16}\Big\{ -27\left(b^2\left(\nu-\frac{\pi}{2}\right)+R_0^4\right)+b\Big[-64+ \\
&+2R_0\left(\nu-\frac{\pi}{2}\right)\left(48+R_0\left(\nu-\frac{\pi}{2}\right)\left(3+4R_0\left(\nu-\frac{\pi}{2}\right)\right)\right)\Big]\Big\} \qquad (4.14)
\end{aligned}
$$

When $\Delta < 0$, the equation has 2 real and 2 complex conjugated roots and when > 0, the equation has 4 real roots. In our case, only the real and positive solutions have a physical meaning. It can be shown that when $\Delta < 0$, there is only one real and positive solution, whereas when $\Delta > 0$, there are 3 positive and real solutions, whose meaning will be discussed later on.

As an example, Fig.4.4 shows the real and positive solutions of R versus the preferred angle ν for $R_0 = 3$ and different b values. For $b = 1$ (see yellow line) we are in the stretching-dominated limit, and the optimal radius is $R^* \approx R_0$, irrespective of the value of the spontaneous curvature. In the opposite limit, for $b = 1000$ (green line) the optimal radius is very close to that dictated by the curvature $R_b = \frac{1}{2\left(\nu-\frac{\pi}{2}\right)}$. The solution for $b = 100$ (blue line) also is close to this limit, specially at large values of ν. Interestingly, for $b = 10$ (red line) we have obtained an intermediate behavior, where for small values of ν corresponds to the stretching-dominated solution $R^* \approx R_0$ (small b), but for large ν tends to bending radius $R^* \approx R_b$. In between, there is a transition region, where 3 possible solutions for the radius exist. It is worth to note that for all values of b, the solutions meet with a common value $R^* \approx R_b$ at $\nu = \frac{1}{2R_0} + \frac{\pi}{2} = 1.73824$, which is the point where

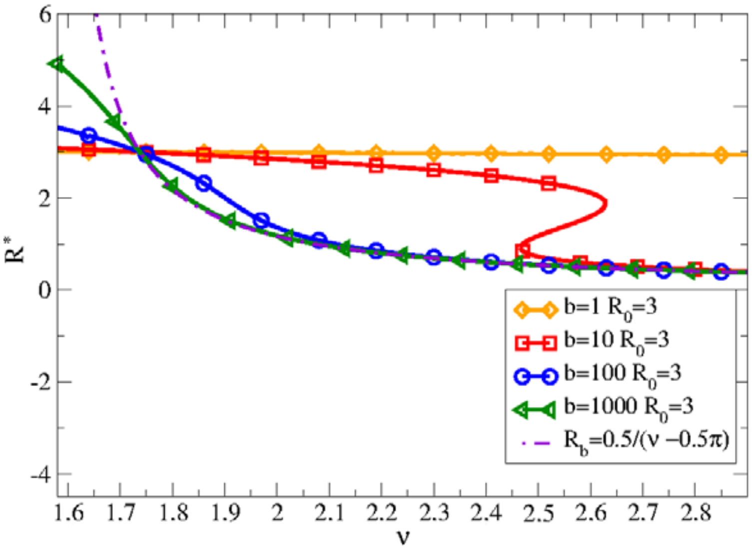


Figure 4.4: Real and positive solutions of R versus ν for $R_0 = 3$ and different b values. All solutions meet at $\nu = \frac{1}{2R_0} + \frac{\pi}{2}$ corresponding to the solution where $R_b = R_0$.

$R_0 = R_b$ (hence being R^* optimal both for stretching and bending).

Fig.4.5 plots the value of R^* as a function of b, for $R_0 = 3$ and different values of ν. As expected, for small b ($b << 1$), R^* tends to the preferred stretching radius R_0. In the opposite limit of $b >> 1$, each solution tends to the bending-determined radius $R_b = \frac{1}{2\left(\nu - \frac{\pi}{2}\right)}$. Remarkably, for values of $\nu < \nu^* = \frac{1}{2R_0} + \frac{\pi}{2}$, there is only one solution, whereas 3 solutions coexist for $\nu > \nu^*$ and intermediate values of b.

Finally, Fig.4.6 shows how R^* depends on R_0 for an intermediate value of $b = 100$ and different preferred angles ν. On one hand, for values of $\nu < \nu^* = \frac{1}{2R_0} + \frac{\pi}{2}$, there is only one solution (blue line) that approaches the stretching-dominated limit of $R^* = R_0$ (yellow line) at large values of R_0. On the other hand, for values of $\nu > \nu^* = \frac{1}{2R_0} + \frac{\pi}{2}$ (green and red lines) there is only one solution corresponding to $R^* = R_b$ for small values of R_0, and there are three solutions for large values of R_0.

4.2.4 Optimal size and metastable solutions

As mentioned in the previous subsection, the quartic equation for the optimal radius, eq.4.13, admits either only one or three real and positive solutions, depending on the values of the parameters. In the cases where there is only one solution, it corresponds to the optimal value of the radius, that tends to be R_0 for $b << 1$ and R_b for $b >> 1$. But there are regions of parameters where there are 3

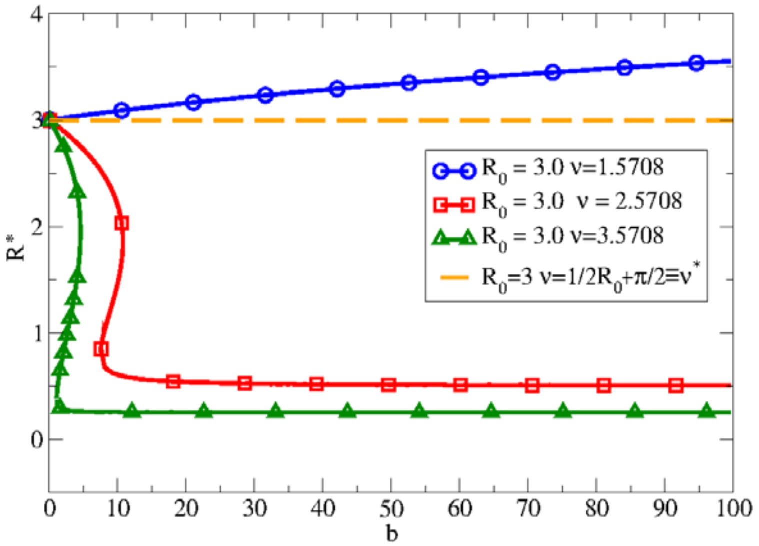


Figure 4.5: Real and positive solutions of R versus b for $R_0 = 3$ and different ν values.

solutions and it is worth to analyze them more carefully.

Fig.4.7 shows the behavior of the energy as function of R for $\nu = 2.5708$ and different values of R_0. By calculating the second derivative, evaluated at each solution $R^*_{1,2,3,4}$, it can be seen that the 3 real and positive solutions, correspond to two minima, separated by a maximum. Thus, out of the 3 solutions, 2 correspond to local minima of the the free energy that are located close to R_0 and R_b. The third solution corresponds to a critical value of the radius that signals the maximum of the free energy barrier that has to be overcome to go from one of the minima to the other.

The energy at the minimum around R_b can be approximated by the stretching energy $E(R_b) \sim \frac{1}{2}(R_b - R_0)^2$ and the energy of the other minimum is approximately the bending energy, $E(R_0) \sim \frac{b}{2}\left(\frac{1}{2R_0} + \frac{\pi}{2} - \nu\right)^2$. The difference of energy between the two minima is then $\delta E_{min} = E(R_0) - E(R_b) \sim \frac{1}{2}\left(b\left(\frac{1}{2R_0} + \frac{\pi}{2} - \nu\right)^2 - \left(\frac{-1}{2cos(\nu_0)} - R_0\right)^2\right)$ and determines which of the 2 solutions is the global minimum. The barrier energy is $E_{barr} = E(R_{max}) - E(R_{min})$, where R_{max} is the value of the solution corresponding to the maximum and R_{min} is one of the two solutions. The barrier of energy depends strongly on the parameter b, as shown in Fig.4.8.

Mathematically, the transition between having one or three solutions occurs when the discriminant is equal to zero, i.e. $\Delta = 0$, in eq.4.14. From this condition we can find the critical values of b, labeled as b_c, for which this occurs. Ignoring

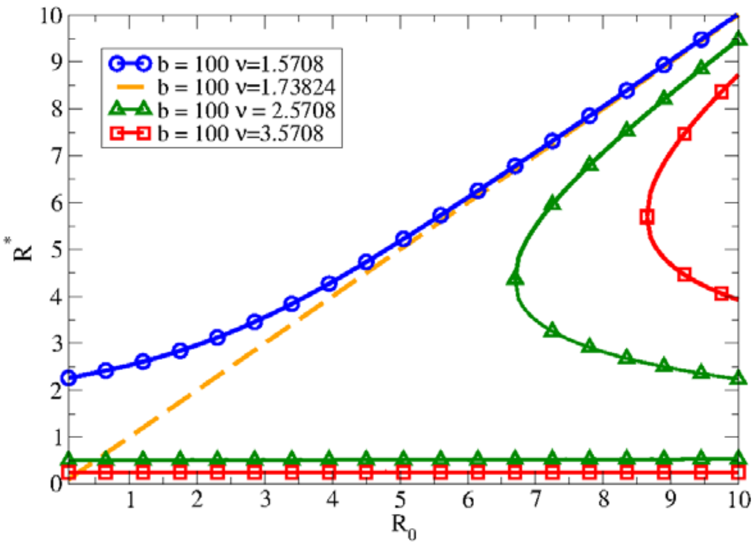


Figure 4.6: Real and positive solutions of R versus R_0 for $b = 100$ and different ν values.

R_0	ν	b_c^1	b_c^2	line
3	2.5708	7.50845	10.7878	red
4	2.5708	10.9138	23.4566	yellow

Table 4.1: Critical values of b calculated using eq. 4.15 for $R_0 = 3$ (red line with squares) and 4 (yellow line) illustrated in Fig. 4.7, by the discontinuous lines, where three solutions for the radius are possible.

the trivial solution $b = 0$, the critical values are:

$$b_c^{1,2} = \frac{1}{27(\nu-\pi/2)^4}\Bigg\{ -32 \pm 2\sqrt{2} \sqrt{\left(R_0(\nu-\pi/2)-2\right)^3 \left(R_0(\nu-\pi/2)+4\right)^2 \left(2R_0(\nu-\pi/2)-1\right)} + R_0(\nu-\pi/2)\left(48 + R_0(\nu-\pi/2)(3+4R_0(\nu-\pi/2)\right)\Bigg\} \tag{4.15}$$

So there will be three solutions for R^*, in the range of $b's$ defined by $b_c^1 < b < b_c^2$, as shown in Fig. 4.7. The presence of a square root in eq. 4.15 also imposes a limitation on the value of the other two parameters, ν and R_0, to have a real solution for b_c. This limitation is the condition $R_0 > 4R_b = \frac{2}{\nu-\pi/2}$. That is the reason why, in Fig. 4.7, there is only one solution for $R_0 = 1$ and $R_0 = 2$ (blue and green line, respectively). For $R_0 = 3$ and $R_0 = 4$, there are three solutions.

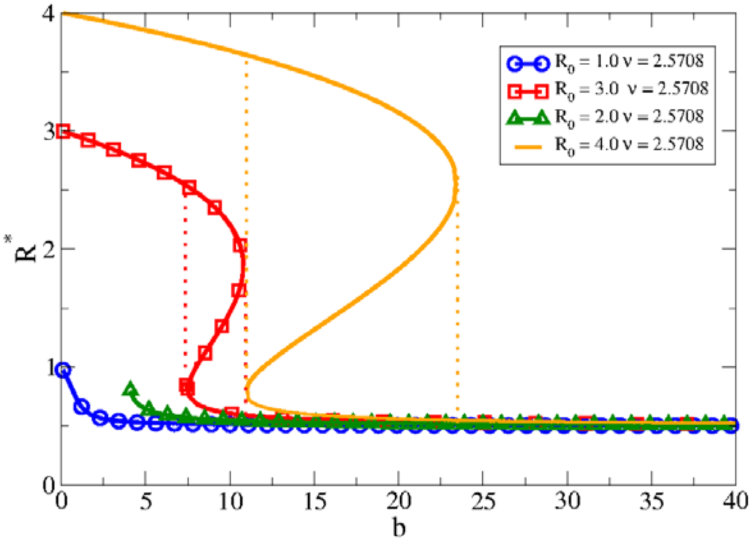
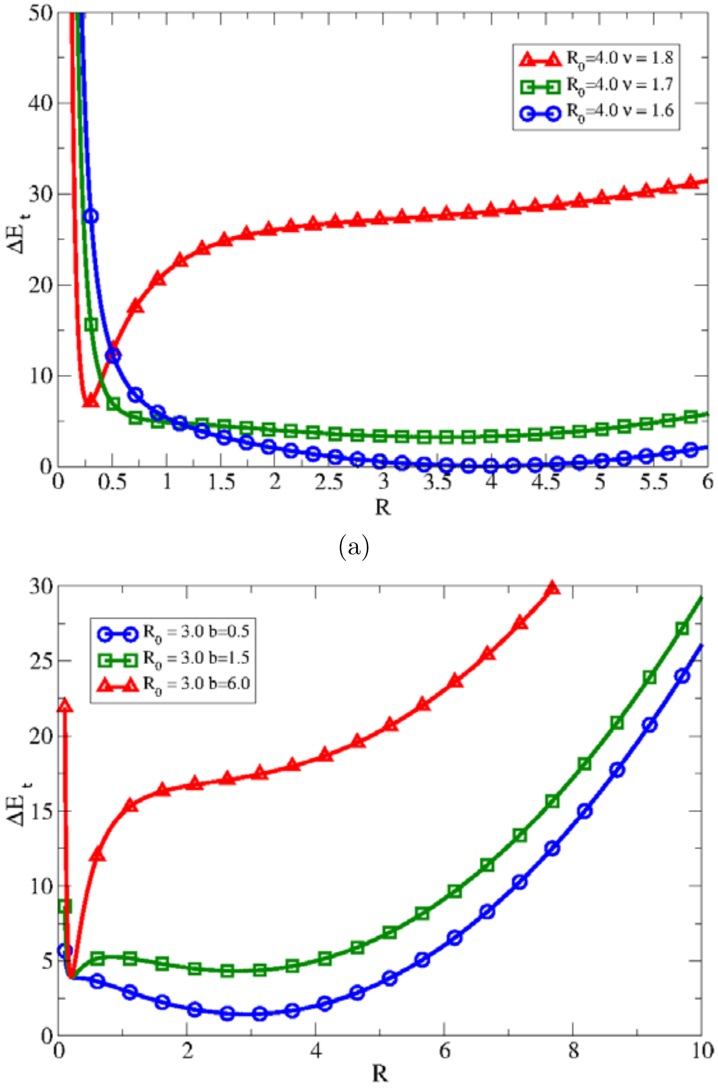


Figure 4.7: Real and positive solutions of R^* versus b for $\nu = 2.5708$ and different R_0 values. The discontinuous lines define the region in values of b where there are 3 different solutions for $R_0 = 3$ (red lines) and $R_0 = 4$ (yellow lines).

Using eq. 4.15 we have calculated the two values b_c^1 and b_c^2 delimiting the region where three real and positive solutions exists, see Table 4.1.

With these critical points we can identify fully the regions where metastable solutions are possible in terms of the spontaneous curvature and the number of capsomers of the viral capsid that ultimately sets the value of R_0. This metastable region is characterized by a strong competition between the bending and stretching energies, see Fig. 4.8. Then, in these regions we expect possible transitions in size that would influence other mechanical properties.



(a)

(b)

Figure 4.8: Energy landscape in the region where there are three solutions of R^* for (a) $b = 20$ and $R_0 = 4$ and different values of ν and (b) for $\nu = 2.7508$, $R_0 = 3$ and different values of b.

4.3 Conclusions

In this chapter, we have presented a very simple elastic model which can be considered as the continuum limit of the coarse-grained interaction potential of Chapter 3 for a complete capsid. By minimizing the elastic energy, we have obtained the value of the optimal radius of the capsid. Using this elastic model, we have obtained and characterized the different regions of physical parameters, where the bending and stretching contributions dominate. Interestingly, we have obtained a region of parameters where the two contributions strongly compete. In this region there are three possible solutions, corresponding to two local minima of energy separated by one maximum. Thus in this region a transition between two different locally optimal radii of the capsid is expected.

This elastic model, provides a very useful starting point to explain the competition between bending and stretching effects. However, this simplified model has some clear limitations. First, the main drawback of this continuum model is that it lacks structural details. The discrete and anisotropic interactions between subunits or proteins cannot be easily incorporated in this type of descriptions [16]. This fact difficults the characterization of plastic or breaking behaviors observed in nanoindentation experiments [17–19].

Another limitation is that we are restricted to uniform radial deformations. So no change in shape can be described with this model. A final limitation is that we have not contemplated the real possibility that capsids can only sustain a limited expansion/contraction. When the expansion exceeds the inflection point of the intercapsomer potential, capsid will break. Therefore the solution corresponding to R_b may in some cases be impossible to reach due to the breaking of the capsid.

Appendix A

Solution of the quartic equation for the optimal radius

The total energy of a capsid is a combination of a bending contribution and stretching contribution and in order to minimize the energy we must solve [20]:

$$\begin{aligned}
&\frac{dE_t}{dR}\Big|_{R=R^*} = 0 \Rightarrow R^* - R_0 + b(\frac{\pi}{2} + \frac{1}{2R^*} - \nu)(\frac{-1}{2R^{*2}} - \frac{d\nu}{dR^*}) = 0 \\
&R^{*4} - R_0R^{*3} + \frac{b}{2}(\nu - \frac{\pi}{2})R^* - \frac{b}{4} = 0
\end{aligned} \tag{4.16}$$

The general solution of the quartic equation $x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0$, is commonly expressed in terms of the four solutions of equation:

$$z^2 + \frac{1}{2}(a_1 \pm \sqrt{a_1^2 - 4a_2 + 4y_1})z + \frac{1}{2}(y_1 \mp \sqrt{y_1^2 - 4a_4}) = 0 \tag{4.17}$$

where y_1 is

$$y_1 = S + T + \frac{1}{3}a_1 \tag{4.18}$$

In the previous expression,

$$\begin{aligned}
S &= (R + \sqrt{Q^3 + R^2})^{1/3} \\
T &= (R - \sqrt{Q^3 + R^2})^{1/3}
\end{aligned} \tag{4.19}$$

where $Q = \frac{-4a_4}{3}(1 - a_3)$ and $R = \frac{27a_4}{54}\left(a_1^2 - \frac{a_3^2}{a_4}\right)$.

In our case we can identify

$$\begin{aligned}
a_1 &= -R_0 \\
a_2 &= 0 \\
a_3 &= \frac{b}{2}(\nu - \frac{\pi}{2}) \\
a_4 &= -\frac{b}{4}
\end{aligned} \tag{4.20}$$

Accordingly, Q and R are given by:

$$\begin{aligned} Q &= \frac{b}{3}\left(1-\frac{bR_0}{2}\left(\nu-\frac{\pi}{2}\right)\right) \\ R &= \frac{-27}{54}\frac{b}{4}\left(b\left(\nu-\frac{\pi}{2}\right)^2+R_0^2\right) \end{aligned} \tag{4.21}$$

where $\Delta = Q^3 + R^3$ is the discriminant. The real solution of y_1 then becomes explicitly:

$$\begin{gathered} y_1 = \left\{ \left(\frac{-b}{8}R_0^2 - b\left(\nu-\frac{\pi}{2}\right)^2\right) + \right. \\ \left. + \sqrt{\frac{b^3}{27}\left(1-\frac{R_0}{2}\left(\nu-\frac{\pi}{2}\right)\right)^3 + \frac{b^2}{64}\left(R_0^2 - b\left(\nu-\frac{\pi}{2}\right)^2\right)^2}\right\}^{\frac{1}{3}} + \\ + \left\{ \left(\frac{-b}{8}R_0^2 - b\left(\nu-\frac{\pi}{2}\right)^2\right) - \right. \\ \left. - \sqrt{\frac{b^3}{27}\left(1-\frac{R_0}{2}\left(\nu-\frac{\pi}{2}\right)\right)^3 + \frac{b^2}{64}\left(R_0^2 - b\left(\nu-\frac{\pi}{2}\right)^2\right)^2}\right\}^{\frac{1}{3}} \end{gathered} \tag{4.22}$$

By substituting y_1 in eq. 4.17, and solving for $z = R^*$, we finally obtained the four solutions of the quartic eq. 4.17. Explicitly:

$$\begin{aligned} R^*_{1,2} &= \frac{-1}{4}\left[-R_0 + \sqrt{R_0^2 - 4y_1}\right] \pm \\ &\pm \frac{1}{2}\left\{\frac{1}{2}\left(R_0^2 - 2y_1 - R_0\sqrt{R_0^2-4y_1}\right) - 2\left(y_1 - \sqrt{y_1^2 - b}\right)\right\}^{\frac{1}{2}} \end{aligned} \tag{4.23}$$

$$\begin{aligned} R^*_{3,4} &= \frac{-1}{4}\left[-R_0 - \sqrt{R_0^2 - 4y_1}\right] \pm \\ &\pm \frac{1}{2}\left\{\frac{1}{2}\left(R_0^2 - 2y_1 + R_0\sqrt{R_0^2 - 4y_1}\right) - 2\left(y_1 - \sqrt{y_1^2 - b}\right)\right\}^{\frac{1}{2}} \end{aligned} \tag{4.24}$$

Replaced eq. 4.22 in eq. 4.23 and eq. 4.24 we have obtained the four solution in function of ν_o, R_0 and b.

References

[1] L. D. Landau and I. M. Lifshitz, Theory of elasticity. Pergamon, New York, 1975.

[2] J. Lidmar, L. Mirny, and D. Nelson, Virus shapes and buckling transitions in spherical shells, *Phys. Rev. E*, **68**, 051910, 2003.

[3] T. Nguyen, R. Bruinsma, and W. Gelbart, Continuum theory of retroviral capsids, *Phys. Rev. Lett.*, **96**, 078102, 2006.

[4] T. Guerin and R. Bruinsma, Theory of conformational transitions of viral shells, *Phys. Rev. E*, **76**, 061911, 2007.

[5] B. Stephanidis, S. Adichtchev, P. Gouet, A. McPherson, and A. Mermet, Elastic properties of viruses, *Biophys. J.*, **93**, 1354-9, 2007.

[6] F. Tama and C. L. Brooks, Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis, *J. Mol. Bio.*, **345**, 299-314, 2005.

[7] T. Nguyen, R. Bruinsma, and W. Gelbart, Elasticity theory and shape transitions of viral shells, *Phys. Rev. E*, **72**, 051923, 2005.

[8] A. Y. Morozov and R. F. Bruinsma, Assembly of viral capsids, buckling, and the Asaro-Grinfeld-Tiller instability, *Phys. Rev. E*, **81**, 041925, 2010.

[9] Z. Yang, I. Bahar, and M. Widom, Vibrational dynamics of icosahedrally symmetric biomolecular assemblies compared with predictions based on continuum elasticity, *Biophys. J.*, **96**, 4438-48, 2009.

[10] A. Siber, Buckling transition in icosahedral shells subjected to volume conservation constraint and pressure: Relations to virus maturation, *Phys. Rev. E*, **73**, 061915, 2006.

[11] M. Widom, J. Lidmar, and D. Nelson, Soft modes near the buckling transition of icosahedral shells, *Phys. Rev. E*, **76**, 031911, 2007.

[12] WS. Klug, RF. Bruinsma, JP. Michel, CM. Knobler, Failure of viral shells, *Phys. Rev. Lett.*, **97**, 228101, 2006.

[13] U. Seifert, Configurations of fluid membranes and vesicles, *Adv. Phys.*, **46**, 1, 13-137, 1997.

[14] S. Schneider and G. Gompper, Shapes of crystalline domains on spherical fluid vesicles, *Europhys. Lett.*, **70**, 1, 136-142, 2005.

[15] J. Camps and R. Emparan, Derivation of the blackfold effective theory, *JHEP*, **1203**, 038, 2012.

[16] C. Carrasco, A. Carreira, I. A. T. Schaap, P. A. Serena, J. G.-H. M. G., Mateu, and P. J. de Pablo, DNA-mediated anisotropic mechanical reinforcement of a virus, *PNAS*, **103**, 13706-11, 2006.

[17] B. Böttchera, M. Vogelb, M. Plossa and M. Nassal, High Plasticity of the Hepatitis B Virus Capsid Revealed by Conformational Stress, *J. Mol. Biol.*, **356**, 812-822, 2006.

[18] IL. Ivanovska, R. Miranda, JL. Carrascosa, GJL. Wuite and CF. Schmidt. Discrete fracture patterns of virus shells reveal mechanical building blocks, *PNAS*, **108**, 31, 12611-12616, 2011.

[19] A. Ortega-Esteban, AJ. Pérez-Berná, R. Menéndez-Conejero, SJ. Flint, C. San Martín and PJ. de Pablo, Monitoring dynamics of human adenovirus disassembly induced by mechanical fatigue, *Sci. Rep.*, **3**, 1434, 2013.

[20] MR. Spiegel, J. Liu and L. Abellanas, Mathematical handbook of formulas and tables, McGraw-Hill, U.S.A, 2000.

CHAPTER 5

Simulation methods

5.1 Introduction

Nowadays, simulations are becoming an essential tool to study and understand better a wide variety of problems in physics, chemistry, biology and complex systems. With simulations it is possible to mimic real processes and systems using computer models. That has facilitated the investigation of scientific questions that would be otherwise expensive or impossible to study in experiments and/or theory. Another interesting area where simulations are becoming specially helpful is in the field of biophysics. Important problems such as protein folding [1], the hydration of proteins and hydrophobic interactions [2, 3] or the elastic properties of membranes [4] are an intense field of study by simulations.

In this chapter we will explain some technical details about the simulation methods used in this thesis to investigate the model described in Chapter 3. Essentially, three different simulation techniques have been implemented. Monte Carlo (MC) simulations have been used for the study of the optimal structures and stability of viral shells. On the other hand, Brownian Dynamics (BD) simulations will be used to analyze the dynamical aspects of the assembly and the mechanical response of viral capsids at the discrete level. To complement this mechanical description at the continuous level, we have also used Finite Element (FE) simulations.

The chapter is distributed as follows. In Section 5.2 we will discuss the most relevant details of the Monte Carlo simulation using the coarse-grained model described in Chapter 3. After that, Section 5.3 will be devoted to describe the Brownian Dynamics method used to study the dynamics of capsid assembly and for the implementation of our "Virtual AFM" in Chapter 10. Then Finite Elements analysis will be introduced in section 5.4, as a tool for the study of some mechanical properties of the capsid, which will be presented in Chapter 9. Finally, we will discuss the advantages and limitations of all the methods in the study of the different properties of viruses.

5.2 Monte Carlo simulations

Monte Carlo methods (MC) are a class of computational algorithms that rely on repeated random sampling to obtain numerical results. This method was introduced by Newmann, Ulam and Metropolis in the forty's [5], to study neutron diffusion in fissionable materials.

MC methods are often used in physical and mathematical problems and they are most suited to be applied when it is impossible to obtain a closed-form expression or a deterministic algorithm. These methods are mainly used in three types of problems: numerical integration, optimization, and generation of samples from a probability distribution [6].

The MC method is very useful to compute high dimensional and definite integrals. While other algorithms usually evaluate the integral at a regular grid, MC algorithms randomly choose the points at which the integral is evaluated in a domain [7]. This algorithm is common in optimization problems too, where one must minimize, or maximize, functions of some vector that often has a large number of dimensions. Many other problems can be phrased in terms of sampling random points, where the goal is to minimize distances [8].

The third type of problem where MC simulations are used is in statistical mechanics [6]. The general motivation to use the MC method in statistical mechanics is to evaluate a multivariable integral representing the average value of a given thermodynamic quantity where the Hamiltonian is known. To obtain the mean value of some macroscopic variable, the idea is to compute, over all the phase space, the mean value according to the relevant statistical weights (i.e. Boltzmann statistics) using the MC method to solve the multivariable integral.

The MC method is intimately related with the generation of random numbers, because this method is based on randoms changes in the system following some arithmetic or logical operation. As a consequence the method has become progressively more efficient with the improvement in the generation of random numbers. These methods of simulations are subject to statistical and systematic errors. Statistical errors control the error bars associated with each measurement. In the Monte Carlo method, the absolute statistical error, decreases with the number of steps or evaluations as $\frac{1}{\sqrt{N}}$ [6].

In a typical MC simulation, the starting point is an initial configuration of particles in a system. Then, a random move is proposed that changes the configuration of the particles. This move is accepted or rejected based on an acceptance criterion that guarantees that configurations are sampled in the simulation with the correct weight according to a statistical mechanics ensemble distribution [6]. After the acceptance or rejection of a move, one calculates the value of a property of interest, obtaining an average value of this property after repeating this procedure a large number of times or “steps”.

In our particular case we have implemented a simple Monte Carlo simulation using the Metropolis algorithm to obtain the equilibrium capsid structures and study their stability. The equilibrium structures in our model are characterized by the energy of the system, which will be used in the acceptance criterion. The total energy in a system containing N model capsomers is given by the expression:

$$E_t = \sum_{i=1}^{N-1} \sum_{j>i}^{N} V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j) \tag{5.1}$$

where $V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)$ is the potential between capsomer i and j given by eq. 3.2. In the potential the interaction is by pairs, and each subunit interacts in principle with all the rest of capsomers.

There are mainly two different strategies to move the subunits in MC simulations; essentially, the update and trial moves can be done randomly or sequentially [6]. The most common strategy is to randomly select which particle will be displaced, but in some cases where the interaction between neighbor subunits is strong, for example in chains, or in atomic model [10] (in order of atom index), this technique will be more appropriate. In our case, we have chosen a random selection of subunits to update the configuration.

In our typical MC simulations, we start from an initial configuration corresponding to a perfect T-number structure, having an energy E_1. Then, one capsomer is randomly selected to attempt a displacement. In our case, two different types of displacement are considered: spatial moves, that attempt to change the coordinates x, y, z by a small random displacement $(2\xi - 1)\Delta r_{max}$, where ξ is a random number between 0 and 1; and orientational moves, where the orientation of the capsomer described by the angles θ, ϕ is altered by an amount $(2\xi - 1)\Delta\varphi_{max}$.

The attempted movement is governed by a transnational and orientational maximum displacements Δr_{max} and $\Delta\varphi_{max}$. If the maximum displacement is too small, a big number of moves are accepted, but the convergence to the optimal configuration is very slow and the states are highly correlated. On the other hand, if the maximum displacement is too large, only a small fraction of moves are accepted and the state is nearly frozen, close to the initial configuration. To minimize this problem the value of the maximum displacement is adjusted every 10000 steps in order to keep the acceptance ratio close to 0.5.

After the random displacement, the new energy of the system, E_2, is calculated. If the new energy (E_2) is lower than the old one E_1, the movement is always accepted; and if it is higher, the movement is still accepted with some Boltzmann probability, see Fig. 5.1, given by.

$$P_0 = e^{-(E_2 - E1)/k_B T} \tag{5.2}$$

where T is the temperature of the system and k_B is Boltzmann's constant. Applying this procedure repeatedly ensures the minimization of the energy [6].

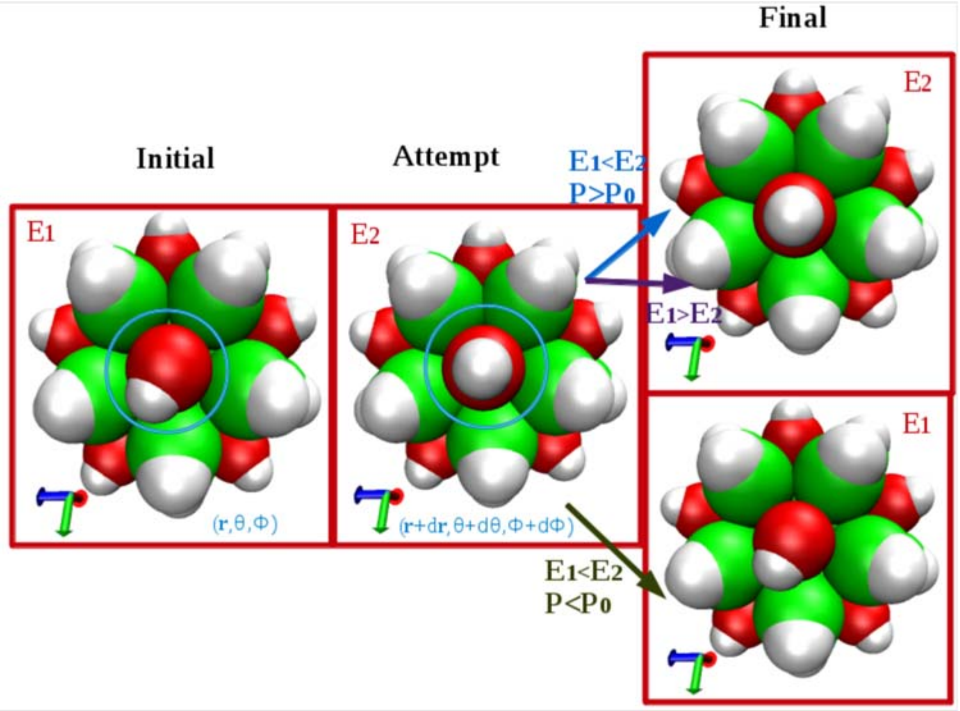


Figure 5.1: Illustration of the acceptance of attempted displacements in a MC simulation. The energy of the initial configuration is E_1, and after a random displacement of the highlighted capsomer becomes E_2. If $E_2 < E_1$, the move is always accepted. If $E_2 > E_1$, the movement is accepted with a probability P_0.

In all MC simulations we have used periodic boundary conditions to minimized finite size and boundary effects.

The Monte Carlo method is appropriate mainly for the study of equilibrium properties of viral capsids. In principle, the optimal structure would be the one that minimizes the free energy. But the free energy of a system is difficult to evaluate. In the case of viral shells the system is dominated by the strong binding interaction between subunits. For that reason we have used and reported the average potential energy of the viral capsid and not strictly the free energy.

It is worth to note that there are other methods which are more efficient to minimize the energy and to find global optima, in complex systems, such as simulated annealing, parallel tempering or the mapping of the complete free energy landscape [12]. However, for systems with a low number of subunits, as the ones that will be the focus of our work, a simple Metropolis algorithm is a good and simpler alternative.

5.3 Brownian Dynamics

One of the goals of this thesis is to characterize the dynamics of viral assembly, and to that end dynamic simulation techniques are more suited than equilibrium techniques as MC. Dynamic simulations can keep track of the movement of subunits as a function of time, which is calculated by solving the corresponding equations of motion.

There are mainly two different simulations techniques for the study of the dynamics: Molecular Dynamics (MD) and Brownian Dynamics (BD) [13]. In MD simulations, given the interaction, Newton's equations of motion of every particle in the system are integrated to calculate their trajectories. This method of simulation is highly realistic because all atoms and molecules, including the solvent, are simulated [10]. However, for the case of viral assembly the long time scales, the presence of the solvent and the large amount of atoms involved even in the smallest viruses, makes this technique computationally inapplicable.

BD simulations can be used as a more efficient alternative to MD to describe complex systems at larger times scales. In this method, Newton's equations of the system are replaced by some simplified version of Langevin dynamics, where the solvent is replaced by a combination of random forces and frictional terms [13]. In addition, for nanoscale biological objects such as proteins, inertia and the associated acceleration term become irrelevant, so the dynamics is overdamped. This approximation leads to even simpler equations of motion [13].

A very simple algorithm that computes the forces only once per time step and is capable of integrating overdamped systems is the stochastic Euler algorithm. In this algorithm, the position of a particle at a time $t + dt$ is given by

$$r_i(t + dt) = r_i(t) + \frac{F_i(t)dt}{\eta} + \sqrt{2Dt}\xi \tag{5.3}$$

where η is the friction coefficient, D is the diffusion coefficient, $F_i(t)$ is the force on particle i, and ξ is a Gausssian distributed random number with zero mean and variance 1. Thus, in the stochastic Euler algorithm, the new position of a given particle depends on the force and a noise term that accounts for random collisions with the solvent.

It is very often convenient to simplify the simulations using reduced units. In our case, we will use σ (the diameter of a capsomer, see Chapter 3) as our unit of length, ϵ_0 as the unit of energy, and $t_0 = \frac{\sigma^2}{D}$ as our time unit. For a typical capsomer size of about $\sigma = 10nm$, and using the Stokes-Einstein [14] equation for the diffusion coefficient in water $D = \frac{k_B T}{\eta}$, we obtain a time unit of $t_0 = 4.6\mu s$. Using these reduced units, the stochastic Euler equation can be rewritten as

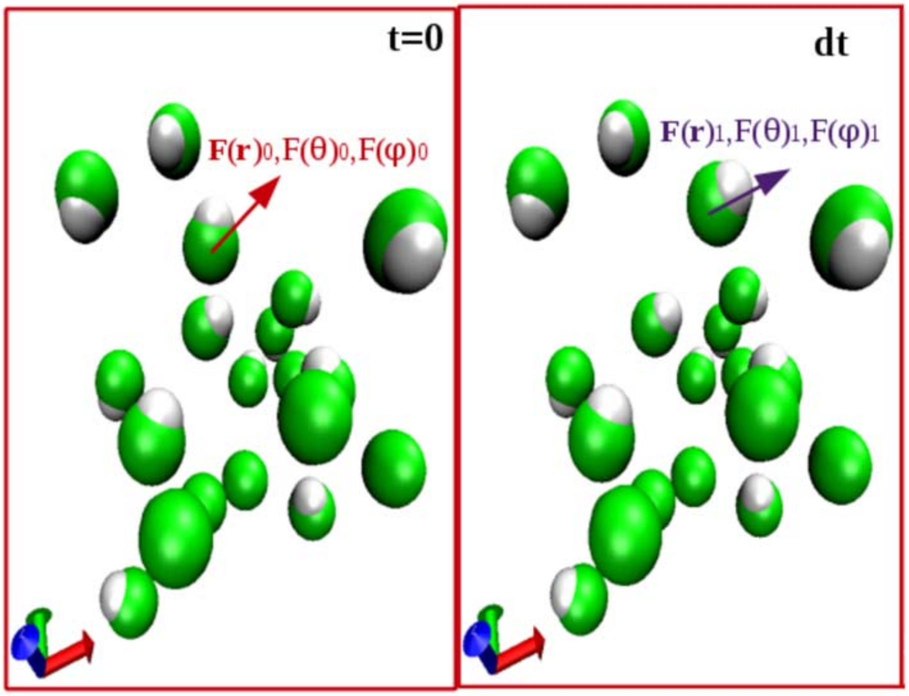


Figure 5.2: Schematic representation of the motion of one capsomer in a Brownian Dynamic simulation. At $t = 0$, the force exerted by all capsomers on capsomer i are calculated. The new position at $t = dt$ is calculated from this force using eq. 5.5.

$$r_i(t+dt) = r_i(t) + \frac{F_i(t)dt}{T} + \sqrt{2dt}\xi \tag{5.4}$$

In a typical BD simulation of assembly, we start by placing N capsomers at random positions and orientations inside our simulation box. We then calculate the total force felt by each capsomer $\overrightarrow{F_i}(t)$. In the next step of simulation corresponding to time $t + dt$, the new positions $\mathbf{r}(t)$ and angles ($\theta(t)$ and $\phi(t)$) for each capsomer are calculated using the following equations, see Fig. 5.2.

$$\begin{aligned}
\mathbf{r}(t+dt) &= \tfrac{dt}{T}\overrightarrow{F_r}(r) + \sqrt{2dt}\overrightarrow{\xi_r} + \mathbf{r}(t) \\
\theta(t+dt) &= \tfrac{dt}{2T}F_\theta(t) + \sqrt{dt}\xi_\theta + \theta(t) \\
\phi(t+dt) &= \tfrac{dt}{2T}F_\phi(t) + \sqrt{dt}\xi_\phi + \phi(t)
\end{aligned} \tag{5.5}$$

where dt is the time step $\overrightarrow{\xi_r}$, ξ_θ, and ξ_ϕ are Gaussian distributed random numbers with zero mean a variance 1. The time step dt has to be small enough so that a subunit cannot travel a distance comparable to σ at each time step. In our simulations we have found that a reasonable value is $dt = 10^{-4}$, in reduced units.

The translational and orientational forces $\overrightarrow{F_r}(t)$, $F_\theta(t)$ and $F_\phi(t)$ are calculated from our model potential between capsomers, $V_{ij}(r_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)$ described in Chapter 3. The explicit expression of the forces and how they were calculated is explained in Appendix A.

As in our Monte Carlo simulations, we have been implemented periodic boundary conditions to minimize finite size effects and mimic more closely an infinite system [15]. We have built cubic boxes replicated through the total space. Using the boundary conditions, the molecules move in the original box, where the periodic images in each of the neighbor boxes moves in the same way. Implementing boundary condition one capsomer leaving to the central box its image will enter by the opposite face. This tool reproduce the topology of the system, without the problem from the surface effects.

BD simulations, as all simulation methods, involves some limitations.The size of the simulation box, the number of particles, the time-step and total time duration must be selected property so that the calculation can finish within a reasonable time period and with reasonable error. Even working with coarse-grained models, the accessible time scales and system sizes are typically smaller that for a real situation. For example for a typically simulation run with $t_0 = 4.6\mu s$, $dt = 10^{-4}t_0$ and 10^9 steps, one can reach only $0.46s$ of real time. Thus, the simulations described in this section can mimic the motion of viral capsids and its subunits, but this method is limited to relatively fast processes. It is possible study the viral self-assembly at relatively strong supersaturations, but larger or more efficient simulations would be needed for other slower viral process. In the specific case of viruses, with our BD simulations we can study the general dynamics and intermediate steps in self-assembly viral capsids.

5.4 Finite Element Methods

The Finite Element Method (FEM) is a mathematical method which was born by the necessity to solve the complex elasticity and structural analysis problems in civil and aeronautical engineering. The idea, by A. Hrennikoff and R. Courant [16], was to simplify the study of a continuous domain (described by complex equations) by a set of discrete sub-domains (with more simple equations), usually called elements.

The FEM began to be used in the 70's, when NASA and UC Berkeley sponsored the first studies of the implementation of the FEM method using computa-

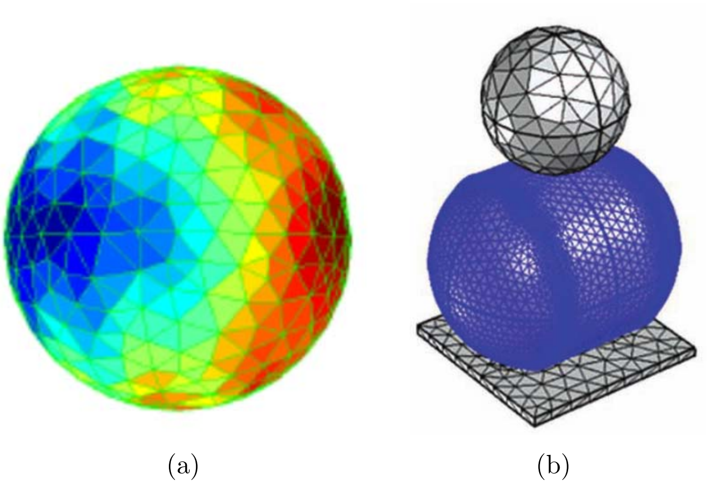

(a) (b)

Figure 5.3: Finite element mesh using triangles to study mechanical properties a) of a thin spherical shell and b) a thin spherocylinder resembling an Atomic Force Microscopy nanoindentation experiment on ϕ29 bacteriophage.

tional software. Strang and Fix [17] published a rigorous mathematical basis to the finite element method. In this way, this method has since then been generalized for the numerical modeling of physical and engineering systems in disciplines such as electromagnetism [18], heat transfer [19] mechanics of solids and fluid dynamics [19].

The key idea of FEM is to approximate the partial derivatives by finite differences. The differential equations of the initial problem are thus simplified into a set of lineal equations, where the correct result can be obtained for a finite number of points (nodes) and then interpolated to the rest of the domain. The set of nodes is obtained by dividing or discretizing the system into various small elements (for instance triangular areas or tetrahedral volumes).

In general, the determination of the mechanical properties of viral capsids is a difficult problem to solve because the interactions and the internal structure of the protein subunits could be extremely complex. However, capsids are made by a large number of atoms and molecules. Therefore one could expect that, collectively and as a first approximation the capsid can be treated as a continuum shell.

The mechanical properties of viruses can then be described approximately by continuum elasticity theory. With this approximation the system is governed by the Navier equations of elasticity. For arbitrary geometries and deformations, these equations, even in the thin shell limit, are impossible to solve analytically. This is the reason why FE methods are commonly used to solve numerically these equations.

Continuum elasticity has been used to model mechanical properties of many viruses [20–23, 25–28]. In our work, we have used FEM to model the mechanical response of viruses, in particular bacteriophage T7 and $\phi29$, to interpret the nanoindentation experiments carried out by Atomic Force Microscopy. In the particular case of $\phi29$ bacteriophage, it has a prolate shape with an architecture $T_{end} = 3$ and $Q_{5F} = 5$ centered on a 5-fold axis [29, 30], see Fig. 5.3(b). The prolate shape of this virus makes difficult to find an analytic solution of the force versus indentation curves. Hence, we opted to implemented FEM simulations, as described in detail in chapter 9.

In the study of the mechanical properties of viral shells we have used the commercial software COMSOL Multiphysics 4.2a (Comsol, Stockholm, Sweden). This program provides a powerful interactive environment for modeling and solving all kinds of scientific problems based on partial differential equations. With this tool the solution of the elasticity equation is performed by the discretization of the space and time using different methods based on finite differences or finite elements. This converts the nonlineal differential equations into a set of discrete dynamic equations that can be solved by different algorithms that try to minimize errors and potential instabilites in the solution.

5.5 Conclusions

Computer simulations have become nowadays a powerful tool to solve complicated analytical problems in natural sciences and to mimic experimentally inaccessible conditions.

The choice of an appropriate simulation method for each problem and area of study is fundamental. The selection of the most efficient or suitable technique depends on the nature of the system, the desired level of description and the kind of information that is sought for.

The implementation of simulations with high level for details is often not necessary for the a study of some general properties of system and it could lead to an expensive waste in computational resources. It is important to get a compromise between realism and insight or simplicity. Using coarse-grained model and simple but efficient simulation methods we can learn many helpful information and get some general information.

In this thesis, we have implemented three different simulation techniques to investigate different aspects of viruses. For the study of optimal structures and stability the most important element is finding the structure of minimal energy and for this type of studies Monte Carlo simulations are particularly. On the other hand, this method cannot realistically the kinetics of viral self-assembly. To that end, we have rather opted to use Brownian Dynamic simulations. Finally, in the study of the mechanical deformation of viruses, we have used a mixed approach that combined BD simulations to analize discrete effects, with FE methods to study the macroscopic continuum response.

In next chapters we will apply all these different simulation methods to the study of viruses in the framework of our coarse-grained model. We will see that these thecniques will provide information on the stability, polymorphism, maturation, self-assembly, AFM experiments, etc, with the hope to get some helpful information.

5.6 Appendix A

Expressions for the force used in the BD simulations

In general, the generalized forces are calculated from the gradient of the potential $\overrightarrow{F} = - \bigtriangledown V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)$, where $\overrightarrow{F} = (F_x, F_y, F_z, F_\theta, F_\phi)$ and $\overrightarrow{\bigtriangledown} = (\partial_x, \partial_y, \partial_z, \partial_\theta, \partial_\phi)$. We have calculated the forces which we need for the BD simulations using the interaction potential $V(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)$ described in chapter 3. At distances ($r < 2^{\frac{-1}{6}}\sigma$), capsomers interact only via the Lennard-Jones potential, which is repulsive for these distances. In this case ($r < 2^{\frac{-1}{6}}\sigma$) the force is given by:

$$F_r = -\frac{dV_{LJ}(r)}{dr} = \frac{-12\epsilon_0}{r}\left[\left(\frac{\sigma_0}{r}\right)^{12} - \left(\frac{\sigma_0}{r}\right)^{6}\right] \tag{5.6}$$

In these conditions the spatial components of the force over capsomer i are $\overrightarrow{F}^r_{ij} = F_r \frac{\overrightarrow{r}_{ij}}{r^2}$ where $\mathbf{r}_{ij} = (x_j - x_i, y_j - y_i, z_j - z_i)$ and $\overrightarrow{F}^r_{ij} = (F_x, F_y, F_z)$. Obviously $\overrightarrow{F}^r_{ij} = -\overrightarrow{F}^r_{ji}$.

For distance between capsomers in the range $[2^{\frac{-1}{6}}\sigma_0, 5\sigma_0]$ the attractive part of the Lennard-Jones is modulated by an angular contribution that depends on the orientation $\overrightarrow{\Omega}_i$ and $\overrightarrow{\Omega}_j$ of the capsomers. In this case, the force has a ratial (as for $r < \sigma_{LJ}$ distances) plus angular part. We will derive for the capsomer i because the translational force for capsomer j is $F^{ij}(r) = -F^{ji}(r)$. The translational forces for capsomer i in the region of distances are:

$$\begin{aligned}\overrightarrow{F}^r_{ij} &= -\frac{\partial(V_{LJ}(r)V_{ang}(\mathbf{r}_{ij}, \overrightarrow{\Omega}_i, \overrightarrow{\Omega}_j)V_{tor}(\Omega_i, \Omega_j))}{\partial r} = \\ &= -\frac{\partial V_{LJ}}{\partial r}V_{ang}V_{tor}\frac{\overrightarrow{r}_{ij}}{r} - \bigtriangledown_{\overrightarrow{r}_{ij}}(V_{ang})V_{LJ}V_{tor} - \bigtriangledown_{\overrightarrow{r}_{ij}}(V_{tor})V_{LJ}V_{ang}\end{aligned} \tag{5.7}$$

The first term in the previous equation is just $\overrightarrow{F}^r_{ij} = F_r \frac{\overrightarrow{r}_{ij}}{r^2}$. The gradients in the second and third terms are given by:

$$\begin{aligned}\frac{\partial V_{ang}}{\partial x_i} = -V_{ang}&\left[\frac{\theta_{ij} - \nu}{\alpha^2 \sin(\theta_{ij})}\left(\frac{dx\cos(\theta_{ij})}{r^2} - \frac{\sin(\theta_i)\cos(\phi_i)}{r}\right) + \right. \\ &\left. + \frac{\theta_{ji} - \nu}{\alpha^2 \sin(\theta_{ji})}\left(\frac{-dx\cos(\theta_{ji})}{r^2} + \frac{\sin(\theta_j)\cos(\phi_j)}{r}\right)\right]\end{aligned}$$

$$\begin{aligned}\frac{\partial V_{ang}}{\partial y_i} &= -V_{ang}\left[\frac{\theta_{ij}-\nu}{\alpha^2\sin(\theta_{ij})}\left(\frac{dy\cos(\theta_{ij})}{r^2}-\frac{\sin(\theta_i)\sin(\phi_i)}{r}\right)+\right.\\ &\left.+\frac{\theta_{ji}-\nu}{\alpha^2\sin(\theta_{ji})}\left(\frac{-dy\cos(\theta_{ji})}{r^2}+\frac{\sin(\theta_j)\sin(\phi_j)}{r}\right)\right]\end{aligned}$$

$$\begin{aligned}\frac{\partial V_{ang}}{\partial z_i} &= \\ &= -V_{ang}\left[\frac{\theta_{ij}-\nu}{\alpha^2\sin(\theta_{ij})}\left(\frac{dz\cos(\theta_{ij})}{r^2}-\frac{\cos(\theta_i)}{r}\right)+\right.\\ &\left.+\frac{\theta_{ji}-\nu}{\alpha^2\sin(\theta_{ji})}\left(\frac{-dz\cos(\theta_{ji})}{r^2}+\frac{\cos(\theta_j)}{r}\right)\right]\end{aligned} \tag{5.8}$$

$$\begin{aligned}\frac{\partial V_{tor}}{\partial x_i} &= -V_{tor}\frac{k_t}{2}\left[\left(\frac{dx\cos(\theta_{ij})}{r^2}-\frac{\sin(\theta_i)\cos(\phi_i)}{r}\right)\cos(\theta_{ji})+\right.\\ &\left.+\cos(\theta_{ij})\left(\frac{-dx\cos(\theta_{ji})}{r^2}+\frac{\sin(\theta_j)\cos(\phi_j)}{r}\right)\right]\\ \frac{\partial V_{tor}}{\partial y_i} &= -V_{tor}\frac{k_t}{2}\left[\left(\frac{dy\cos(\theta_{ij})}{r^2}-\frac{\sin(\theta_i)\sin(\phi_i)}{r}\right)\cos(\theta_{ji})+\right.\\ &\left.+\cos(\theta_{ij})\left(\frac{-dy\cos(\theta_{ji})}{r^2}+\frac{\sin(\theta_j)\sin(\phi_j)}{r}\right)\right]\\ \frac{\partial V_{tor}}{\partial z_i} &= -V_{tor}\frac{k_t}{2}\left[\left(\frac{dz\cos(\theta_{ij})}{r^2}-\frac{\cos(\theta_i)}{r}\right)\cos(\theta_{ji})+\right.\\ &\left.+\cos(\theta_{ij})\left(\frac{-dz\cos(\theta_{ji})}{r^2}+\frac{\cos(\theta_j)}{r}\right)\right]\end{aligned} \tag{5.9}$$

where $dx = x_j - x_i$, $dy = y_j - y_i$ and $dz = z_j - z_i$.

Now, we will calculate the angular forces for the capsomer i, F_i^θ and F_i^ϕ and for capsomer j, F_j^θ and F_j^ϕ. These forces are calculated as minus the partial derivative of θ and ϕ as:

$$\begin{aligned}F_{i,j}^\theta &= -V_{LJ}\left(V_{tor}\frac{\partial V_{ang}}{\partial\theta_{i,j}}+V_{ang}\frac{\partial V_{tor}}{\partial\theta_{i,j}}\right)\\ F_{i,j}^\phi &= -V_{LJ}\left(V_{tor}\frac{\partial V_{ang}}{\partial\phi_{i,j}}+V_{ang}\frac{\partial V_{tor}}{\partial\phi_{i,j}}\right)\end{aligned} \tag{5.10}$$

In the case of angular forces we must calculate them for capsomer i and j separately because $F_{ij}(\theta) \neq F_{ji}(\theta)$ and $F_{ij}(\phi) \neq F_{ji}(\phi)$. The partial derivates of angular parts $\frac{\partial V_{ang}}{\partial \theta_i}$ and $\frac{\partial V_{ang}}{\partial \theta_j}$ are given by:

$$\begin{aligned}\frac{\partial V_{ang}}{\partial \theta_i} &= V_{ang}\left(\frac{\theta_{ij}-\nu}{\alpha^2\sin(\theta_{ij})}\right)\frac{1}{r}\Big[dx\cos(\theta_i)\cos(\phi_i)+dy\cos(\theta_i)\sin(\phi_i)-dz\sin(\theta_i))\Big]\\ \frac{\partial V_{ang}}{\partial \theta_j} &= V_{ang}\left(\frac{\theta_{ji}-\nu}{\alpha^2\sin(\theta_{ji})}\right)\frac{1}{r}\Big[-dx\cos(\theta_j)\cos(\phi_j)+dy\cos(\theta_j)\sin(\phi_j)-dz\sin(\theta_j))\Big]\end{aligned}\tag{5.11}$$

The partials of torsion part of the potential with respect to θ are given by:

$$\begin{aligned}\frac{\partial V_{tor}}{\partial \theta_i} &= \frac{k_t}{2}\Bigg[\frac{1}{r}(dx\cos(\theta_i)\cos(\phi_i)+dy\cos(\theta_i)\sin(\phi_i)-dz\sin(\theta_i))\cos(\theta_{ji})+\\ &\quad +\cos(\theta_i)\sin(\theta_j)(\cos(\phi_i)\cos(\phi_j)+\sin(\phi_i)\sin(\phi_j))-\sin(\theta_i)\cos(\theta_j)\Bigg] \qquad (5.12)\\ \frac{\partial V_{tor}}{\partial \theta_j} &= \frac{k_t}{2}\Bigg[\frac{1}{r}(-dx\cos(\theta_j)\cos(\phi_j)+dy\cos(\theta_j)\sin(\phi_j)-dz\sin(\theta_j))\cos(\theta_{ij})+\\ &\quad +\sin(\theta_i)\sin(\theta_j)\cos(\phi_i)\cos(\phi_j)+\cos(\theta_i)\sin(\phi_i)\cos(\theta_j)\sin(\phi_j)-\sin(\theta_j)\cos(\theta_i)\Bigg]\end{aligned}$$

Similarly, the potential derivates with respect to ϕ_i and ϕ_j are:

$$\begin{aligned}\frac{\partial V_{ang}}{\partial \phi_i} &= V_{ang}\left(\frac{\theta_{ij}-\nu}{\alpha^2\sin(\theta_{ij})}\right)\frac{1}{r}(-dx\sin(\theta_i)\sin(\phi_i)+dy\sin(\theta_i)\cos(\phi_i))\\ \frac{\partial V_{ang}}{\partial \phi_j} &= V_{ang}\left(\frac{\theta_{ji}-\nu}{\alpha^2\sin(\theta_{ji})}\right)\frac{1}{r}(dx\sin(\theta_j)\sin(\phi_j)-dy\sin(\theta_j)\cos(\phi_j))\end{aligned}\tag{5.13}$$

and

$$\begin{aligned}\frac{\partial V_{tor}}{\partial \phi_i} &= V_{tor}\frac{k_t}{2r}\Big[-dx\sin(\theta_i)\sin(\phi_i)+dy\sin(\theta_i)\cos(\phi_i))\cos(\theta_{ji})-\\ &\quad -\sin(\theta_i)\sin(\phi_j)(\sin(\phi_i)\cos(\phi_j)+\cos(\phi_i)\sin(\phi_j))\Big]\\ \frac{\partial V_{tor}}{\partial \phi_j} &= V_{tor}\frac{k_t}{2r}\Big[dx\sin(\theta_j)\sin(\phi_j)-dy\sin(\theta_j)\cos(\phi_j))\cos(\theta_{ij})-\\ &\quad -\sin(\theta_i)\sin(\phi_j)(\sin(\phi_i)\sin(\phi_j)+\cos(\phi_i)\cos(\phi_j))\Big]\end{aligned}\tag{5.14}$$

Replacing the previous expressions into eq. 5.10, we finally get for the angular forces on capsomer i:

$$
\begin{aligned}
F_i^{\theta} \;=\; & -V_{LJ}V_{ang}V_{tor}\Bigg(\frac{k_t}{2}(1-\cos\phi_i) \\
& \Bigg[\cos\theta_{ji}\frac{\theta_{ij}-\nu}{\alpha^2}(dx\cos\theta_i\cos\phi_i + dy\cos\theta_i\sin\phi_i - dz\sin\theta_i) + \\
& (\cos\theta_i\cos\phi_i\sin\theta_j\cos\phi_j + \cos\theta_i\sin\phi_i\sin\theta_j\sin\phi_j - \sin\theta_i\cos\theta_j)\Bigg] \\
& + \Bigg(\frac{\theta_{ij}-\nu}{\sin\theta_{ij}\alpha^2}(dx\cos\theta_i\cos\phi_i + dy\cos\theta_i\sin\phi_i - dz\sin\theta_i)\Bigg)\Bigg)
\end{aligned}
\tag{5.16}
$$

$$
\begin{aligned}
F_i^{\phi} \;=\; & -V_{LJ}V_{ang}V_{tor}\Bigg(\frac{k_t}{2}(1-\cos\phi_i) \\
& \Bigg[\cos\theta_{ji}\frac{\theta_{ij}-\nu}{\alpha^2}(dx\sin\theta_i\sin\phi_i + dy\sin\theta_i\cos\phi_i) + \\
& (-\sin\theta_i\sin\phi_i\sin\theta_j\cos\phi_j + \sin\theta_i\cos\phi_i\sin\theta_j\sin\phi_j)\Bigg] \\
& + \Bigg(\frac{\theta_{ij}-\nu}{\sin\theta_{ij}\alpha^2}(-dx\sin\theta_i\sin\phi_i + dy\sin\theta_i\cos\phi_i)\Bigg)\Bigg)
\end{aligned}
\tag{5.18}
$$

And the corresponding forces on capsomer j are:

$$
\begin{aligned}
F_j^{\theta} \;=\; & -V_{LJ}V_{ang}V_{tor}\Bigg(\frac{k_t}{2}(1-\cos\phi_j) \\
& \Bigg[\cos\theta_{ij}\frac{\theta_{ji}-\nu}{\alpha^2}(-dx\cos\theta_j\cos\phi_j - dy\cos\theta_j\sin\phi_j + dz\sin\theta_j) + \\
& (\sin\theta_i\cos\phi_i\cos\theta_j\cos\phi_j + \theta_i\sin\phi_i\cos\theta_j\sin\phi_j - \cos\theta_i\sin\theta_j)\Bigg] \\
& + \Bigg(\frac{\theta_{ji}-\nu}{\sin\theta_{ji}\alpha^2}(-dx\cos\theta_j\cos\phi_j - dy\cos\theta_j\sin\phi_j + dz\sin\theta_j)\Bigg)\Bigg)
\end{aligned}
\tag{5.20}
$$

$$
\begin{aligned}
F_j^{\phi} \;=\; & -V_{LJ}V_{ang}V_{tor}\left(\frac{k_t}{2}(1-\cos\phi_j)\right. \\
& \left[\cos\theta_{ij}\frac{\theta_{ji}-\nu}{\alpha^2}(-dx\sin\theta_j\sin\phi_j - dy\sin\theta_j\cos\phi_j) + \right. \\
& \left.(-\sin\theta_i\cos\phi_i\sin\theta_j\sin\phi_j + \sin\theta_i\sin\phi_i\sin\theta_j\cos\phi_j)\right] \\
& \left. + \left(\frac{\theta_{ji}-\nu}{\sin\theta_{ji}\alpha^2}(dx\sin\theta_j\sin\phi_j + dy\sin\theta_j\cos\phi_j)\right)\right)
\end{aligned}
\tag{5.22}
$$

References

[1] V. Voelz, GR. Bowman, K. Beauchamp and VS. Pande, Molecular Simulation of ab Initio Protein Folding for a Millisecond Folder NTL9(1-39), *JACS*, **132**, 5, 1526, 2010.

[2] MW. Mahoney and WL. Jorgensen, A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions, *J. Chem. Phys.*, **112**, 20, 8910-8922, 2000.

[3] V. Bianco. Thesis, Statistical Physics of Water in Hydrophobic Nano-Confinement and at Proteins Interfaces, 2013.

[4] R. Goetz and R. Lipowsky, Computer simulations of bilayer membranes: Self-assembly and interfacial tension, *J. Chem. Phys.*, **108**, 17, 7397-7409, 1998.

[5] N. Metropolis and S. Ulam, The Monte Carlo Method, *J. Amer. Stat. Assoc.*, **44**, 247, 335-341, 1949.

[6] D. Frenkel and B. Smit, Understanding molecular simulations. *Academic Press*, 2002.

[7] P. Del Moral, Mean Field Simulation for Monte Carlo Integration. *Chapman Hall/CRC*, 2013.

[8] K. Sabelfeld, Monte Carlo Methods in Boundary Value Problems, *Springer-Verlag*, 1991.

[9] MN. Barber, RB. Pearson, D. Toussaint and JL. Richardson, Finite-size scaling in the three-dimensional Ising model, *Phys. Rev. B*, **32**, 1720, 1985.

[10] W. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, 97-109, 1970.

[11] AM. Ferrenberg and RH. Swendsen, Optimized Monte-Carlo Data-Analysis, *Phys. Rev. Lett.*, **63**, 12, 1195-1198, 1989.

[12] BS. Gonzalez, J. Hernandez-Rojas and J. Breton *et al.* Global Potential Energy Minima of (H2O)(n) Clusters on Graphite: A Comparative Study of the TIPNP (N=3, 4, 5) Family, *J. Phys. Chem. C*, **112**, 42, 16497-16504, 2008.

[13] W. Coffey, PY. Kalmykov and JT. Waldron, The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry, and Electrical Engineering, *World Scientific*, 2004.

[14] AL Bozic, A. Siber and R. Podgornik, Statistical analysis of sizes and shapes of virus capsids and their resulting elastic properties, *J. Biol. Phys.*, **39**, 2, 215-228, 2013.

[15] MP. Allen and DJ. Tildesley, Computer Simulation of Liquids, Oxford, 1987.

[16] A. Hrennikoff, Theory of Inelastic Bending with Reference to Limit Design, *Transactions of the American Society of Civil Engineers*, **113**, 213-247, 1948.

[17] G. Strang and G. Fix, *An Analysis of The Finite Element Method*, Prentice Hall, 1973.

[18] J. JIN, The finite element method in electromagnetics, (2nd Ed.), Lavoisier, Paris, 2002.

[19] GF. Pinder, and WG. Gray, Finite element simulation in surface and subsurface hydrology, Academic Press, New York, 1977.

[20] L. D. Landau and I. M. Lifshitz, Theory of elasticity. Pergamon, New York, 1975.

[21] J. Lidmar, L. Mirny, and D. Nelson, Virus shapes and buckling transitions in spherical shells, *Phys. Rev. E*, **68**, 051910, 2003.

[22] T. Nguyen, R. Bruinsma, and W. Gelbart, Continuum theory of retroviral capsids, *Phys. Rev. Lett.*, **96**, 078102, 2006.

[23] T. Guerin and R. Bruinsma, Theory of conformational transitions of viral shells, *Phys. Rev. E*, **76**, 061911, 2007.

[24] B. Stephanidis, S. Adichtchev, P. Gouet, a. McPherson, and a. Mermet, Elastic properties of viruses, *Biophys. J.*, **93**, 1354-9, 2007.

[25] F. Tama and C. L. Brooks, Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis, *J. Mol. Biol.*, **345**, 299-314, 2005.

[26] T. Nguyen, R. Bruinsma, and W. Gelbart, Elasticity theory and shape transitions of viral shells, *Phys. Rev. E*, **72**, 051923, 2005.

[27] A. Y. Morozov and R. F. Bruinsma, Assembly of viral capsids, buckling, and the Asaro-Grinfeld-Tiller instability, *Phys. Rev. E*, **81**, 041925, 2010.

[28] Z. Yang, I. Bahar, and M. Widom, Vibrational dynamics of icosahedrally symmetric biomolecular assemblies compared with predictions based on continuum elasticity, *Biophys. J.*, **96**, 4438-48, 2009.

[29] Y. Tao, NH. Olson, W. Xu, DL. Anderson, MG. Rossmann and TS. Baker, Assembly of a Tailed Bacterial Virus and Its Genome Release Studied in Three Dimensions, *Cell*, **95**, 3, 431-437, 1998.

[30] A. Luque, R. Zandi, and D. Reguera, Optimal architectures of elongated viruses, *Proc. Nat. Acad. Sci. USA* **107**, 53235328, 2010.

Part II

Self-assembly of viruses

CHAPTER 6

Physical ingredients controlling the polymorphism and structural selection of empty viral capsids

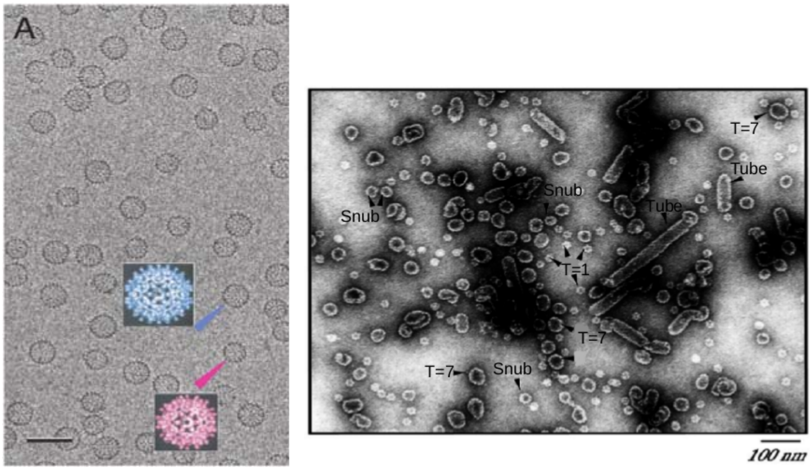
6.1 Introduction

Viruses are generally characterized by the architecture of their capsids. Typically, each native virus self-assembles into a unique structure, with some exceptions like Hepatitis B virus [1] which makes T=3 and T=4 capsids. However, the proteins of many viruses have the capability of self-assemble *in vitro* by changing the assembly conditions even in the absence of their genetic material, forming hollow capsids with different T-numbers and tubes [2–5], see Fig. 6.1. They can even self-assemble and encapsulate other types of (non-genetic) cargoes, thus opening the door to a wealth of nanotechnological applications of viruses including *in vivo* imaging, controlled reactions or drug delivery. A proper understanding and control of *in vitro* assembly is essential to get shells with well-defined size and structure, as required for the different aforementioned applications [6–9].

In recent decades, the self-assembly process of viruses has been investigated from different viewpoints, including *in vitro* experiments [2, 2–5, 12–14], theoretical modeling [15–18] and simulations. As described in Chapter 3 different simulation techniques and models have been used to study capsid structure and assembly. In most of these models the interactions have multiple sites whose location, strength, and directionality have to be tuned in a delicate way to reproduce a specific target structure. For this reason, it is no clear what are the essential mechanisms and general ingredients of the interaction which control the selection of a particular size or structure.

That is precisely the reason why we decided to develop the simplest model capable of describing the relevant mechanisms involved in viral assembly. In this spirit, the goal of this chapter is to use the coarse-grained model developed in Chapter 3 to understand the physical mechanisms controlling viral size and structure selection. We have analyzed the tolerance of the capsid to changes in the physical parameters, related to ambient conditions, to characterize viral stability and possible mechanisms to induce misassembly or failure. We will see that the effective bending stiffness and spontaneous curvature are the most relevant parameters determining the structure and stability of the capids. In addition, elongated, faceted, and decapsidated structures arise in a natural way with this model which sheds important light on our understanding of viral assembly and could guide future assembly experiments.

The chapter is distributed as follows. In section 6.2 we briefly describe some details about our coarse-grained model and the simulation. Section 6.3 is devoted to present the results of stability of $T = 1, 3, 4, 7$ and snub cube structures made exclusively by one type of capsomers. We will also describe their tolerance to changes in bending and spontaneous curvature and the resulting phase diagram representing the most stable shape and the most common aberrant. We will then repeat our analysis for viral shells made by two types of capsomers (hexamers and



(a) (b)

Figure 6.1: Examples of viral polymorphism *in vitro*. Electron microscopy observations of (a) Hepatitis B self-assembly into $T = 3$ and $T = 4$ structures [10] and (b) Simian Virus 40 self-assembly into $T = 1$, snub cubes, $T = 7$ and tubes of different sizes. Figures taken from Refs. [10] and [11].

pentamers), discussing the main differences. Finally, in section 6.5, we summarize the main conclusions and discuss their potential biological interpretation.

6.2 Model and Monte Carlo simulations

The coarse-grained model described in Chapter 3 has been used in the present chapter to analyze and understand the physical mechanisms dictating the size, the structure selection, and the stability of viral capsids. Both the models with one capsomer type (corresponding to all-pentamer viruses), and two types of capsomers (i.e. hexamers and pentamers) have been used and the corresponding results will be described in two separate sections. Monte Carlo simulations have been used in order to study the stability of preformed capsids and the range of parameters that would in principle guarantee a successful assembly. We have studied the smaller T-number structures, namely $T = 1, 3, 4$ and 7, plus the snub cube for all-pentamers viruses, which is a stable $N = 24$ structure that appears in *in vitro* assembly of polyoma [19] and SV40 virus [11].

Our code is based on a standard Metropolis algorithm described in Chapter 5. The following parameters have been fixed in all simulations (in reduced units): density $\rho = 0.1$, temperature $T = 0.05$, and torsion constant $k_{tor} = 1.5$. Each simulation consists of a total of 10^7 Monte Carlo steps sampled every 10^5 steps. From the Monte Carlo simulations we obtained the average energy per capsomer, the position of every capsomer and the integrity and order of the shell for each structure and for each value of α and ν. The integrity of the shell was analyzed automatically by monitoring the size of the shell using the following cluster criteria. Two capsomers are considered to be part of the same cluster if their distance is smaller or equal to 1.2 σ, and their relative angles are in the range $[\nu - 45\%\nu, \nu + 45\%\nu]$. In this way, we can check that capsomers are bound and have the proper orientation. Additionally, we have verified the integrity of the structure analyzing different snapshots of the simulation.The breaking of the structure can be also detected by an abrupt change in its energy.

In the case of capsids with one type of capsomer, the starting point of our MC simulations were the positions and angles of the most stable T-number structure harvested from the Brownian Dynamic assembly simulations described in Chapter 7. Using these structures with the optimal parameters as initial configuration, we have repeated the simulations for a set of different values of ν and α until the structure looses its integrity. We have used a different protocol to get our initial $T = 4$ structure, because it is not an optimal structure for one type of capsomers [21, 22] and was not obtained in our BD self-assembly simulations. In this case, we have used as initial guess the theoretical $T = 4$ positions and angles from the coordinates of the icosahedral spherical code corresponding to the solution of the maximum volume problem for $N = 42$ in Ref. [23]. We equilibrated this initial configuration using a short Monte Carlo simulation using $\alpha = 0.1$ and the predicted value $\nu_{opt} = 1.88496$ (see Table 3.1, in Chapter 3), and verified that this configuration was indeed locally stable. From it, we did a sequence of MC simulations changing ν and α.

The Monte Carlo simulations were used to analyze the energy and order of structures as a function of the two parameters ν and α. Changes in ν and α are intended to resemble changes in the environmental conditions (pH or salt concentration). As we will see, ν is the main parameter in the selection of a particular structure for a viral capsid. In fact, as described in Chapter 3, we can estimate the theoretical values of the preferred angle that are expected to be optimal for each T-number structure, that will be denoted as ν_{opt}. The resulting ν_{opt} for each T are listed in Table 3.1, and simulations have been made in a range of ν that encompasses structures from $T = 1$ to $T = 7$.

6.3 Stability of all-pentamer viruses

In this section, we will focus on the assembly and stability of viruses formed exclusively by one type of capsomer such as papillomavirus [24] and polyoma [19]. We will first analyze separately what is the influence of the bending stiffness and of the preferred angle ν on the energy and integrity of the capsids. We will then focus on the conditions favoring elogated, faceted or decapsidated shells; and end up by showing the phase diagram of each structure and the complete stability

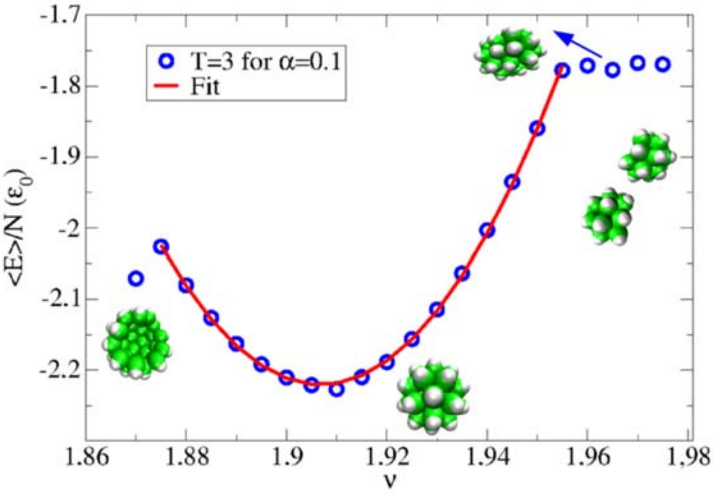


Figure 6.2: Average energy per capsomer for $k_{tor} = 1.5$, $T = 0.05$, and $\alpha = 0.1$, as a function of the preferred angle ν for the $T = 3$ structure with $N = 32$ capsomers. The inset figures show snapshots of the optimal capsid, an elongated shell, and the typical aberrant structures obtained when the capsid loses its stability at too low and too high values of ν. The red line corresponds to a fit: $< E > /N = -2.23 + \frac{1}{2}386.8(\nu - 1.907)^2$.

phase diagram in the range of parameters studied.

6.3.1 Stability of different structures

In the following, we summarize the results obtained for the stability of $T = 1, 3, 4, 7$ and snub cube structures made by one type of capsomers, using Monte Carlo simulations. We have analyzed the tolerance of each structure to changes in ν and α and the influence of the parameters in different morphology changes, and in the appearance of aberrant structures.

Influence of the preferred angle ν

First, we have analyzed the stability of all structures by fixing $\alpha = 0.10$ and changing the preferred angle ν, related to the spontaneous curvature, for each structure. We have chosen $\alpha = 0.1$, because at this value of α all structures self-assembled successfully in our BD simulations (see Chapter 7). For all the structures we have observed the same qualitative behavior upon changing ν, which is illustrated in Fig. 6.2 for the case of $T = 3$. The average energy per capsomer as a function of ν has a clear parabolic behavior for all structures. By fitting the energy per capsomer to a parabola $\frac{E}{N} = \frac{E_0}{N} + \frac{K}{2}(\nu - \nu_0)^2$ around the optimal value

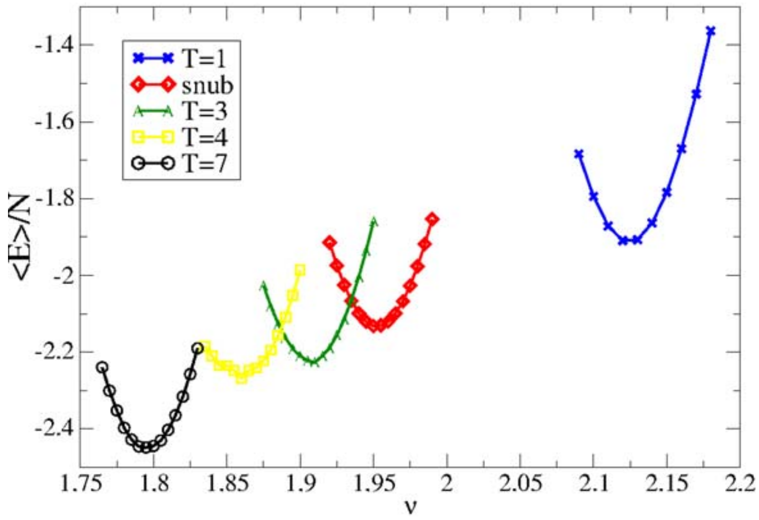


Figure 6.3: Average energy per capsomer for $k_{tor} = 1.5$, $T = 0.05$, and $\alpha = 0.1$, as a function of the preferred angle ν for all structures.

Structure	ν range	α range	E_0/N (ϵ_0)	ν_{opt}	ν_0	K
T=1	2.09-2.22	0.03-1.81	-1.91	2.12437	2.125	362
Snub	1.92-2.01	0.04-1.15	-2.13	1.954	1.953	416
T=3	1.88-1.98	0.06-1.40	-2.23	1.93566	1.907	387
T=4	1.84-1.92	0.06-1.25	-2.26	1.88496	1.858	316
T=7	1.77-1.85	0.03-1.75	-2.45	1.79667	1.796	441

Table 6.1: Stability range and parameters for the different structures. The columns indicate: the range of ν at $\alpha = 0.10$, and the range of α at $\nu = \nu_{opt}$, at which each structure maintains its integrity; the predicted value of the angle ν_{opt} for each structure; the optimal preferred angle ν_0, the optimal average energy per capsomer E_0/N, and the effective stiffness K, obtained from the parabolic fit of the energy per capsomer vs ν for each structure.

of the preferred angle, we have obtained a minimum energy per capsomer E_0/N, the optimal angle ν_0, and the effective stiffness K for every structure. The results are compiled in Table 6.1. Remarkably, the values of the optimal spontaneous curvature ν_0 obtained from the fit (see Table 6.1) agree, within the error, with the expected values listed in Table 3.1, except for the $T = 3$ structure, whose ν_0 obtained from the simulations is smaller than the predicted ν_{opt}.In addition, the values of the effective stiffness K are similar to the naive estimate $K \approx \frac{z\epsilon_0}{2\alpha^2} = 300$ for the average bending stiffness per capsomer of a hexagonal network proposed in the model of Chapter 4.

The parabolic behavior ends abruptly at both extremes, indicating that there is a finite range of spontaneous curvatures for which each structure keeps stable.

Outside this range, structures suffer a morphological change or loose their integrity by different mechanisms. When the preferred angle ν is too small, the bending contribution forces the structure to be less curved until eventually it opens up (see Fig. 6.2). In this case, breaking occurs when the energy cost of being an open structure (associated to the line tension due to the missing interactions at the rim of the open capsid) is less than the cost of feeling angularly frustrated. On the other hand, for large spontaneous curvatures, the structures loose their stability by different mechanisms. In Fig. 6.2 we first see a plateau at big ν in the average energy per capsomer for the $T = 3$ structure, corresponding to the appearence of an elongated structure, as described in more detail below. If we keep increasing the preferred angle ν, the bending contribution will tend to make the shell smaller until the energetic cost needed to keep the integrity of the $T = 3$ structure becomes too large, and the shell eventually breaks up typically into two smaller incomplete structures (see Fig. 6.2).

Finally, we have compared the behavior of the energy per capsomer as a function of the preferred angle ν for a fixed $\alpha = 0.10$ and for all structures, see Fig. 6.3. All structures have a parabolic behavior of the energy. But as the structure becomes bigger, its tends to be more stable, i.e. large structures have lower optimal average energy per capsomer E_0/N. Note that this optimal values of the energy depend on the chosen value of α. As we will see in the next section, E_0/N gets smaller with α, since the bending penalty becomes smaller, plus the interaction becomes slightly longer ranged. Another important observation is that large structures, like $T = 7$, tolerate less changes in ν than small structures, as $T = 1$ (see Table 6.1). That is because in big structures a small change in ν means a big change in the radii, as indicated by eq. 3.10. But the most important feature of Fig. 6.3 is the overlap in the stability range of $T = 3$, $T = 4$ and the snub cube. That is a clear indication that, for intermediate values of the spontaneous curvature, these three structures compete.

Influence of the bending stiffness parameter α

In this section we have repeated our analysis of the stability of all structures now by fixing ν to the predicted optimal angle ν_{opt}, listed in Table 6.1, and changing α, in different ranges up to the breaking of the structure, which is detected both in the cluster statistics as well as by an abrupt change in the average energy per capsomer (see Fig. 6.4). For all the structures we have observed the same qualitative behavior upon changing α, which is illustrated in Fig. 6.4 for the case of $T = 3$. The energy per capsomer gets monotonically smaller with α, since the bending penalty becomes smaller, plus the interaction becomes slightly longer ranged and more isotropic. The main difference between structures is the range of α in which they are stable, that is listed in Table 6.1. Outside this stability range, the shell loses its integrity by two different mechanisms. For very small α the

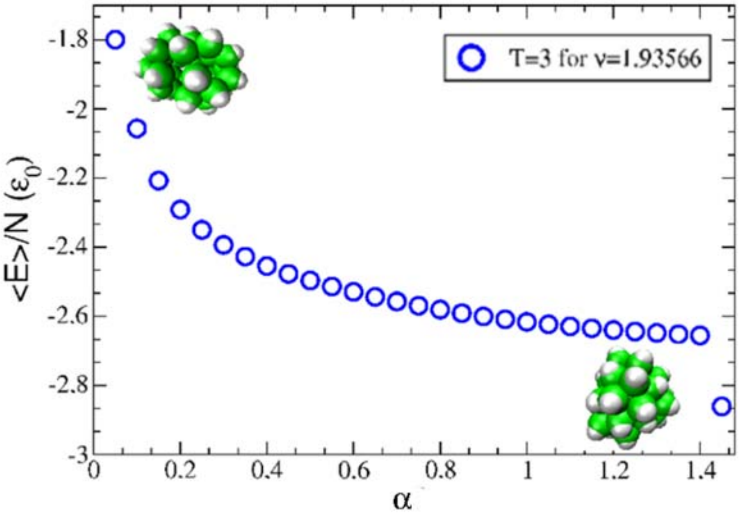


Figure 6.4: Average energy per capsomer versus α for the $T = 3$ structure with $\nu = 1.93566$, $N = 32$, $k_{tor} = 1.5$, and T=0.05.

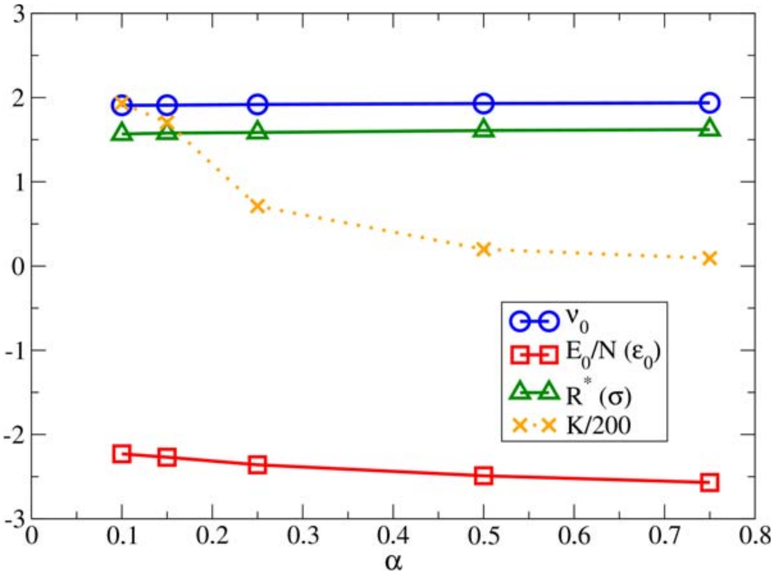


Figure 6.5: Optimal angle, ν_0 (blue line), optimal energy per capsomer, E_0/N (red line), optimal radius of the shell R^* (green line) and its effective stiffness K (orange line) as a function of α for a $T = 3$ with $\nu = \nu_{opt}$.

angular interaction tolerance is very small so that any small misalignment driven by thermal fluctuations leads to a progressive loss of capsomers. For too large α, the angular tolerance of the Lennard-Jones potential is large, becoming progressively more isotropic, and the shell collapses into a compact cluster or concentric layer structures, as can be seen for the $T = 3$ example in Fig6.4.

From the stability ranges listed in Table 6.1, it is clear that $T = 1$ and $T = 7$

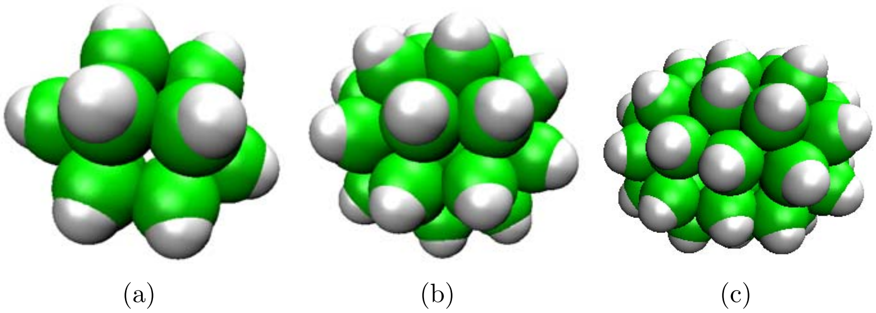

(a) (b) (c)

Figure 6.6: Snapshots of different elongated structures for $\alpha = 0.10$, $k_{tor} = 1.5$ and T=0.05 for (a) $T = 1$ for $\nu = 2.2$, (b) Snub cube for $\nu = 2.01$ and (c) $T = 3$ structure for $\nu = 1.97$.

are the structures that have the largest tolerance to changes in α, whereas the snub cube and $T = 4$ have the least tolerance to changes in α.

Finally, we have analyzed how the value of α modifies the optimal angle, ν_0, the optimal energy per capsomer, E_0/N, the optimal radius of the shell, and its effective stiffness K. The results for the case of $T = 3$ obtained in the range $\alpha \in [0.05, 1.00]$ are plotted in Fig. 6.5. In that range of α the optimal angle ν_0, and radius R^* are practically constant (ν only increases by 0.033 rad and the radius by 0.1σ), but the optimal energy per capsomer E_0/N decreases by $0.39\epsilon_0$. The effective stiffness K changes as $\sim \frac{1}{\alpha^2}$, as predicted for the bending rigidity according to eq. 3.22, obtained in Chapter 4.

6.3.2 Elongated structures

Bacilliform structures are a common shape adopted by some native viruses and also by some spherical viruses assembled in *in vitro* experiments. These elongated morphologies have been extensively reported and studied showing a great variety of sizes. Some examples for all pentamer viruses include SV40, see Fig. 6.1(b), or human papillomavirus, where elongated and spherical shapes coexist in *in vitro* experiments.

For all the structures that we studied, we found a small region of parameters where elongated structures emerge as optimal structures. In particular for $T = 1$ we saw a morphological change from the $T = 1$ spherical capsid to an elongated structure (see Fig. 6.6 (b)) in the range $\alpha = 0.1$ and $\nu \in [2.185, 2.22]$. For the snub cube, we have obtained elongated structures in the range $\alpha = 0.1$ and $\nu \in [2.005, 2.010]$, see Fig. 6.6(b). For $\alpha = 0.05$ there are elongated structures in the ranges $\nu \in [1.925, 1.94]$ and $\nu \in [1.885, 1.890]$ for $T = 3$, see Fig. 6.6(c), and $T = 4$ structures, respectively. Finally, in the region $\alpha = 0.03$ and $\nu \in [1.80, 1.81]$

T	R (σ)	ν	CR
1	0.951	2.185	2.19
Snub cube	1.602	2.005	2.43
T=3	1.957	1.925	2.71
T=4	1.99	1.89	2.50
T=7	2.93	1.80	2.66

Table 6.2: Values of the radius R and the preferred angle ν, where the transition to an elongated structure occurs for the different structures studied. The last column list the value of CR, that is very close to the prediction $CR \approx 3$ of Nguyen and co-workers [29].

is where $T = 7$ structure becomes elongated.

In all cases, the appearance of elongated structures occurs at small α and $\nu > \nu_{opt}$. Presumably, for $\nu > \nu_{opt}$ the capsid is forced to increase its curvature, and if the bending penalty is important, as occurs for small α, a compromise solution is to change the shape into an elongated structure, rather than a compressed spherical shell.

Nguyen and co-workers [29] estimated using a simple continuum elastic model that the transition from sphere to spherocylinder takes place, for small FvK numbers, when $CR \simeq 3$. To check that prediction, we have calculated the value of $CR = \frac{-4cos\nu}{\sigma}R$ for every structure at the value of α and ν at which the transition to an elongated structure takes place, that can be located by the kink in the optimal energy versus ν plot (see Fig. 6.2). The results are listed in Table 6.2, and are very close to the theoretical prediction, deviating slightly more from the predictions for the cases of $T = 4$ and the snub-cube.

6.3.3 Buckling

Many spherical viruses, before becoming infective, go through an intermediate stage of maturation, where they undergo a structural transition that flattens the faces of their capsids to form an icosahedron. This morphological change is known as *buckling transition* and it has been characterized in the framework of continuum elasticity [29, 30].

In continuum elasticity theory a viral capsid is considered as a thin elastic shell with two competing energies: The bending energy, which takes into account the deviations from the spontaneous curvature imposed by the proteins, and the stretching energy, which involves the in-plane deformations of the network of proteins. The ratio of these contributions defines the Föppl-von Kárman number,

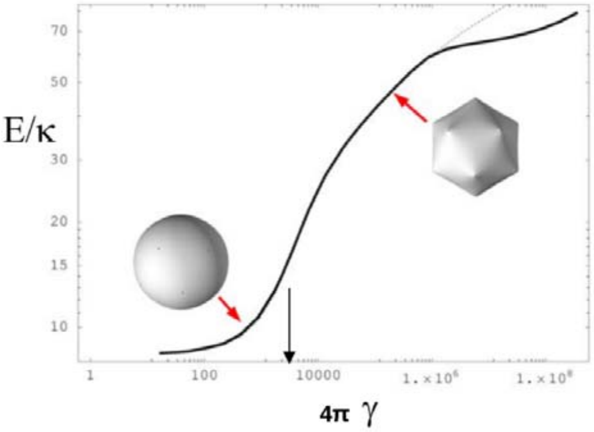


Figure 6.7: Elastic energy E of an icosahedral shell expressed in units of the bending constant κ versus Foppl-von Kármán Number. Near a critical value around $4\pi\gamma = 3000$ (i.e $\gamma \approx 239$) a buckling transition takes place, with the shell shape transforming from spherical to icosahedral. The dotted line shows the result of a fit to the theory [29].

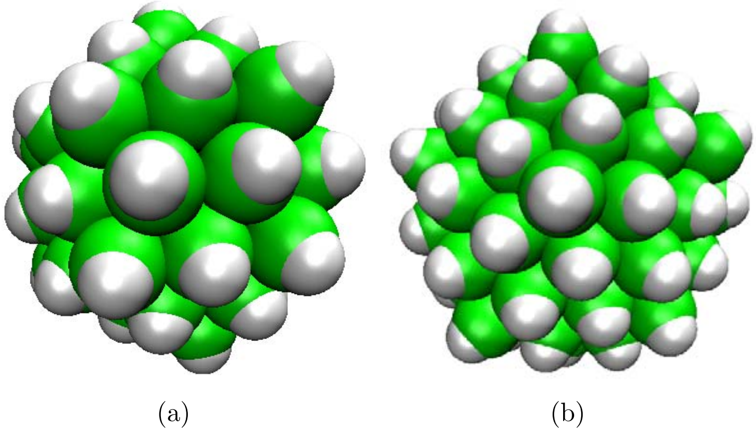

(a) (b)

Figure 6.8: Snapshots of different faceted structures obtained in our simulations corresponding to a (a) $T = 4$ for $\nu = 1.875$ and $\alpha = 1.2$ and (b) $T = 7$ for $\nu = 1.797$ and $\alpha = 1.6$.

$\gamma = YR^2/\kappa$, a nondimensional parameter that determines the shape of the virus, where Y is the Young modulus and κ the bending rigidity. For $\gamma_B \approx 240$ it is predicted the occurrence of a buckling transition from a spherical shape, at $\gamma < 240$ to an increasingly faceted polyhedral capsid for $\gamma > 240$, see Fig. 6.7.

One of the features of our coarse-grained model is that it naturally accounts for this buckling transition. In the context of our model, using the values of Y and κ derived in Chapter 3, the FvK number γ becomes:

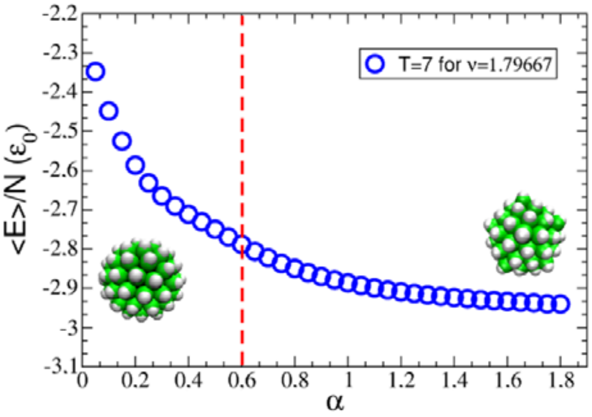


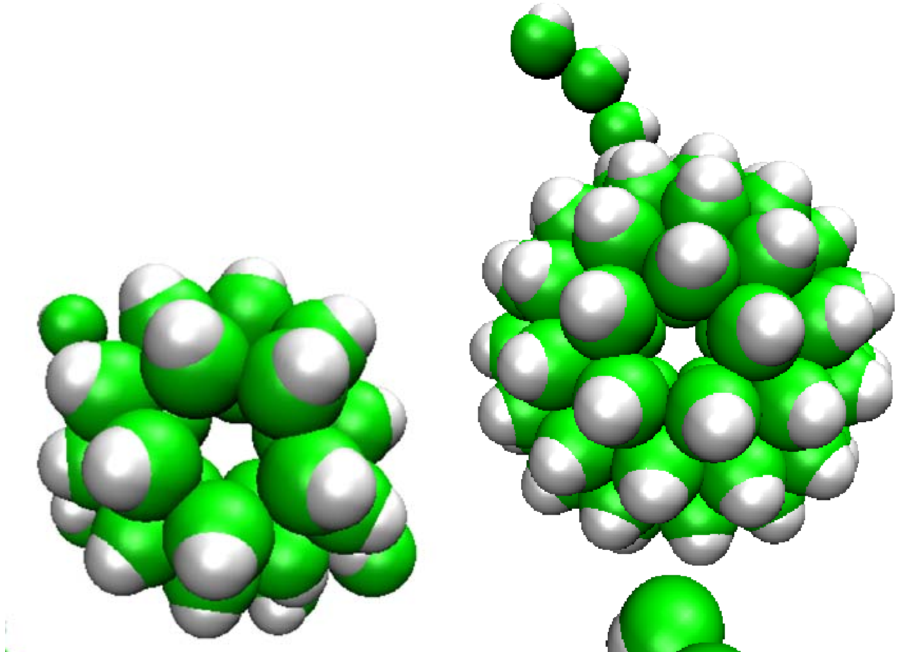
Figure 6.9: Average energy per capsomer for a $T = 7$ structure versus α at $\nu_{opt} = 1.79667$, $k_{tor} = 1.5$ and $T = 0.05$. The red dashed line indicates the transition from a spherical capsid to a faceted shell, as illustrated by the inset snapshots taken from the MC simulations.

$$\gamma = 128\alpha^2 \left(\frac{R}{\sigma}\right)^2 \tag{6.1}$$

For a given T-number the radius will be approximately constant, so the occurrence of a buckling transition at which we will start noticing the faceting of our shells will be directly related to the value of α. Since the FvK is proportional to R^2, buckling will be observed more easily for large capsids. Indeed, we have observed this buckling transition in our largest capsids, i.e. $T = 4$ and $T = 7$ having radii $R = 1.99\sigma$ and $R = 2.93\sigma$, respectively, for the corresponding ν_{opt}, see Fig. 6.8.

Fig. 6.9 represents the average energy per capsomer as a function of α for $T = 7$ at the optimal spontaneous angle ν_{opt}, see Fig. 6.8(b). In this case, there is a critical α value of around 0.6 (indicated by the red line in Fig. 6.9), where the equilibrium structure starts to become noticeably and progressively more faceted upon increasing α. Thus, this value of $\alpha = 0.6$ characterizes the onset of buckling for ν_{opt}, and corresponds to a value of $\gamma \approx 225$, which compares well with the predicted threshold γ_B. For $T = 4$ at ν_{opt}, the radius is $R(T = 4) = 1.99\sigma$, and a noticeable buckling is observed also starting at $\alpha_B = 0.6$, corresponding to $\gamma = 125$, see Fig. 6.8(a).

The critical value of α increases slightly with ν, becoming for instance around $\alpha = 0.65$ for bigger ν angles. This is due to the fact that larger values of ν slightly reduce the radius, thus requiring larger α's to compensate this effect. It is worth noticing that faceted structures have less energy than the spherical shape.



(a) (b)

Figure 6.10: Snapshots of decapsidated structures obtained (a) for $T = 3$ with $\nu = 1.95$ and $\alpha = 0.02$. (b) For $T = 7$ at $\nu = 1.81$ and $\alpha = 0.05$.

6.3.4 Decapsidation

There a some viruses that *in vitro* suffer a decapsidation process driven typically by changes in pH or temperature. Examples of $T = 3$ viruses that undergo decapsidation are Turnip Yellow Mosaic Virus (TYMV) [31], Belladonna Mottle virus (BelMV) [32] and Eggplant Mosaic Virus (EMV) [32]. These viruses loose their pentamers *in vitro* when the pH is ≤ 5.0.

We have also seen the decapsidation process in our simulation for $T = 3$ in a small region of parameters, corresponding to $\alpha = 0.02$ and $\nu \in [1.945, 1.950]$ and around $\alpha = 0.03$ and $\nu = 1.955$, see Fig. 6.10(a). There are also all-pentamers $T = 7$ viruses as Human Papillomaviruses Type 18 [33, 34] or polyoma virus [35] that loose their pentons in experiments. That corresponds to the behavior which we have obtained for our $T = 7$ structure at $\nu \in [1.79, 1.82]$ and $\alpha \in [0.03, 0.05]$, see Fig. 6.10(b). Thus decapsidation seems to be observed at very small values of α, i.e. large bending rigidities, where the small difference in the angle between pentons and hexons has a large energetic cost, so large that the shell is eventually more stable without them.

6.3.5 Stability phase diagram

In the previous section, simulations were performed either at a fixed α or at a fixed ν. In this section we present the full phase diagram of each structure in the range of values of α and ν where each of them have been found to be stable. This phase diagram has been obtained by repeating all the MC simulations of each structure, at steps of 0.05 in α and 0.01 in ν.

Figs. 6.11 to 6.15 compile the phase diagrams obtained for $T = 1$, the snub cube, $T = 3$, $T = 4$, and $T = 7$. We have also included in the diagram the typical aberrant structures. We can see in all cases 5 different morphologies, illustrated as an example, for the case of $T = 7$ Fig. 6.15: 1) the perfect spherical structure (indicated by the blue region), 2) one (or two) open structures, smaller than the normal one (the violet region) obtained at small α and large ν, 3) burst structures (indicated in red in Fig. 6.15) at small α and ν, 4) compact clusters or multi-layered structures (the orange region in Fig. 6.15) when the bending penalty is low, and 5) elongated capsids (in the narrow light green area in Fig. 6.15). For the largest $T = 4$ and $T = 7$ structures, there is also a region at $\alpha > 0.6$ where the equilibrium structure is a faceted icosahedral capsid (indicated by the green region. For $T = 3$ and $T = 7$ we have seen also a small region where decapsidation of the pentameric capsomers occurs (yellow area in Figs. 6.13 and 6.15, see previous subsection).

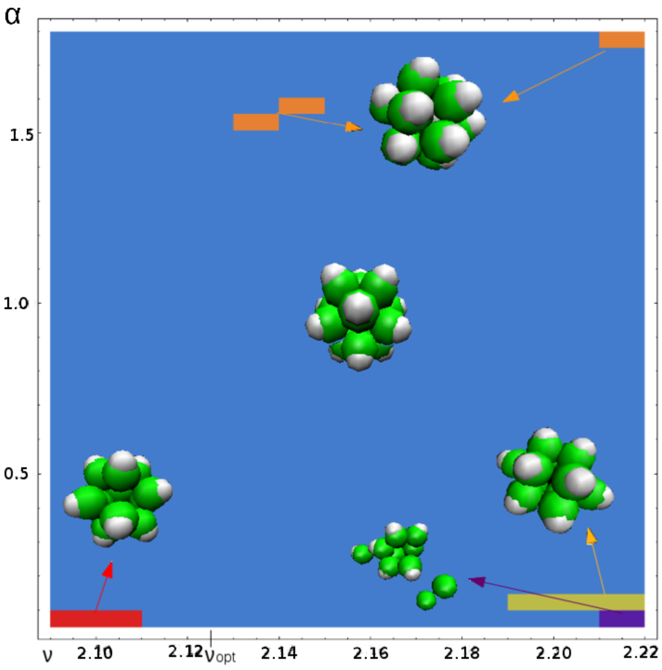


Figure 6.11: Phase diagram of the different structures obtained from a $T = 1$ as a function of α and ν, for $N = 12$, $k_{tor} = 1.5$ and $T = 0.05$.

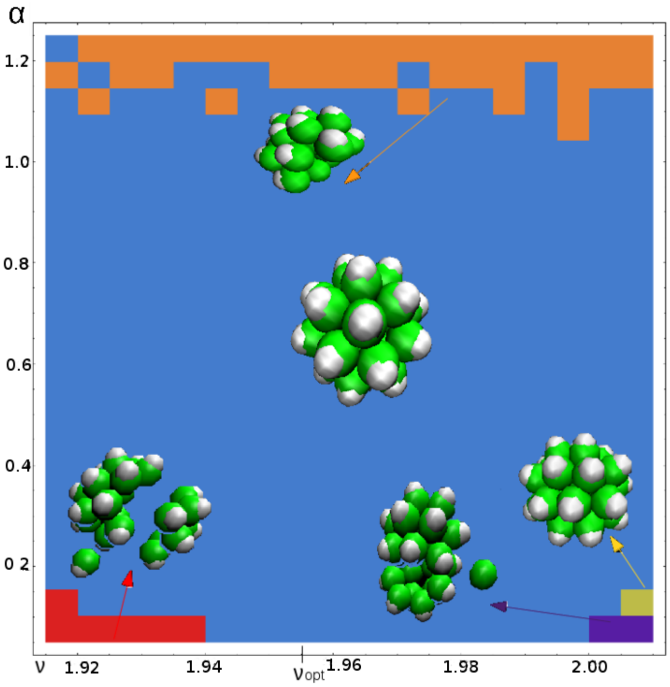


Figure 6.12: Phase diagram of the different structures obtained from a snub cube as a function of α and ν, for $N = 24$, $k_{tor} = 1.5$ and $T = 0.05$.

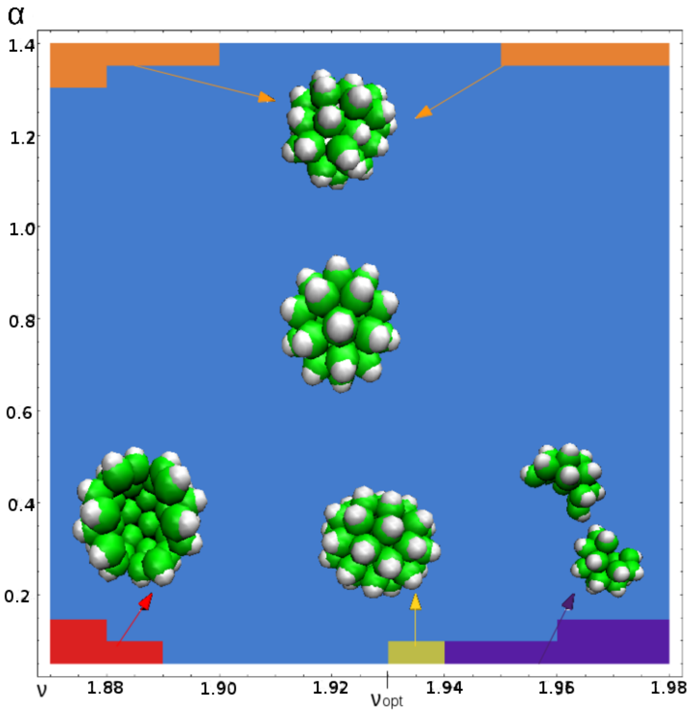


Figure 6.13: Phase diagram of the different structures obtained from a $T = 3$ as a function of α and ν, for $N = 32$, $k_{tor} = 1.5$ and $T = 0.05$.

The case of $T = 4$ is special. In principle, this structure is not an optimal configuration for capsids made out of a single type of capsomer [21]. As a consequence, the phase diagram of $T = 4$ shows another typical morphology. For small α (i.e. large bending penalty) and $\nu < \nu_{opt}$ (less than optimal curvature), we obtain as optimal morphology a shell corresponding to the optimal structure of $N = 44$ but with two missing capsomers, normally one in front of the other. As described in Chapter 7, all-pentamer $T = 4$ structures do not self-assemble spontaneously, due to the strong competence of two structures very close in size and in energy, corresponding to $N = 40$ and $N = 44$, which have cubic simmetry [21]. It seems that, in this range of ν and α, it is more stable to have a 44 structure with 2 missing capsomers than a $T = 4$ structure.

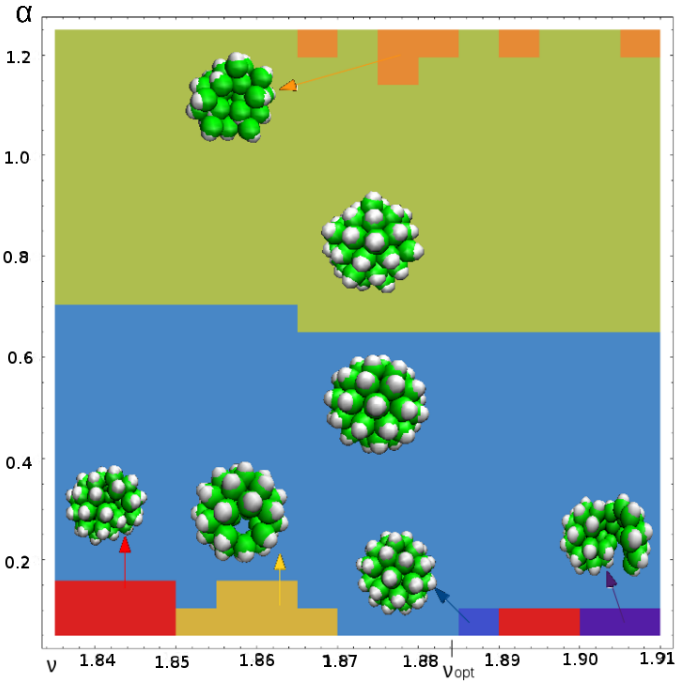


Figure 6.14: Phase diagram of the different structures obtained from a $T = 4$ as a function of α and ν, for $N = 42$, $k_{tor} = 1.5$ and T=0.05.

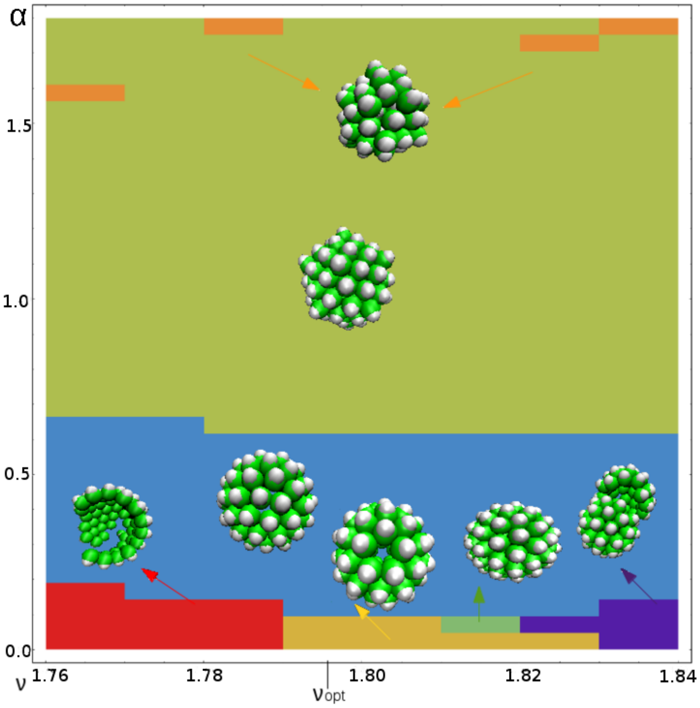


Figure 6.15: Phase diagram of the different structures obtained from a $T = 7$ as a function of α and ν, for $N = 72$, $k_{tor} = 1.5$ and T=0.05.

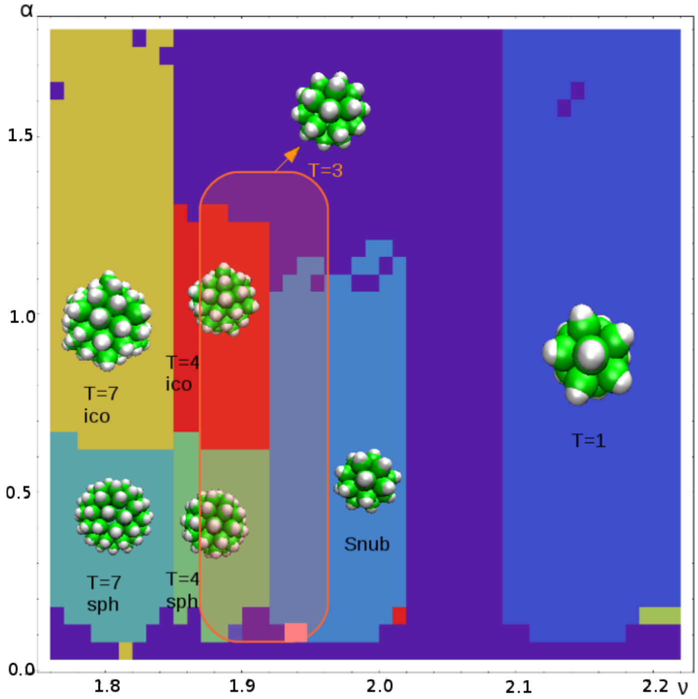


Figure 6.16: Complete stability phase diagram of the different structures obtained as a function of α and ν, for $k_{tor} = 1.5$ and T=0.05. From small to big ν and from small to large α: $T = 7$; elongated (yellow), spherical (blue) and icosahedral (dark yellow); $T = 4$, elongated (dark blue), spherical (green), and icosahedral (red); $T = 3$, elongated (red), and spherical (orange); snub cube, elongated (light orange) and spherical (blue) and $T = 1$, elongated (light green) and spherical (dark blue).

Combining the phase diagrams of each structure, we have compiled the complete stability phase diagram of the dominating structures in the range $0.03 < \alpha < 1.80$ and $1.76 < \nu < 2.22$ (see Fig. 6.16). In the range where different structures compete, the diagram shows the one with the lowest energy. We can see that the preferred angle ν is the main parameter selecting a particular structure. In order from smaller to larger ν, the optimal shells are $T = 7$, then $T = 4$ in a very narrow range, $T = 3$, the snub cube, and finally $T = 1$. There is a region of preferred angles ν where none of these structures is stable (although other local optimal structures like $N = 27$ or $N = 16$ could be stable) (see Fig. 6.16, violet region). At $\alpha \sim 0.65$, $T = 7$ and $T = 4$ have a transition between a spherical shell (at small α) and faceted structures (at large α).

Elongated structures appear at $\nu > \nu_{opt}$ and a narrow range of small α before the structure breaks. In particular, $T = 1$ and the snub cube are elongated for $\alpha = 0.1$ (Fig. 6.16, orange and light green regions). $T = 3$ and $T = 4$ adopt an elongated morphology for $\alpha = 0.05$ (Fig. 6.16, red and blue regions). Finally, the

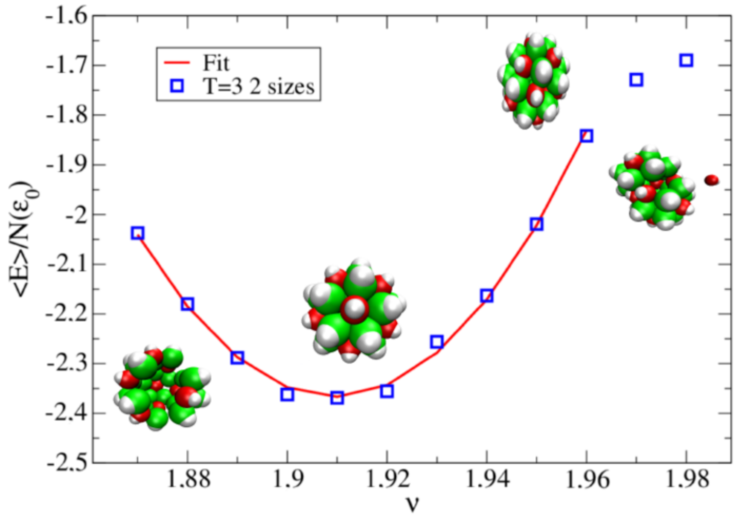


Figure 6.17: Average energy per capsomer versus ν angle with $\alpha = 0.10$, $k_{tor} = 1.5$ and T=0.05 for a $T = 3$ structure made of two different kinds of subunits (pentamers in red and hexamers in green).

elongated $T = 7$ structure is obtained for $\alpha = 0.03$.

In terms of their tolerance in changes of parameters, Fig. 6.16 shows that $T = 1$ and $T = 7$ are the structures which tolerate largest changes both in α and ν. $T = 3$ capsids are also in principle stable in a wide range of angles, but its stability range is strongly limited by the competence with $T = 4$ and the snub cube. That might be one of the reasons why polyoma forms snub cubes instead of $T = 3$ *in vitro*. $T = 4$ has the smallest stability region in the phase diagram due to its non-optimal nature.

6.4 Stability of capsids made by two types of capsomers

In this section we discuss the results obtained when two different types of capsomers, pentamers and hexamers, are considered. This is a more general scenario particularly valid for viruses that form from pentamers and hexamers previously assembled in solution, as it is the case of HK97 [27]. In our model, the only difference between pentamers and hexamers is that they differ in size (see Fig. 3.3). The ratio between their effective radii is $\frac{\sigma_p}{\sigma_h} = \frac{tan(\pi/6)}{tan(\pi/5)}$, and was determined by inscribing their equatorial circles, respectively, into a pentagon or a hexagon of the same edge, as described in section 3.3.

We have assumed that the strength of the binding interactions between hexamers, pentamers, and between hexamers and pentamers, are all the same. We used the same MC algorithm to analyze the stability of these structures in the

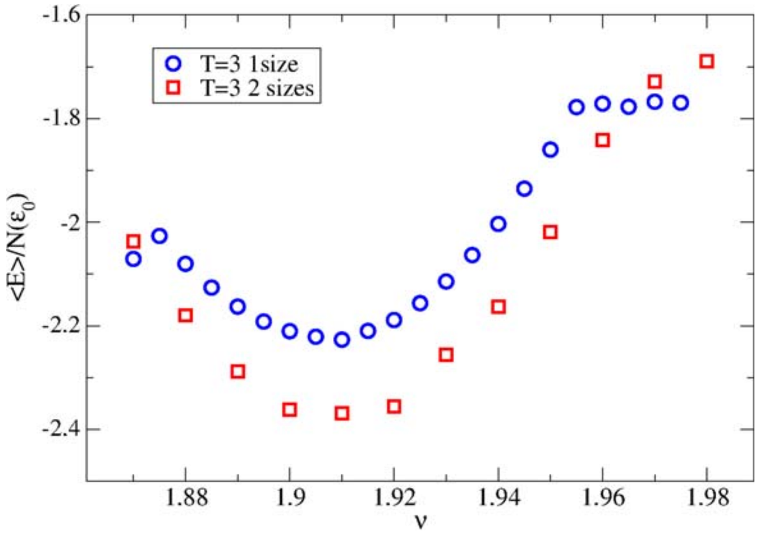


Figure 6.18: Average energy per capsomer versus ν for $\alpha = 0.10$, $k_{tor} = 1.5$ and T=0.05, for a $T = 3$ structure with only one (blue circles) o two types of subunits (red squares).

range of parameters α and ν. In all simulations, the density $\rho = 0.1$, temperature T=0.05, torsion constant $k_{tor} = 1.5$, total number of MC steps 10^7, and sampling frequency (every 10^5 steps), were fixed to the same values as in our previous analysis. In this case, the initial configurations used in our MC simulations were obtained from the the coordinates of the icosahedral spherical code corresponding to the solution of the maximum volume problem in Ref. [23]. Using this model, we repeated the previous stability analysis performed for all-pentamer shells, focusing on the differences associated to the presence of two different morphological units.

Influence of the preferred angle ν

First, we studied the behavior of the different structures with ν as we did for the case of one type of morphological subunit. We have obtained the same qualitative behavior of the average energy versus ν, and the same kind of aberrant as for the case of all pentamer viruses. The main differences are that, in general, the structures with two capsomers types tolerate larger changes in ν and their average energy per capsomer is always smaller than that for the same structure with all subunits of the same size. As an example, Figs. 6.17 and 6.18 show the results for the average energy per capsomer as a function of ν for a $T = 3$ structure at fixed $\alpha = 0.1$. The energy also exhibits a parabolic behavior as a function of ν around and optimal value ν_0. The results of a parabolic fit for each structure at $\alpha = 0.1$ are compiled in Table 6.3.

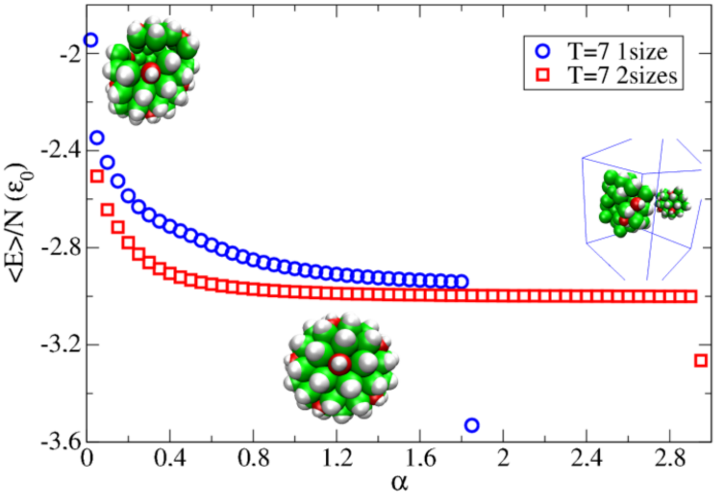


Figure 6.19: Average energy per capsomer versus α for the $T = 7$ structure with $\nu = 1.79667$, $N = 72$, $k_{tor} = 1.5$ and T=0.05 made by two different subunits (hexamers in green and pentamers in red) differing in their size.

Influence of the bending stiffness parameter α

In this case, we have fixed the value of ν in each structure to ν_{opt} and we changed the parameter α, associated to the bending stiffness. One again, comparing with the results in the previous section for all pentamer viruses we have observed the same qualitative behavior of the energy and aberrant structures at too large α (compact clusters or concentric layers) or too small α (disrupted shells). There are, however, two noticeable differences. First, the average energy per capsomer is smaller for shells made of hexamers and pentamers and the structure tolerates larger changes in α, specially the bigger structures, as shown in Table 6.3 and in Fig. 6.19 for $T = 7$. The second important difference is that we did not observe any faceted structure, in contrast to the buckling obtained for $T = 4$ and $T = 7$ with all subunits of the same size for large values of α. This is most probably due to the fact that the smaller size of pentamers alleviates the accumulated stress [37], which is the main driving force for the buckling.

Stability phase diagram

In the previous subsections, simulations were performed either at α or at ν fixed. Now, we construct the full phase diagram of each structure with two capsomers types in the range of values of α and ν where they keep their integrity at steps of 0.05 in α and 0.01 in ν. Obviously, the phase diagram for $T = 1$, made exclusively by pentamers, is the same as the one in the previous section, see Fig. 6.11. In addition, we did not analyze the stability of snub cubes structures because this structure with two different subunits is not optimal and has not been observed

Structure	ν range	α range	ν_0	ν_{opt}	E_0/N (ϵ_0)	K
T=1	2.09-2.22	0.03-1.81	2.125	2.12437	-2.12	362
T=3	1.87-1.99	0.06-1.90	1.91	1.89698	-2.37	419
T=4	1.84-1.92	0.06-2.10	1.86	1.84758	-2.45	426
T=7	1.75-1.87	0.03-2.90	1.79	1.78005	-2.64	485

Table 6.3: Stability range and parameters for the different structures with two capsomer types. The columns indicate: the range of ν at $\alpha = 0.10$, and the range of α at $\nu = \nu_{opt}$, at which each structure maintains its integrity; the predicted value of the angle ν_{opt} for each structure (see Table 3.1); the optimal preferred angle ν_0, the optimal average energy per capsomer E_0/N, and the effective stiffness K, obtained from the parabolic fit of the energy per capsomer vs ν for each structure at $\alpha = 0.1$.

in experiments [21].

Figs. 6.20 to 6.22, show the phase diagrams obtained for $T = 3$, $T = 4$ and $T = 7$, with two types of capsomers. The figures show the most characteristic stable morphologies and also the typical aberrant structures. We can see, in all cases, 5 different morphologies, illustrated as an example, for the case of $T = 7$ in Fig. 6.22: 1) the perfect spherical structure (indicated by the blue region), 2) open aberrant structures, smaller than the normal one (indicated by the violet region) obtained at small α and large ν, 3) burst structures (indicated in red) at small α and ν , 4) clusters or multi-layered structures (the orange region in Fig. 6.22) when the bending penalty is low, and finally 5) elongated capsids (yellow regions). Remarkably, unlike the simulations with only one type of capsomers, here no faceted structures appear in the phase diagram, not even for the largest $T = 4$ and $T = 7$. In this case, the presence of the smaller pentamers prevents the accumulation of stress that triggers the faceting of the shells.

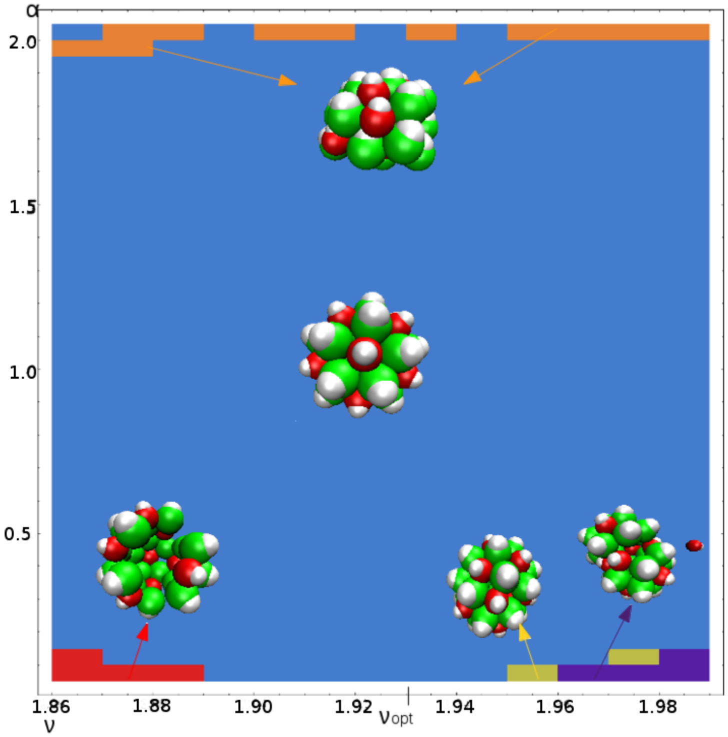


Figure 6.20: Phase diagram of the different structures obtained from a $T = 3$ with pentamers (in red) and hexamers (in green) as a function of α and ν, for $N = 32$, $k_{tor} = 1.5$ and T=0.05.

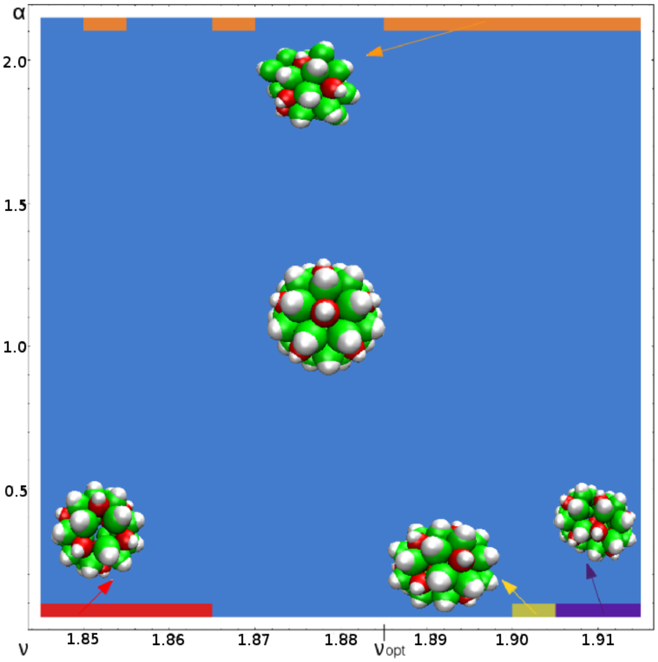


Figure 6.21: Phase diagram of the different structures obtained from a $T = 4$ with pentamers (red spheres) and hexamers (green spheres) as a function of α and ν, for $N = 42$, $k_{tor} = 1.5$ and T=0.05.

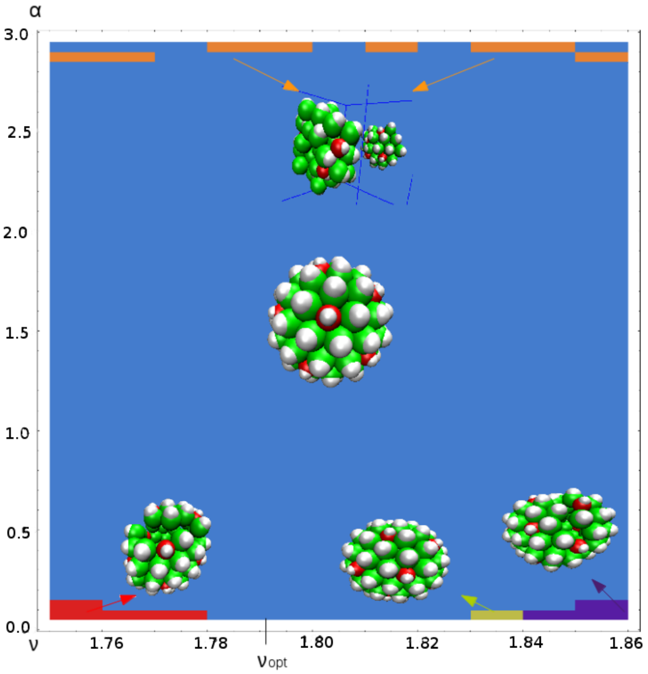


Figure 6.22: Phase diagram of the different structures obtained from a $T = 7$ with pentamers (in red) and hexamers (in green) as a function of α and ν, for $N = 72$, $k_{tor} = 1.5$ and T=0.05.

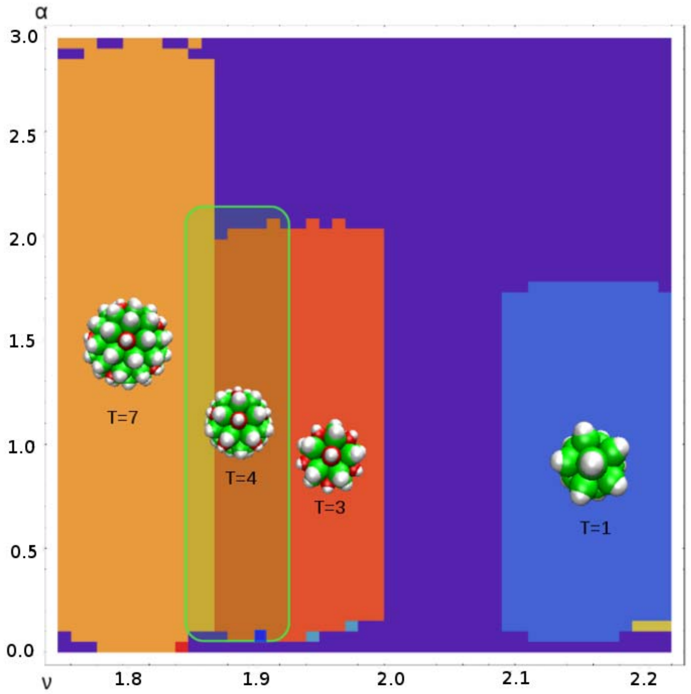


Figure 6.23: Complete stability phase diagram of the different structures obtained as a function of α and ν, for $k_{tor} = 1.5$ and T=0.05. From small to big ν and from small to large α: $T = 7$; elongated (red) and spherical (light orange); $T = 4$, elongated (dark blue) and spherical (light green); $T = 3$, elongated (light blue) and spherical (orange) and $T = 1$, elongated (yellow) and spherical (dark blue).

Putting together the phase diagrams of each structure, we have constructed the complete stability phase diagram of the dominating structures with two capsomer types in the range $0.02 < \alpha < 3.0$ and $1.75 < \nu < 2.22$ (see Fig. 6.23). In the range where different structures compete, the diagram shows the one with the lowest energy. Again we see that the preferred angle ν is the main parameter selecting a particular structure. In order from smaller to larger ν, the optimal shells are $T = 7$ (light orange), then $T = 4$ (light green), $T = 3$ (orange) and finally $T = 1$ (dark blue). There is a region of preferred angles ν where none of these structures is stable (violet).

There is a significant overlap of $T = 7$ and $T = 4$ shells for angles in the range $\nu \in [1.845, 1.86]$ and of $T = 4$ and $T = 3$ capsids in the range $\nu \in [1.87, 1.91]$. Elongated structures appear at $\nu > \nu_{opt}$ and relatively small α but strong enough to prevent their breaking. $T = 1$ is elongated for $\alpha = 0.1$ (see Fig. 6.23 yellow region). $T = 3$ and $T = 4$ have elongated morphology for $\alpha = 0.05$ (see Fig. 6.23 light blue and blue regions). Finally, an elongated version of $T = 7$ is obtained for $\alpha = 0.03$ (in red).

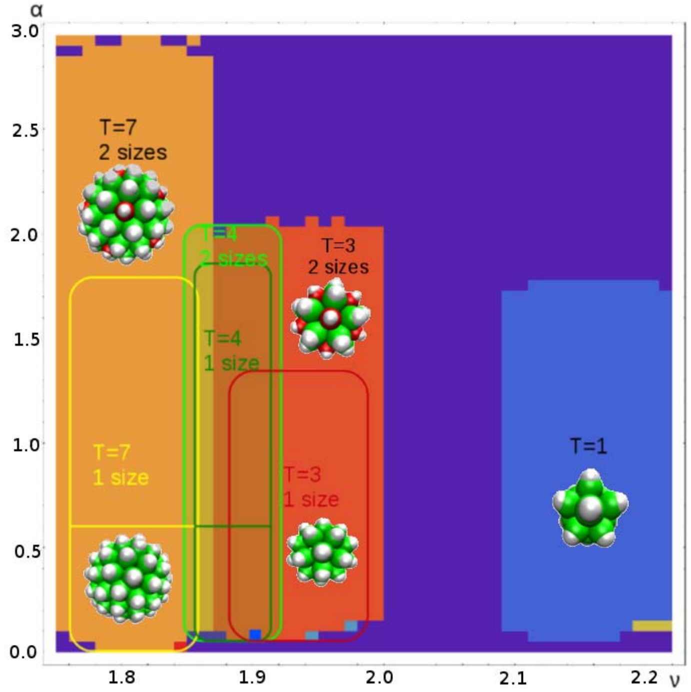


Figure 6.24: Comparison of the stability phase diagrams of the different structures obtained as a function of α and ν, for one type and two types of capsomers, and $k_{tor} = 1.5$ and T=0.05. The coloured areas indicate the stability of $T = 7$ (orange and yellow), $T = 4$ (dark and light green), $T = 3$ (orange and red) and $T = 1$ (blue) with pentamers and hexamers, whereas the boxed areas correspond to the stability of all-pentamers viruses.

Fig. 6.24 compares the phase diagrams obtained for capsids with one or two different capsomer types. The most noticeable difference is that for all-pentamers $T = 7$ and $T = 4$ shells we obtained faceted structures at high α, whereas for the same structures but using two types of subunits there are not faceted structures. Moreover, all structures with pentamers and hexamers tolerate larger changes in α and ν keeping its integrity, especially in α. The stability range of $T = 7$ and $T = 4$ structures with hexamers and pentamer overlap, unlike what happened for the same structures with only one type of subunit, where this overlap did not exist.

6.5 Conclusions

In this chapter we have characterized the phase diagram of stability of viral shells corresponding to the lowest CK structures plus the snub cube, using MC simulations. We have verified that our model is able to reproduce the right structures adopted by viruses as energy minima. In addition, we have found the optimal parameters required for stability in each case both for all-pentamers viruses as well as for viruses made by two different morphological units (pentamers and hexamers). We have seen that the parameter α associated to the bending stiffness, and specially the preferred angle ν of capsomer-capsomer interaction are the main parameters controlling the selection of a specific structure by a virus. These macroscopic physical parameters are directly linked to the interaction between proteins, that can be in turn modified by changing the assembly conditions. Thus our model provide useful insight on how the modification of assembly conditions will determine the fate of the assembly.

In this context, a relevant example is Polyoma, whose capsid is made exclusively by pentamers, that can be assembled *in vitro* forming $T = 1$, snub cube and $T = 7$ structures [19]. It is remarkable that this virus adopts a snub cube rather than a $T = 3$ structure. One possible explanation for this fact could be the overlap in the phase diagram between the snub cube, $T = 3$ and $T = 4$ observed in Fig. 6.16. This diagram shows that for a specific range of values of the preferred angle between capsomers, a snub cube structure would be preferred over a $T = 3$, as occurs in the case of polyomavirus.

We have also seen that the model reproduces not only the optimal spherical structures of capsids, but also elongated, faceted and decapsidated structures that are also characteristic of many viruses. There are many studies of polymorphism in viral capsids, where spherical and elongated shapes appear as a function of environmental conditions, e.g., salt concentration or pH. Some examples are CCMV, alphalpha mosaic virus, polyoma or Simian 40 virus. Especially interesting for us are the polyoma virus and Simian 40 virus because these two virus assemble only with pentamers. Indeed, we have located small regions in the $\nu\alpha$ phase diagram for all structures where the preferred structure is elongated. In all cases, this occurs for small α (or large bending constant k_b) and ν bigger than optimal angle ν_{opt}, in agreement with theoretical predictions.

We have also observed the faceting of $T = 4$ and $T = 7$ all-pentamers structures occurring at a critical value of α that nicely coincides with the prediction of where the buckling transition occurs. For $T = 3$ and $T = 7$ we have also obtained decapsidation by the loss of pentamers, and not hexamers, because pentamers have only 5 neighbors and they suffer more stress than the hexamer for these structures [37].

Thus our model could yield very useful insight on the physical mechanisms regulating viral assembly. Moreover, we have also characterized what happens to the structures when the parameters are not right, leading to aberrant morphologies. These aberrant structures, and the physical mechanisms involved in their formation, could also provide guidance to find the right conditions for the correct assembly of viruses, or alternatively, the most relevant factors that will lead to misassembled and inactive virus.

The results found in this chapter refer to an equilibrium phase diagram of stability of preformed structures. But viral self-assembly is a truly kinetic non-equilibrium process. The actual kinetics of assembly and the modifications that it introduces in the picture are the topic of the next chapter.

References

[1] RA. Crowther, NA. Kiselev, B. Bottcher, JA. Berriman, GP. Borisova, V. Ose and P. Pumpens, Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell*, **77**, 943-950, 1994.

[2] 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**, 559-568, 1998.

[3] KW. Adolph and PJG. Butler, Studies on the assembly of a spherical plant virus (I. States of aggregation of the isolated protein). *J. Mol. Biol.*, **88**, 327-338, 1974.

[4] L. Lavelle, M. Gingery, M. Phillips, W. M. Gelbart, C. M. Knobler, R. D. Cadena-Nava, J. R. Vega-Acosta, L. A. Pinedo-Torres and J. Ruiz-Garcia, Phase diagram of self-assembled viral capsid protein polymorphs, *J. Phys. Chem. B*, 2009.

[5] C. Chen, C. C. Kao, and B. Dragnea, Self-assembly of brome mosaic virus capsids: insights from shorter time-scale experiments, *J. Phys. Chem. A*, **112**, 9405-9412, 2008.

[6] M. Comellas-Aragonès, *et al.*, A virus-based single-enzyme nanoreactor, *Nat. Nanotech.*, **2**, 635-639, 2007.

[7] SW. Lee, CB. Mao, CE. Flynn and AM. Belcher, Ordering of quantum dots using genetically engineered viruses, *Science*, **296**, 892-895, 2002.

[8] T. Douglas, M. Young, Host-guest encapsulation of materials by assembled virus protein cages, *Nature*, **393**, 152-155, 1998.

[9] SK. Dixit, *et al.*, Quantum dot encapsulation in viral capsids, *Nano letters*, **6**, 1993-1999, 2006.

[10] C. Uetrecht , C. Versluis, NR. Watts, WH. Roos, GJL. Wuite, PT. Wingfield, AC. Steven and AJR. Heck, High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids, *PNAS*, **105**, 9216-9220, 2008.

[11] SN. Kanesashi, K. Ishizu, M. Kawano, S. Han, S. Tomita, H. Watanabe, K. Kataoka and K. Handa , Simian virus 40 VP1 capsid protein forms polymorphic assemblies *in vitro*, *J. Gen. Virol.*, **84**, 1899-1905, 2003.

[12] GL. Casini, D. Graham, D. Heine, RL. Garcea and DT. Wu, In vitro papillomavirus capsid assembly analyzed by light scattering, *Virology*, **325**, 320-327, 2004.

[13] S. Mukherjee, S. Kler, A. Oppenheim, and A. Zlotnick, Uncatalyzed assembly of spherical particles from SV40 VP1 pentamers and linear dsDNA incorporates both low and high cooperativity elements, *Virology*, **397**, 199-204, 2010.

[14] C. Uetrecht, IM. Barbu, GK. Shoemaker, E. van Duijn and AJR. Heck, Interrogating viral capsid assembly with ion mobility-mass spectrometry, *Nature Chemistry*, **3**, 126-132, 2010.

[15] 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, 2000.

[16] A. Zlotnick and S. J. Stray, How does your virus grow? Understanding and interfering with virus assembly, *Trends Biotechnol.*, **21**, 536-542, 2003.

[17] S. Singh and A. Zlotnick, Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly, *J. Biol. Chem.*, **278**, 18249 -18255, 2003.

[18] R. Zandi, P. van der Schoot, D. Reguera, W. Kegel and H. Reiss, Classical nucleation theory of virus capsids, *Biophys. J.*, **90**, 1939-1948, 2006.

[19] DM. Salunke, DLD. Caspar and RL. Garcea, Polymorphism in the assembly of polyomavirus capsid protein VP1, *Biophys. J.*, **56**, 887-900, 1989.

[20] YL. Miao, JE. Johnson and PJ. Ortoleva, All-Atom Multiscale Simulation of Cowpea Chlorotic Mottle Virus Capsid Swelling, *J. Phys Chem. B*, **114**, 11181-11195, 2010.

[21] R. Zandi, D. Reguera, RF. Bruinsma, WM. Gelbart and J. Rudnick, Origin of icosahedral symmetry in viruses, *PNAS*, **101**, 15556-15560, 2004.

[22] A. Luque, Thesis. Structure, Mechanical Properties, and Self-Assembly of Viral Capsids, 2011.

[23] RH. Hardin, NJA. Sloane and WD. Smith, Tables of spherical codes with icosahedral symmetry, published electronically at http://www.research.att.com/∼njas/icosahedral.codes/.

[24] GL. Casini, D. Graham, D. Heine, RL. Garcea, and DT. Wu, In vitro papillomavirus capsid assembly analyzed by light scattering, *Virology*, **325**, 320-327, 2004.

[25] W. Zhang, NH. Olson, TS. Baker, L. Faulkner, M. Agbandje-McKenna, MI. Boulton, JW. Davies, and R. McKenna, Structure of the maize streak virus geminate particle, *Virology*, **279**, 471-477, 2001.

[26] S. Mukherjee, S. Kler, A. Oppenheim, and A. Zlotnick, Uncatalyzed assembly of spherical particles from SV40 VP1 pentamers and linear dsDNA incorporates both low and high cooperativity elements, *Virology*, **397**, 199-204, 2010.

[27] Z. Xie and RW. Hendrix, Assembly in vitro of bacteriophage HK97 proheads, *J. Mol. Biol.*, **253**, 74-85, 1995.

[28] RR. Rueckert, Fields, B. N. and D. M. Knipe (ED.). Fundamental Virology, Second Edition. New York, USA. Illus. *International Congress Series* 409-450, 1991.

[29] T. Nguyen, RF. Bruinsma and WM. Gelbart, Elasticity theory and shape transitions of viral shells, *Phys. Rev. E.*, **72**, 051923, 2005.

[30] J. Lidmar, L. Mirny and DR. Nelson, Virus shapes and buckling transitions in spherical shells, *Phys. Rev. E.*, **68**, 051910, 2003.

[31] M. Adrian, PA. Timmins and J. Witz, In vitro decapsidation of turnip yellow mosaic virus investigated by cryo-electron microscopy: a model for the decapsidation of a small isometric virus, *J. Gen. Virol.*, **73**, 2079-2083, 1992.

[32] K. Mutombo, B. Michels, H. Ott, R. Cerf and J. Witz, The Thermal-Stability and Decapsidation Mechanism of Tymovirus - a Differential Calorimetric Study, *Biochimie*, **75**, 667-674, 1993.

[33] Y. Stauffer, K. Raj, K. Masternak and P. Beard, Infectious human papillomavirus type 18 pseudovirions, *J. Molec. Biol.* , **270** 2, 238-246, 1998.

[34] K. Münger, A. Baldwin, KM. Edwards, H. Hayakawa, CL. Nguyen, M. Owens, M. Grace and K. Huh, Mechanisms of Human Papillomavirus-Induced Oncogenesis, *J Virol* , **78** , (21), 11451-11460, 2004.

[35] E. Frost, D. Bourgaux-Ramoisy and P. Bourgaux, Decapsidation of polyoma virus mutants, *Virology*, **68** , 1, 256-259, 1975.

[36] KW. Adolph and PJG. Butler, Studies on the assembly of a spherical plant virus (I. States of aggregation of the isolated protein), *J. Mol. Biol.*, **88**, 327-338, 1974.

[37] M. Aznar A. Luque and D. Reguera, Relevance of capsid structure in the buckling and maturation of spherical viruses, *Phys. Biol.*, **9**, 036003, 2012.

CHAPTER 7

Simulations of the assembly kinetics of empty viral shells

7.1 Introduction

In the previous chapter, we analyzed the range of bending stiffness and preferred angles for which the different capsids in our coarse-grained model are expected to be stable. But the stability of a preformed capsid and the conditions required for its assembly are in general different. The formation of a capsid is in reality a kinetic process and there is no warranty that conditions at which a capsid can keep its integrity will be ideal for its formation out of a solution of free subunits.

In this chapter we are going to analyze the kinetics of capsid formation using BD simulations. The idea is to mimic, using simulations, the standard conditions and procedure of *in vitro* assembly experiments of empty capsids. We will thus start from a solution containing our model capsomers randomly dispersed at a given concentration and conditions (in terms of T, α and ν) and will monitor the successful formation of stable structures, corresponding to the ones adopted by viruses. In particular, we will focus on the assembly of stable structures in the range of parameters between $T = 1$ and $T = 7$. For the sake of simplicity, we will restrict our simulations to the assembly of empty capsids for which identical capsomers are the fundamental assembly units. This will be the case, for instance, of the assembly of papilloma [1, 2] and polyomavirus [3].

After describing some details of the model and simulations in Section 7.2, the first goal is to confirm that the direct assembly of the different optimal structures is indeed spontaneously feasible with our coarse-grained model (Section 7.3). We will then analyze in Section 7.4 the conditions required for assembly and compare them with those corresponding to the range where preformed capsids are stable. In Section 7.5 the kinetics of assembly in more detail, focusing on the particular case of a $T = 3$ viral shell. We will use simulations to determine the equilibrium conditions, as well as the critical cluster size, the activation barrier and the nucleation rates, and compare them with the CNT predictions. Finally, we summarize the main conclusions in Section 7.5.

7.2 Model and Brownian dynamics simulations

The study of the kinetics of capsid assembly will be done using our coarse-grained model introduced in Chapter 3. In this study, only one type of capsomer will be considered, corresponding to the standard conditions of *in vitro* assembly of viruses such as polyoma, SV40 and papillomavirus [1–5].

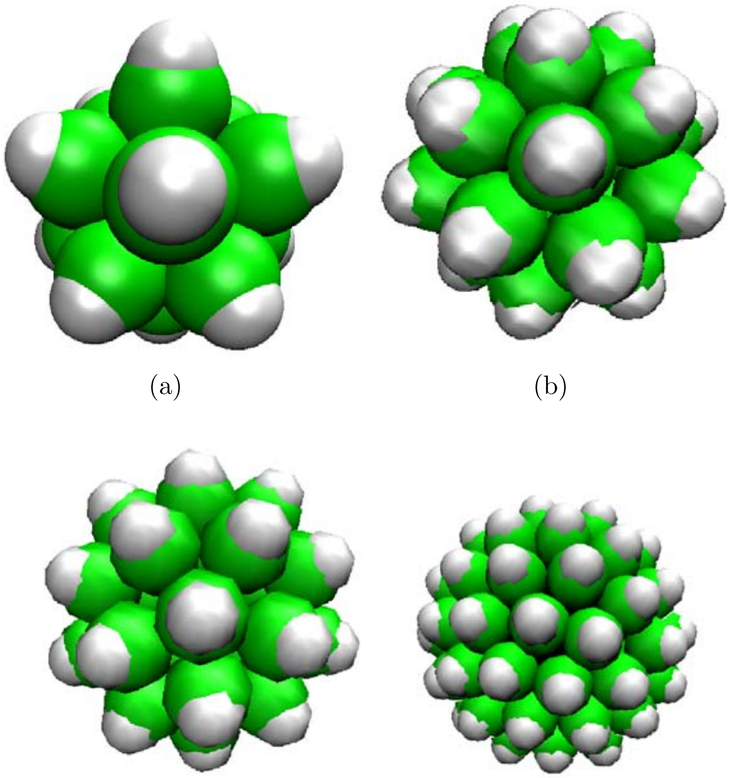
Since in this chapter the interest is on the dynamics of the process, BD simulations will be used instead of equilibrium MC simulations. The general philosophy and details of the simulation are described in Chapter 5. It is important to emphasize that typically the self-assembly of a virus takes place at time scales that range from seconds to minutes or hours. These timescales are not easy to reach in simulations, not even for simple coarse-grained models. Therefore, one should keep in mind that our simulations will be done at much more supersaturated conditions than the typical conditions for assembly *in vivo* or *in vitro*. Nevertheless, the results of the simulations will provide a useful insight on the mechanism and kinetics of viral assembly. In addition, assembly at those extreme conditions will constitute a more stringent test of the validity of the theory.

We implemented four different types of BD simulations. First, some preliminary simulations using a fixed number of capsomers equal to those in the desired structure, were performed. The second type of simulations were done for $N = 100$, and changing the values of the parameters of the potential α and ν to map the phase diagram of the assembly. The third type of simulations were designed to find out the conditions corresponding to equilibrium between free capsomers and a complete $T = 3$ capsid. For those simulations, we started from a half $T = 3$ capsid and monitored whether this structure disassemble or grow for different values of the density. The last type of simulations were targeted to analyze accurately the kinetics of assembly of a $T = 3$ structure. To that end, simulations with $N = 100$ capsomers at fixed α and ν, corresponding to optimal values for the formation of $T = 3$, were repeated with different initial configurations a hundred times. By monitoring the size of the largest cluster in each case, it is possible to evaluate the assembly rate, the critical cluster size, and the nucleation barrier. The specific details of each simulation as well as the results obtained with them will be described in the corresponding section.

7.3 Preliminary simulations of assembly

We first did some preliminary simulations to verify the feasibility of the assembly and to determine the approximate ranges adequate for the assembly of each structure. These preliminary simulations were done using a total number of subunits equal to that of the desired structure (namely, $N = 12$ for $T = 1$, $N = 32$ for $T = 3$, $N = 24$ for the snub cube, and $N = 72$ for $T = 7$). With this strategy, one limits the competence with other aberrant structures, and also the formation of more than one partial structures, so the simulations are faster. In this preliminary simulations, we were able to obtain successful assembly in all cases at $\alpha \approx 0.1$ and values of the preferred angle ν close to the expected optimal values ν_{opt} listed in Table 3.1. The density was set in all cases to $\rho = 0.1$ in reduced units. The structures resulting from these simulations are plotted in Fig. 7.1.

In those preliminary simulations, we analyzed first the effect of the reduced temperature (which in our reduced units is analogous to a change in the binding strength ϵ_0. At very low temperatures, i.e. T=0.05 (corresponding to $\epsilon_0 = 20 k_B T$), the assembly of the structures was not very successful. The system immediately



(a) (b)

(c) (d)

Figure 7.1: Snapshot of the structures obtained at the end of a BD assembly simulation for: (a) a $T = 1$ shell with $N = 12$, (b) a snub cube with $N = 24$, (c) a $T = 3$ shell with $N = 32$, and (d) a $T = 7$ shell with $N = 72$. In all cases, the simulations were done at $\rho = 0.1$, T=0.1, $k_t = 1.5$.

started the formation of multiple partial shells until no more free subunits were available. Thus at these conditions, and specially for many subunits, aberrant and open structures get formed, that for very long simulations runs were sometimes able to merge into the correct structure. This corresponds to very high supersaturation conditions, where all capsomers would like to stick almost irreversible, leading to the formation of many nucleation sites that deplete the free subunits before any of them is able to form a complete capsid. That is why we decided to increase the reduced temperature, which is equivalent to weakening the binding interactions. This is in line with previous observations that weak interactions are enough and more efficient to warrant viral assembly [6]. We were able to get consistently high yields of correctly assembled structures for T$\approx$ 0.1 (see Fig. 7.1). Accordingly, further analysis of assembly were performed around this temperature. We also performed simulations at high values of the temperature. In that case, for $T > 0.2$ no closed structures were formed in the limited time span of the simulations (typically, 10^8 steps) and essentially no significant clustering was observed.

Those temperatures correspond to too weak binding strengths $\epsilon_0 < 5k_BT$, that will require very high subunit concentrations or extremely long times to nucleate into closed shells.

7.4 Phase diagrams of the assembly of all-pentamer viruses

Once established the appropriate range of temperatures and the feasibility of the assembly, the interest was to construct a phase diagram of assembly, identifying the regions in terms of the parameters α and ν for which successful assembly is achieved at given conditions. For those simulations, we used a total of N=100 capsomers that were initially placed at random positions and orientations inside a cubic box with periodic boundary conditions. The density of capsomers was in all cases fixed at $\rho = 0.1$, to enhance and speed up the assembly. Other parameters that were fixed in the simulations were $dt = 10^{-4}$ and $k_t = 1.5$. In this case, the simulations were run for a total of 10^8 steps. A cluster algorithm was used to identify the size of the largest cluster in the system, and the final outcome of the simulation (whether there was assembly or not, and the size of the resulting shells) was monitored for each value of α and ν. Simulations were repeated for a whole range of $\alpha \in [0.09, 0.15]$ and $\nu \in [1.78, 2.15]$, sampled at steps of 0.01 in α and of 0.005 in ν. With that, we constructed the assembly phase diagrams at three different temperatures. In principle, the assembly is an stochastic process, thus many repetitions of each simulation at any given conditions will be required to get significative results. However, in our preliminary studies, and mainly by lack of time, only one repetition was made for each conditions.

7.4.1 Assembly phase diagram at temperature T$= 0.09$

First, we analyzed the phase diagram of assembly at relatively high binding strengths, corresponding to a reduced temperature T=0.09. The explored range of $\nu \in [1.78, 2.15]$ encompasses the values of ν_{opt} corresponding to structures from $T = 1$ to $T = 7$, so we expect to find these structures in our simulations. In our preliminary simulations the assembly of every structure happened at $\alpha \sim 0.12$, so in the exploration of the phase diagram the range $\alpha \in [0.09, 0.15]$ was choosed as appropriate.

Fig. 7.2 shows the resulting phase diagram of assembly, where the different regions corresponding to the successful assembly of the CK plus the snub cube structures are indicated by different colors. In particular, the assembly of $T = 1$ was obtained in the range $\alpha \in [0.11, 0.15]$ and $\nu \in [2.055, 2.15]$ (light green region); snub cubes were found in the region $\alpha \in [0.09, 0.15]$ and $\nu \in [1.925, 1.980]$ (in red); $T = 3$ was recovered between $\alpha \in [0.09, 0.15]$ and $\nu \in [1.875, 1.935]$ (violet) and a $T = 7$ structure was the result at $\nu = 1.780$ and $\sigma = 0.14$ (orange).

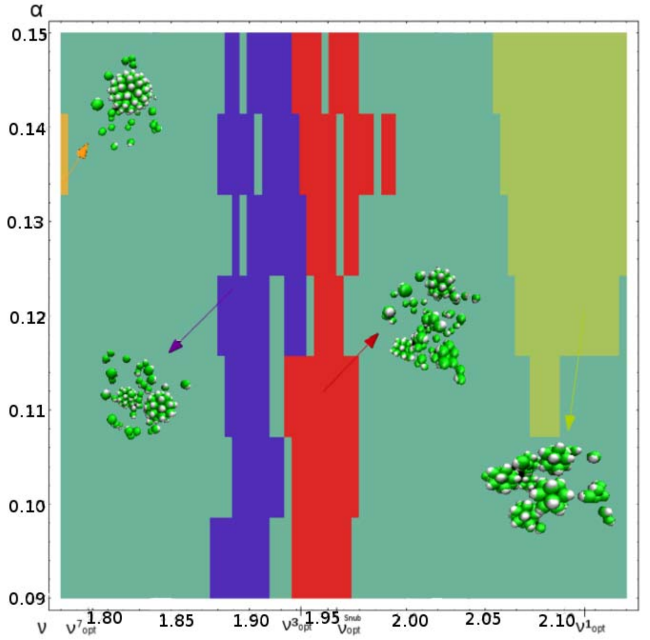


Figure 7.2: Assembly phase diagram of the different structures obtained as function α and ν at T=0.09. The region of successful assembly of each structure is indicated by different colors: orange for $T = 7$, violet for $T = 3$, red for the snub cube, and light green for $T = 1$. The figure also includes a representative snapshot of the simulation in each case.

An important general observation is that, for every structure, the optimal assembly conditions at this temperature were found at values of ν smaller than the expected ν_{opt}, listed in Table 3.1. That indicates that a slightly larger radius of curvature is required for optimal assembly, as suggested by previous studies of templated and equilibrated assembly [7]. This slightly larger radius of curvature will facilitate the incorporation of the right number of subunits and prevent the "closure catastrophe" [7].

A second remark from the simulations is that, for the smallest structures, multiple mono-disperse structures were obtained at the end of the simulation. For instance near the optimal conditions for $T = 3$, up to three $T = 3$ shells, using almost all the $N = 100$ initial capsomers, were found at the end of the simulations.

As discussed in the previous chapter, the preferred angle ν is the main parameter in the selection of a particular structure. But in contrast to the stability simulations, the range of values of ν and α where each structure can self-assemble from solution, is smaller. This is particularly evident for T=7 that appears only at a very limited range of parameters.

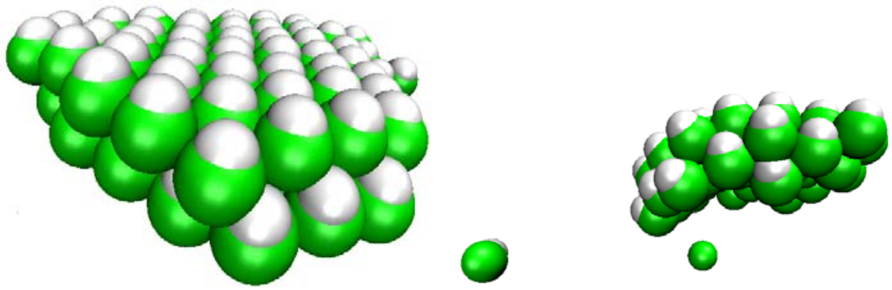

(a) (b)

Figure 7.3: Snapshot of different layers obtained in our simulations for large values of α at T=0.09. (a) Hexagonally-ordered bilayer obtained at $\alpha = 0.75$ and $\nu = 1.795$. (b) Portion of a curved shell obtained at $\alpha = 1.0$ and $\nu = 1.92$.

We have also explored the outcome of the simulations for values of α outside the range shown in Fig. 7.2. For very small values of α, the subunits essentially do not aggregate easily or form small open clusters. A small value of α means a very large bending penalty and a very restrictive angular conditions for binding that difficult the formation of large structures. On the other hand, for big values of α, e.g. $\alpha = 0.75, 1.0$ or 1.25, the local bending stiffness is very small, and the potential is more isotropic. As a result, the typical outcome of the simulations at those conditions are multilayer structures that tend to be hexagonally ordered specially at small preferred angles ν, as shown in Fig. 7.3(a). For large ν, i.e. favouring more curvature, curved multishells or compact clusters often form, see Fig. 7.3(b).

7.4.2 Assembly phase diagram at temperature T$= 0.10$

Fig. 7.4 shows the assembly phase diagram obtained when our simulations were repeated at a slightly higher temperature T=0.10. At this temperature, we have obtained $T = 7$ shells only at $\alpha = 0.14$ and $\nu = 1.800$, indicated by the orange region in Fig. 7.4. $T = 3$ shells were found at all values of α between $\nu \in [1.880, 1.915]$ (violet region) and snub cubes at all α between $\nu \in [1.925, 1.955]$ (red region). Remarkably, no $T = 1$ were formed in the explored range. This can be understood using the ideas of CNT described in Chapter 2. In particular, it was shown that the driving force for the assembly is the chemical potential, which is related to the effective binding energy per capsomer Δg. At higher reduced temperatures, the effective binding energy between capsomers decreases. In addition, as described in Chapter 6, $T = 1$ shells are the ones having the highest energy per capsomer Δg. The combination of these two factors is the underlying reason why assembly of $T = 1$ structures at high temperatures and at fixed densities, might not be favored.

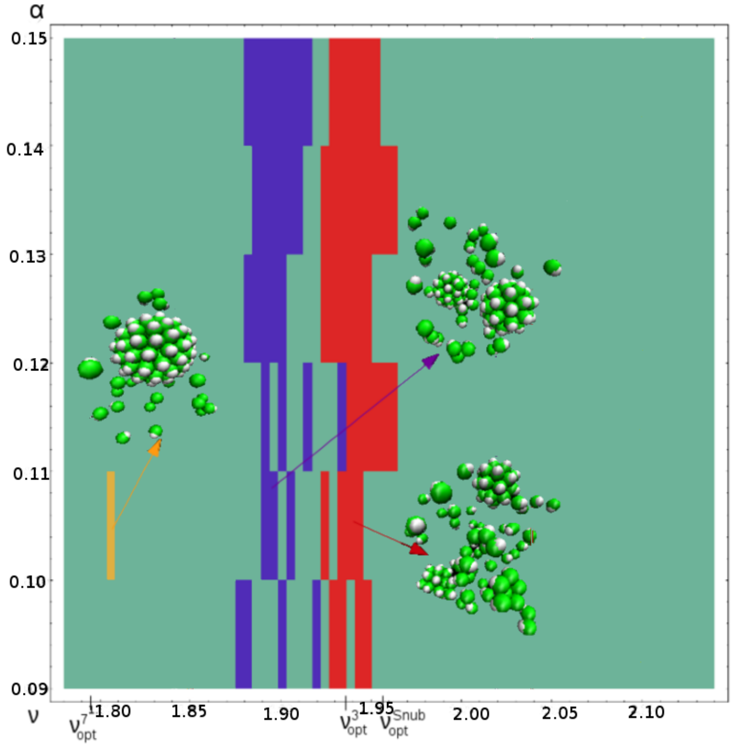


Figure 7.4: Assembly phase diagram of the different structures obtained as function α and ν at T=0.10. The region of successful assembly of each structure is indicated by different colors: orange for $T = 7$, violet for $T = 3$, and red for the snub cube. The figure also includs a representative snapshot of the structures in each case.

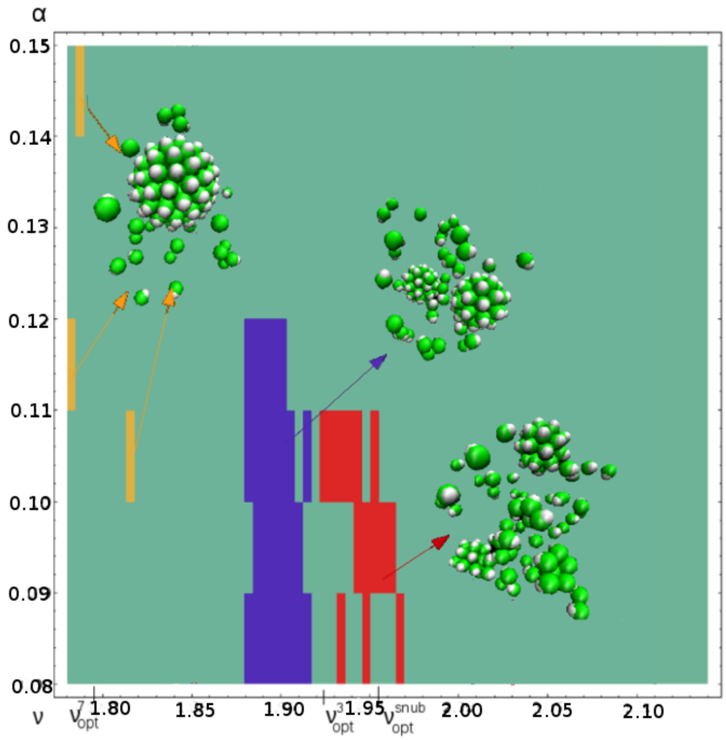


Figure 7.5: Assembly phase diagram of the different structures obtained as function α and ν at T=0.12. The region of successful assembly of each structure is indicated by different colors: orange for $T = 7$, violet for $T = 3$, and red for the snub cube. The figure also includ a representative snapshot in each case.

7.4.3 Assembly phase diagram at temperature T$= 0.12$

Finally Fig. 7.4 illustrates the assembly phase diagram obtained at a temperature T=0.12. In this case, $T = 7$ shells have been found at regions corresponding to $\alpha = 0.09$ - $\nu = 1.785$, $\alpha = 0.12$ - $\nu = 1.780$, and $\alpha = 0.13$ - $\nu = 1.810$, indicated in orange in Fig. 7.4. To get $T = 3$ capsids (violet regions in Fig. 7.4) α must be $\alpha \geq 0.12$, since in the range $\alpha < 0.12$ the structure does not form consistently. The preferred angle is $\nu \in [1.880, 1.915]$, in the same range found for lower temperatures. In turn, snub cubes appear for $\alpha \geq 0.13$ and $\nu \in [1.925, 1.955]$ (red region in Fig. 7.5). Once again, $T = 1$ structures are not present at high temperatures. In general, it is evident that the range of parameters that guarantee the assembly of each structure is significantly reduced, since at high temperatures, the weaker binding energies and the larger thermal fluctuations difficult the clustering process.

7.4.4 Aberrant and non Caspar-Klug structures

At all temperatures, outside the regions where the $T = 1, 3, 7$ and snub cubes were formed, some other structures compete and appear consistently in the simulations. Very often, they are aberrant structures similar to the CK ones, but either disordered or with excess or deficit of capsomers. This is specially the case at the lower temperature T=0.09, where the strong relative strength of the binding interactions favors the clustering and formation of closed structures basically at all values of the spontaneous curvature ν. In many cases the closed structures that get formed are nicely regular, appear consistently, and they turn out to be locally optimal arrangements of capsomers in the free energy landscape of one type of capsomer [8, 9]. Specifically, closed shells with 20, 27, 38, 44, 48, 54 and 60 subunits often appear at all simulated temperatures. The size (i.e. number of capsomers) of these structures is controlled by the value of ν, having more subunits, the smaller the value of ν. Remarkably, the regular non-CK structures that get formed nicely coincide with the deepest minima in the free energy landscape of the templated model in Ref. [9], see Fig. 7.6. At non-optimal conditions, specially in terms of the preferred curvature ν, these structures can strongly compete with the formation of T-number structures.

7.4.5 Comparison with the stability phase diagram

In Section 6.3 of the previous chapter, we used Monte Carlo simulations in order to study the stability and the range of α and ν parameters than keep the integrity of preformed shells made by one capsomer type. Now, we can compare this predicted range of stability with the optimal conditions for assembly found in the BD simulations. Fig. 7.2 compares the full stability diagram obtained in the MC simulations with the assembly phase diagram at T=0.09. The Monte Carlo stability phase diagram was made at T=0.05, a temperature much lower that the

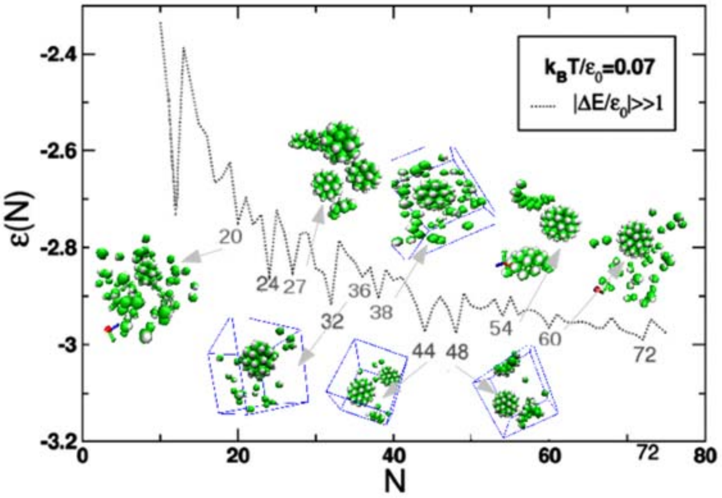


Figure 7.6: Regular non-CK structures obtained in the assembly simulations. The figure represents the free energy landscape of the optimal energy per capsomer ϵ as a function of the total number of capsomers N obtained by MC simulations for the model of Ref. [9] at T=0.07 using one type of capsomer. The inset figures are snapshots of our BD assembly simulations showing the correspondence between the non-CK structures found in the simulations and the deepest minima in the free energy landscape.

one used for the assembly $T \in [0.09, 0.12]$. We analyzed how temperature modifies the results of the MC simulations. Essentially, only the tolerance in the parameter α is altered upon altering the temperature. In particular, for a $T = 1$ shell, the range of stability is $\alpha \in [0.05, 1.80]$ at a temperature T=0.05 and becomes $\alpha \in [0.05, 1.15]$ at a temperature T=0.09. The changes in the stability range in terms of the preferred angle ν are not significant. Thus, in terms of stability the main effect of temperature is to reduce the range of values of α, but in a region that has not been explored in the assembly simulations. Hence, we can still make a sensible comparison between the stability phase diagram (Fig. 6.16) and assembly phase diagrams (Fig. 7.2, 7.4 and 7.5), in the reduced range of small α explored.

The first noticeable difference between the stability and the assembly diagrams is the absence of $T = 4$ structures. As expected, even though an artificially preformed $T = 4$ is locally stable, it is not a structure that can be self-assembled directly from a solution. A second remarkable difference is that the stability range of $T = 7$ is much larger that the optimal conditions for its assembly. In particular $T = 7$ structures have a strong competence with structures with similar or larger number of capsomers, since for them, the energetic gain of incorporating more subunits compensates the cost that has to be paid in terms of bending energy. This finding is also in line with the fact that the optimal energies per capsomer for large N become quite similar, indicating that a very monodisperse distribution of large shells will require the help of other auxiliary mechanisms to discriminate between structures with similar curvatures. The biological implication of this result is that large T-number viruses might require additional help, often in the form of scaffolding proteins, to select the right size and curvature. That could be the reason why many $T = 7$ viruses as P22 [10] use scaffolding proteins in their assembly.

Another important difference is that, in the assembly, we did not obtain elongated structures. That could be due to the fact that they are expect to appear in a very narrow and restricted set of conditions, that could not be capture in our coarse sweep of parameters. Another reasonable possibility is that energetically more optimal spherical structures with a different number of capsomers may be more favorable than elongated structures, when they have to be assembled from scratch. That also suggest that some sort of template or additional ingredient (e.g. the stiffness of the genetic material) may be needed for the formation of elongated structures.

In terms of stability ranges, it is clear that consistently the structures assemble more efficiently at values of spontaneous curvature ν smaller than the theoretically optimal value for the final shell ν_{opt}. As mentioned before, this may be a way to avoid the closure catastrophe described in Ref. [7]. In terms of α, assembly occurs in a much restricted range of α's than the conditions required to keep the integrity.

A final difference is that the regions in the assembly phase diagrams are not as well defined as in the stability diagram. Assembly is a competitive, stochastic and non-equilibrium phenomena where many things can go wrong preventing a successful outcome, specially at the strongly supersaturated conditions where our simulations were done. In addition, many repetitions at a given set of conditions have to be done to improve the accuracy of the results.

7.5 Kinetics of viral assembly

In this section, we will analyze in more detail the kinetics of viral assembly using BD simulations. In particular, we will first try to determine the equilibrium conditions for assembly. Then, we will evaluate the most important quantities in the assembly, specifically the critical cluster size, the assembly rate, and the height of the energy barrier towards assembly. All these characteristics depend not only on the conditions in terms of density, temperature and values of α and ν, but also on the target structure. For the sake of simplicity and concreteness, we will focus on the particular case of the kinetics of assembly of a $T = 3$ shell at the near optimal values of $\alpha = 0.1$ and $\nu = 1.90$.

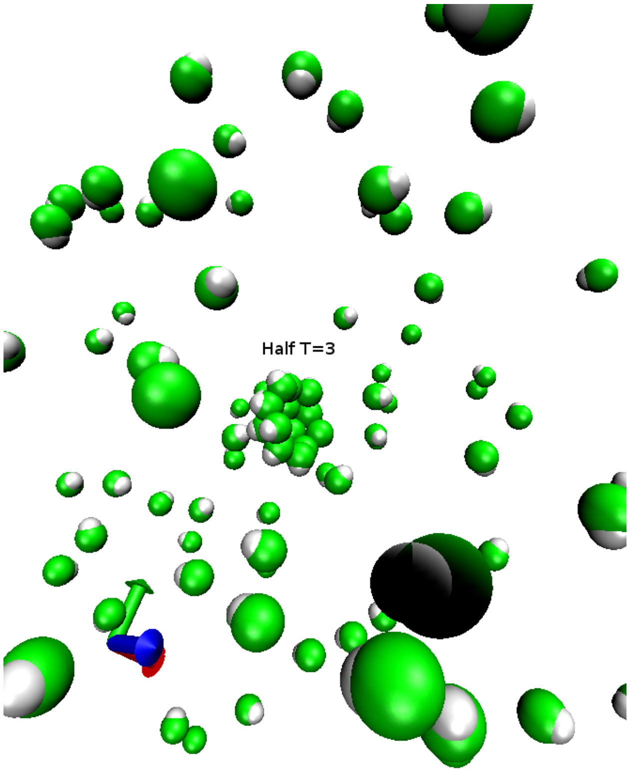


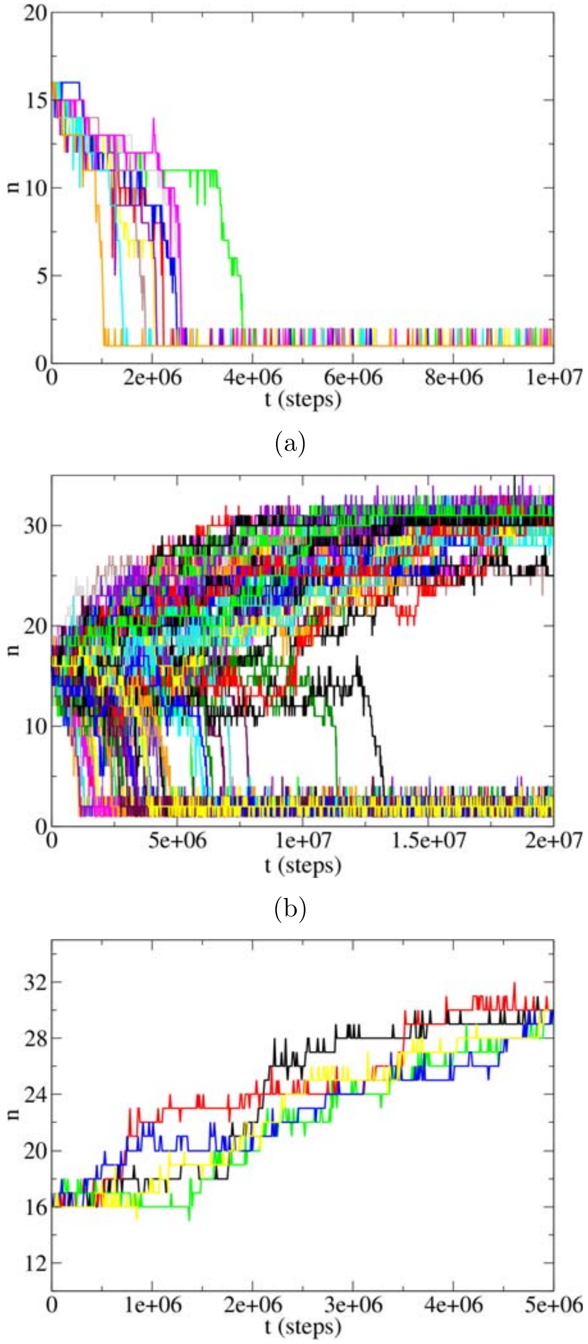
Figure 7.7: Snapshot of the initial configuration in the BD simulations to determine the equilibrium conditions. A half $T = 3$ shell with 16 capsomers is placed at the center of the simulation box and surrounded by 100-16 subunits placed at random positions and orientations.

7.5.1 Equilibrium conditions for the assembly

The density of subunits is a very important factor for the assembly. As described in Chapter 2, assembly experiments *in vitro* are characterized by a steep dependence on the initial total protein concentration. There is a threshold concentration below which no capsids are formed and there is also another special concentration corresponding to equilibrium conditions, i.e. the situation where a capsomer in solution has the same chemical potential as being part of a complete capsid. CNT provides a nice, clear, and simple way to characterize this equilibrium in simulations. As shown in Chapter 2, equilibrium conditions correspond to $\Delta\mu = 0$. At these conditions, the critical cluster size is half the size of the complete capsid, i.e. 16 capsomers for a $T = 3$ shell with $N = 32$, and being at unstable equilibrium at the top of the barrier, it has a 50% chance of growing or decaying.

Accordingly, we developed a simulation where initially, half a perfect $T = 3$ was placed in the simulation box surrounded by $100 - 16$ particles placed and oriented randomly (see Fig. 7.7). These simulations were run at a reduced temperature T=0.1 for 210^7 time steps where the size of the cluster was monitored.

100 different realizations of the simulations were performed at each density and the final size of the cluster after $2 * 10^7$ times steps was recorded. The goal was to ascertain whether the half capsid grows or disassembles. As shown in Fig. 7.8(a), at very low densities, the partial capsid always lost subunits at the end of the simulation. On the contrary, at high densities (Fig. 7.8(c)) the size of the partial shell is always larger at the end of the run. The equilibrium density corresponds to the situation where there is roughly a 50% chance of decaying or growing. This density was found to be approximately $c^* = 0.0085 \pm 0.0002$, as shown in Fig. 7.8(c).



(a)

(b)

(c)

Figure 7.8: Size of the largest cluster in the simulation (initially half a $T = 3$ shell with 16 capsomers) as a function of time for (a) $\rho = 0.0001$ and 10 different repetitions, where the size of the cluster always decays (b) $\rho = 0.0085$, corresponding to the equilibrium conditions, and (c) at $\rho = 0.04$, where the shell always grows.

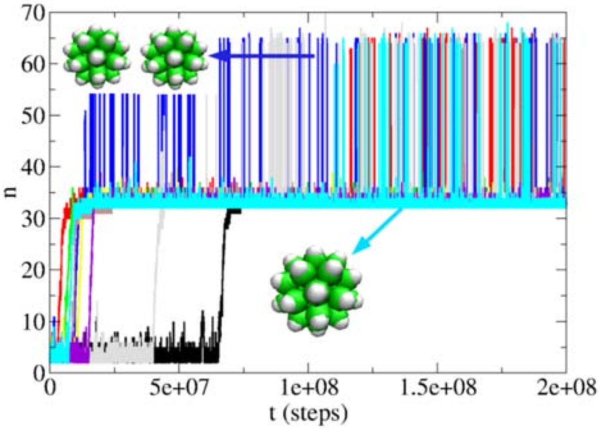


Figure 7.9: Size of the largest cluster as a function of time in five different realizations of an assembly simulation at T=0.1, and $\rho = 0.0085$.

7.5.2 Kinetics of assembly of a $T = 3$ shell

Once determined the equilibrium conditions, we analyzed the kinetics of capsid formation for different values of the initial density of subunits. All simulations were performed with $N = 100$ subunits at a reduced temperature T=0.1. The simulation starts with all subunits located at random positions and orientations, and the size of the largest cluster was monitored for the whole duration of the simulation (typically up to $2 * 10^9$ steps). Fig. 7.9 shows the size of the largest cluster as a function of time for five different realizations of the simulations at $\rho = 0.04$. One can see that the formation of a $T = 3$ shell, characterized by 32 subunits, occurs at different times due to the stochastic and activated nature of the process. A noticeable time lag can be appreciated before the formation of the first shell. In some of the simulations, two $T = 3$ shells or sometimes three $T = 3$ shells get formed. Another remarkable observation is the fact that the addition of the last one or two subunits requires often a significant span of time to get a complete capsid. This slower kinetics of the addition of the last subunits has been reported in other simulations [11, 12] and is due to the several attempts required for a subunit to find the small hole in an almost complete capsids.

Fig. 7.10 shows the size of the largest cluster in the simulation at five different densities. At high densities, the formation of the shell occurs almost immediately, whereas at low densities long delays where no large cluster are formed, are observed before the successful formation of the $T = 3$ shell.

To get accurate evaluation of the relevant kinetics of assembly, for each system and set of conditions we performed 100 independent realizations, using different initial conditions.

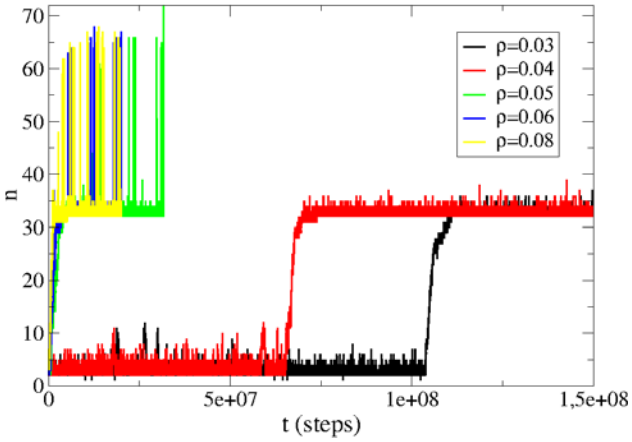


Figure 7.10: Size of the largest cluster as a function of time in an assembly simulation at T=0.1, and different densities: $\rho = 0.03$ (black line), $\rho = 0.04$ (red line), $\rho = 0.05$ (green line), $\rho = 0.06$ (blue line), and $\rho = 0.08$ (yellow line).

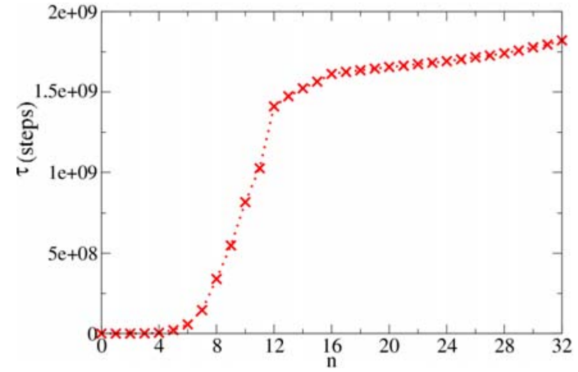
We obtained the rate of capsid formation (i.e. the number of capsids formed per unit volume and unit time) with high accuracy using the method of mean first-passage times (MFPT) [13]. We first calculate the MFPT $\tau_{max}(n)$, defined here as the average time at which the largest cluster in the simulation reaches or exceeds the size n for the first time. From it we can immediately infer the values of the critical cluster size n^*, the Zeldovich factor Z and the nucleation rate $J = (\tau_J V)^{-1}$, by fitting the simulation data to the function

$$\tau_{max}(n) = \frac{\tau_J}{2}\left[1 + \operatorname{erf}\left(Z\sqrt{\pi}\,(n - n^*)\right)\right]. \tag{7.1}$$

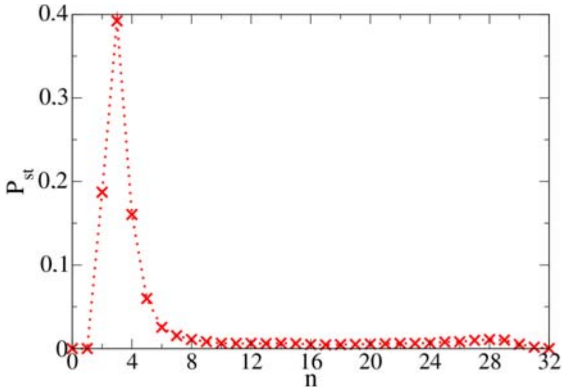
In addition we also reconstructed the relevant free energy barrier directly from the BD simulations using the kinetic reconstruction method that only requires knowledge of the MFPT and the steady-state probability distribution of clusters in the system, as described in Ref. [14]. The steady-state probability distribution $P_{\mathrm{st}}^{\mathrm{max}}(n)$ is obtained easily by sampling the size n of the largest cluster in the system in every repetition up to the first time that a complete $T = 3$ shell is formed, and making a histogram, normalized by the total number of sampled cluster sizes. Using these two inputs, the free-energy landscape is reconstructed from the equation

$$\beta\Delta G_{\mathrm{max}}(n) = \beta\Delta G_{\mathrm{max}}(n_1) + \ln\left(\frac{B(n)}{B(n_1)}\right) - \int_{n_1}^{n_{max}} \frac{dn'}{B(n')} \tag{7.2}$$

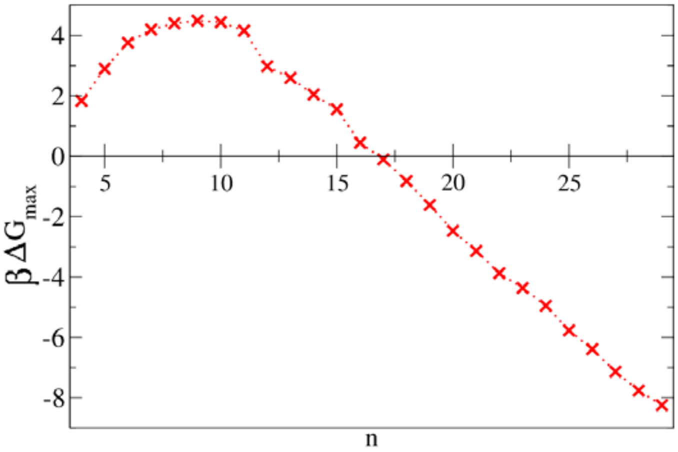
where $\beta = 1/k_{\mathrm{B}}T$ and



(a)



(b)



(c)

Figure 7.11: (a) Mean first-passage times τ(n) as a function of the cluster size for $\rho = 0.04$. (b) Steady-state probability distribution of the cluster size obtained for the same conditions from the numerical solutions. (c) Free-energy landscape reconstructed using eqs. 7.3 and 7.2.

$$B(n) = -\frac{1}{P_{\mathrm{st}}^{\mathrm{max}}(n)} \left[\int_{n_{max}}^{b} P_{\mathrm{st}}^{\mathrm{max}}(n')dn' - \frac{\tau(b) - \tau(n)}{\tau(b)} \right], \tag{7.3}$$

being $b = 32$ the absorbing boundary size up to which we sample both $P_{\mathrm{st}}^{max}(n)$ and $\tau_{max}(n)$. The integrals in Eqs. 7.3 and 7.2 are evaluated by standard discretization techniques (e.g. Ref. [15]). Figure 7.11 (b) and (c) show the steady-state probability and the reconstructed free-energy landscape for the same system for which the MFPT are shown in (a).

The quantity of interest in nucleation theory as well as in the remainder of this work is not the barrier of the largest cluster, $\Delta G^*_{\mathrm{max}}(n)$ but rather the one posed to any individual cluster in the system, ΔG^*. But provided that the barrier for the largest has not yet vanished, i.e., $\Delta G_{\mathrm{max}}(n) > 0$, both quantities are related by the simple relation [15]

$$\beta \Delta G^* = \beta \Delta G^*_{\mathrm{max}} + \ln N. \tag{7.4}$$

Using these MFPT techniques, we have calculated the rate of capsid assembly, the critical cluster size, and the nucleation barrier in a range of densities $\rho \in [0.02, 0.1]$. The results are plotted in Fig. 7.12.

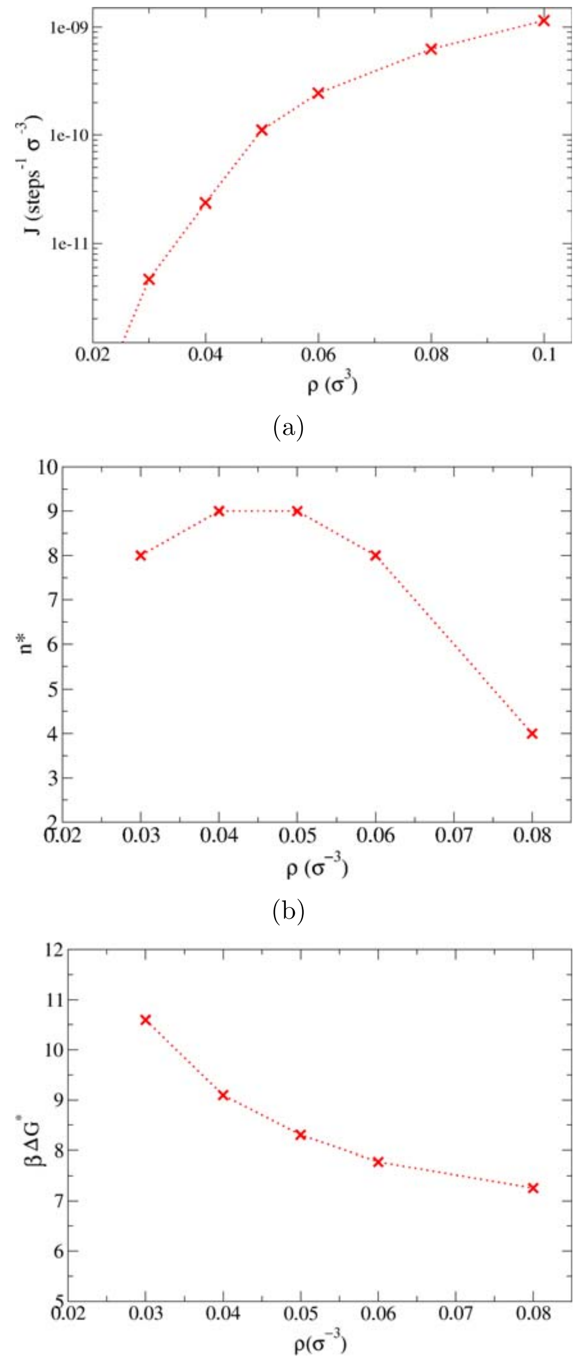
We can clearly see that, as the density is reduced, both the size of the critical cluster and the nucleation barrier increases, leading to an exponential reduction of the nucleation rate. This is in qualitative agreement with the predictions of CNT formulated in Chapter 2. A more careful comparison with CNT is performed in the next subsection.

7.5.3 Comparison with CNT

The results of the previous subsection clearly indicate the nucleation nature of the self-assembly process. A large enough critical size has to be reached and an energy barrier has to be overcome to succeed in the formation of a complete capsid. According to CNT, the height of the nucleation barrier, the size of the critical cluster and the nucleation rate are given by (see Chapter 2)

$$\begin{aligned} \triangle G^* &= \frac{q}{2}\left(\sqrt{\triangle\mu^2 + a^2} + \triangle\mu\right) \\ n^* &= \frac{q}{2}\left(1 + \frac{\triangle\mu}{\sqrt{\triangle\mu^2 + a^2}}\right) \\ J(t) &= c_s\beta^* Z e^{-\Delta G^*/k_B T} \end{aligned} \tag{7.5}$$

where $\Delta\mu = -k_B Tln(c_1/c^*)$ and a was related to the line tension for a capsid. As described in ref. [7], a good approximation for this parameter for a $T = 3$ shell was $a = \frac{4\pi}{3\sqrt{3}}\frac{|\mu_e|}{q}\frac{R}{\sigma_0}$, where $|\mu_e| = 2.23$ is the energy per capsomer in a complete capsid, $q = 32$ is the number of capsomers in a complete $T = 3$ shell and $R = 2.67\sigma_0$ is its radius. Using these approximations yielding $a = 0.45$, and the value of the equilibrium density $c^* = 0.0085$ obtained from the simulations, we can compare the predictions of CNT with the results of the kinetic simulations. Fig. 7.12 shows this comparison for the critical cluster size and the nucleation barrier. The predictions for the critical cluster size and the nucleation barrier are approximately $n^* \approx 10$ and $\Delta G^* \approx 7$. Both predictions seem to be only roughly qualitatively accurate. A more careful and accurate comparison, including also the predictions for the nucleation rate, will be performed in the future.



(a)

(b)

(c)

Figure 7.12: a) Rate of capsid formation J b) critical size n^* and c) nucleation barrier height $\beta\Delta G^*$ as a function of the density of the system, obtained from the BD simulations.

7.6 Conclusions

In this chapter we have analyzed the kinetic aspects of the viral self-assembly, using our coarse grained model and BD simulations. From these simulations, we have obtained a lot of interesting and useful general information about this important process.

In particular, we have verified the importance of the initial density of subunits in the assembly and characterized the equilibrium conditions for a $T = 3$ capsid. We then analyzed how the outcome of viral assembly depends on the relevant parameters of the interaction: the effective binding energy (or, equivalently, the reduced temperature), the preferred angle ν and the local bending stiffness α. From extensive simulations, we were able to construct the phase diagram of the assembly, which turned out to be quite different from the stability phase diagram discussed in the previous chapter. Specifically, the assembly takes place preferentially at $\nu < \nu_{opt}$, presumably to avoid the closure catastrophe. The assembly range is much smaller than the stability range, specially at high temperatures and for the largest shell $T = 7$, that seem to require very restricted conditions to assemble. In addition, no elongated or $T = 4$ structures are formed. In contrast, aberrant and locally optimal non-CK structures are found and strongly compete in some regions of parameters with the CK structures.

We have also analyzed and characterized accurately the kinetics of assembly of $T = 3$ shells. All features of the assembly point out to nucleation as the physical mechanisms controlling shell formation. Using a powerful technique based on the concept of MFPT, we have determined the critical cluster sizes, the height of the nucleation barrier and the rate of capsid formation directly from our assembly simulations, at different conditions. The results agree at least qualitatively with the predictions of CNT and open the door to a proper understanding and control of viral assembly.

The same analysis can be repeated for the different T-number shells, gaining an important insight on the competitive assembly of different structures. In addition, the ideas developed in this chapter can be used for the analysis of *in vitro* experiments, bringing promise to a more accurate characterization and prediction of the assembly of real viruses.

References

[1] S. Mukherjee, MV. Thorsteinsson, LB. Johnston, PA. DePhillips, and A. Zlotnick, A quantitative description of in vitro assembly of human papillomavirus 16 virus-like particles, *J. Mol. Biol.*, **381**, 229-237, 2008.

[2] GL. Casini, D. Graham, D. Heine, R. L. Garcea, and D. T. Wu, In vitro papillomavirus capsid assembly analyzed by light scattering, *Virology*, **325**, 320-327, 2004.

[3] DM. Salunke, DLD. Caspar, and RL. Garcea, Polymorphism in the assembly of polyomavirus capsid protein VP1, *Biophys. J.*, **56**, 887-900, 1989.

[4] S. Kanesashi, K. Ishizu, M. Kawano, S. Han, S. Tomita, H. Watanabe, K. Kataoka, and H. Handa, Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro, *J. Gen. Virol.*, **84**, 1899-1905, 2003.

[5] MS. Kumar and R. Schwartz, A parameter estimation technique for stochastic self-assembly systems and its application to human papillomavirus self-assembly, *Phys. Biol.*, **7**, 045005, 2010.

[6] P. Ceres and A. Zlotnick, Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids, *Biochemistry*, **41**, 11525-11531, 2002.

[7] A. Luque, D. Reguera and A. Morozov *et al.*, Physics of shell assembly: Line tension, hole implosion, and closure catastrophe, *J. Chem. Phys.*, **136**, 18, 2012.

[8] R. Zandi, D. Reguera, RF. Bruinsma, WM. Gelbart and J. Rudnick, Origin of icosahedral symmetry in viruses, *PNAS*, **101**, 44, 15556-15560, 2004.

[9] A. Luque, Thesis. Structure, Mechanical Properties, and Self-Assembly of Viral Capsids, 2011.

[10] KN. Parent, A. Zlotnick, and CM. Teschke, Quantitative analysis of multi-component spherical virus assembly: scaffolding protein contributes to the global stability of phage P22 procapsids, *J. Mol. Biol.*, **359**, 1097-106, 2006.

[11] DC. Rapaport, Role of Reversibility in Viral Capsid Growth: A Paradigm for Self-Assembly, *Phys. Rev. Lett.*, **101**, 186101, 2008.

[12] MF. Hagan and D. Chandler, Dynamic Pathways for Viral Capsid Assembly, *Biophys. J.*, **91**, 1, 42-54, 2006.

[13] J. Wedekind, R. Strey, and D. Reguera, New method to analyze simulations of activated processes, *J. Chem. Phys.*, **126**, 134103, 2007.

[14] J. Wedekind and D. Reguera, Kinetic reconstruction of the free-energy landscape, *J. Phys. Chem. B*, **112**, 11060, 2008.

[15] BP. Flannery, WH. Press, SA. Teukolsky and W. Vetterling, Numerical recipes in C. Press Syndicate of the University of Cambridge, New York, 1992.

[16] J. Wedekind, G. Chkonia, J. Wölk, R. Strey and D. Reguera, Crossover from nucleation to spinodal decomposition in a condensing vapor, *J. Chem. Phys.*, **131**, 114506, 2009.

Part III

Mechanical properties of viral capsids

CHAPTER 8

The importance of the structure in the buckling and maturation of spherical viruses

8.1 Introduction

One of the main roles of the viral capsid is to protect the genetic material against external aggressions and to ensure its safe release at the right location in the host cell. Thus, the mechanical properties of viral capsids play an important role in their stability and infective process. In particular, from the production of a virus to the invasion of a new host, the viral capsid must be stable through environments with different pH or salt concentrations. In addition, the capsid of many viruses that penetrate the host membrane through endocytosis is subjected to a substantial mechanical stress [1]. Moreover, double stranded DNA (dsDNA) viruses should withstand up to tens of atmospheres arising from the confined genetic material [2, 3]. For all these reasons many viruses in the final stage of assembly undergo a maturation step that properly tunes the properties of the initial shell, converting an innocuous procapsid into an infectious agent. This maturation process involves structural changes that often lead to a noticeable change in the shape of the capsid from an initial spherical capsid with icosahedral symmetry into a polyhedral shell with flat triangular faces and the shape of an icosahedron. This transformation from a spherical to a polyhedral shape of the viral capsid is known as the buckling transition.

The transformation of the capsid taking place during virus maturation is a complex process that can be associated to different biological pathways, which may include cleavage, conformational transitions and even covalent bonding of the capsid proteins. For instance, bacteriophage T7 assembles first a procapsid helped by scaffolding proteins, and then undergoes an expansion and a buckling transition during the packaging of the viral dsDNA [4, 5]. Instead, bacteriophage HK97 undergoes a cleavage in the coat protein of the prohead that leads to a buckling transition that can be triggered *in vitro* by a change in the pH [6–8]. Once triggered, the buckling transition seems to be related to an energy cascade of steps that lowers the energy of the capsid [7, 8].

To understand separately the influence of the different ingredients that could be relevant to and responsible of the buckling transition, we will restrict our analysis to the simplest case where the transition does not require auxiliary proteins, cleavage of the coat protein, nor cross-linking between subunits.

The buckling transition originates from the competition between the bending and the stretching energy, and can be described in physical terms by using continuum elasticity theory [10–12]. As mentioned in Chapter 4, in this framework, the viral capsid is considered a thin elastic shell with two competing energies: The bending energy, which takes into account the deviations from the spontaneous curvature imposed by the proteins, and the stretching energy, which involves the in-plane deformations of the network of proteins. The ratio of these contributions defines the Föppl-von Kárman number, $\gamma = YR_0^2/\kappa_b$ [11, 12], a nondimensional

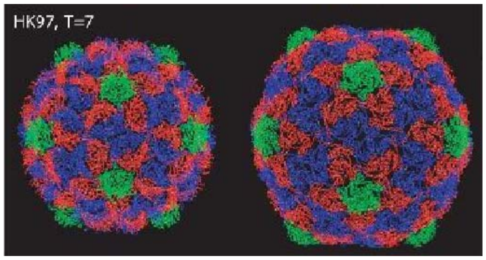


Figure 8.1: Buckling transition during maturation of Hong Kong 97. Figure taken from Ref. [9].

parameter that determines the shape of the virus. Here Y is the 2D Young modulus, κ_b is the bending rigidity (both intrinsic properties of the shell) and R_0 is the radius of the spherical shell. For small values of γ the spherical shape is energetically favored. But when exceeds a certain threshold, a buckling transition takes place and the polyhedral shape, with flat faces, becomes more stable because it reduces the stretching energy. Since $\gamma \approx R^2$, big icosahedral capsids are expected to be polyhedral rather than spherical, which is in agreement with experimental observations [14]. Nevertheless, the discrete nature and arrangement of capsomers in the viral shell might have a significant influence in the buckling phenomenon, specially for small viruses, which cannot be easily described in the framework of continuum elasticity theory.

In this chapter we will study the influence of the T-number on the capsid shape, on its tendency to buckle, and on the resulting mechanical properties of the viral shell. To that end, we will first describe (in sections 8.2 and 8.3) the simple coarse-grained model and the mechanical characterization used in this analysis. We will then compare in section 8.4 the energetics and mechanics of spherical and icosahedral viral capsids for different triangulation numbers T and determine how sensitive is the buckling transition to the discrete nature and precise arrangement of protein subunits in the shell. We will also analyze the tendency to buckle in terms of the distribution of lateral stresses in section 8.5. Section 8.6 we will be devoted to study the morphological and energetic changes that occurs upon expansion. The mechanical properties of the resulting capsid will be then characterized in terms of the bulk modulus (section 8.7) and the maximum sustainable pressure (section 8.8). Finally, we will discuss the potential mechanical and biological advantages of polyhedral versus spherical capsids and compare our main conclusions with data from real viruses.

We will focus on the simplest case where the integrity and the strength of the effective interactions among coat proteins are not changed, since our main goal is to isolate the pure relevance of their geometric arrangement on the buckling transition. In addition, we will study this problem by comparing the properties of the (initial) spherical shape against the (final) icosahedral shape. Our simplified model is thus not aimed to study how the real dynamic transition between these two states takes place. We will see that, despite its simplicity and limitations, our model is still able to provide useful insight on this very complex problem that seem to be confirmed in some cases by experimental evidences on real viruses that comply with the aforementioned restrictions.

8.2 Minimal coarse-grained model of spherical and polyhedral shells

To study the buckling transition and the role of the arrangement of capsomers on viral maturation, we have used the minimal model for capsid architecture described in Section 3.2.1, comparing the energy and mechanical properties of spherical and polyhedral shapes for the lowest T-numbers and P-classes listed in Table 1.6.

In the model, pentamers and hexamers are represented by spheres of different sizes with isotropic pairwise interactions described by the potential:

$$V_{LJ}(r_{ij}) = \varepsilon_0 \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{8.1}$$

where ε_0 is the binding energy between capsomers, σ_{ij} is the equilibrium distance, and r_{ij} is the separation between capsomer i and j. We take ε_0 to be $15k_BT$ [15], with k_B being the Boltzmann constant and T the absolute temperature. Hexamers and pentamers are modeled as spheres with radii σ_0^h and σ_0^p, correspondingly. Their ratio $\sigma_0^h/\sigma_0^p = tan(\pi/6)/tan(\pi/5) \simeq 0.8$ is the same as that between the radii of circles inscribed in a pentagon and a hexagon of the same edge, as follows from the fact that pentamers and hexamers are made up of the same protein. The equilibrium distances for hexamer-hexamer, pentamer-pentamer, and hexamer-pentamer are $\sigma_{h-h} = 2\sigma_0^h$, $\sigma_{p-p} = 2\sigma_0^p$ and $\sigma_{p-h} = \sigma_0^p + \sigma_0^h$, respectively. We have used the same interaction energy between capsomers. We note that all results in this chapter are expressed in units of $\sigma_0 \equiv \sigma_0^h$ and ϵ_0. It is important to emphasize that we will take capsid size and binding energy ϵ_0 as constant and equal for all structures. Accordingly, buckling induced by alterations in protein size or interactions (caused, for instance, by cleavage) are beyond the scope of the present model. The Lennard-Jones-like potential is not truncated and interactions are allowed between all capsomers. Moreover, capsomer positions are restricted to lie on the surface of the corresponding sphere or icosahedron. Without this constraint, this type of potential will produce compact clusters rather than shells as optimal configurations. In addition, we will only compare the (static) properties of purely spherical versus strictly polyhedral capsids. We will neither analyze a continuum change of shape between these two states nor the dynamics of this process.

With these approximations the total energy of the shell will be the sum over all pairs of capsomers: $E = \sum_{i=1}^{N} \sum_{j>i} V_{LJ}$. We will analyze the problem of buckling by comparing the properties of the (initial) spherical shape against the (final) polyhedral shape.

To build spherical shells we use as starting point the exact coordinates of the icosahedral spherical code corresponding to the solution of the maximum volume problem from [16]. Those sets of coordinates correspond to energy minima of the Lennard-Jones interaction model described above as it has been shown by Monte Carlo simulations [17, 18]. We have also verified using a conjugated gradient minimization algorithm that these configurations at the optimal radii are indeed in mechanical equilibrium. Using these fixed angular coordinates, we have explored the energies and mechanical properties of spherical capsids for different radii R.

To form the polyhedral shells, an icosahedron is reconstructed using the basic triangles defined by the Caspar and Klug construction [19]. Starting from the h and k values of a given T-number, we build the associated equilateral triangle on a hexagonal lattice and keep track of the x, y coordinates of all capsomers, i.e. nodes of the hexagonal lattice, that lie inside this triangle of edge $l = \sqrt{T}$. By applying a rotation of $arctan(\frac{1}{\sqrt{3}}\frac{k-h}{h+k})$ around the z-axis, and a translation of $(l/2, l/(2\sqrt{3}), R)$ we place the center of the equilateral triangle at $(0, 0, R)$, where $R = \frac{\sqrt{3}}{12}(3 + \sqrt{5})l$ is the distance from the center to each face of an icosahedron of edge length l. We use then 20 copies of this triangle to reconstruct the icosahedron. The angular coordinates of the center of those triangles in the final icosahedron are given by the position of the vertexes of the dodecahedron of unit edge, which is the dual structure of the icosahedron: $(\pm 1, \pm 1, \pm 1)$, $(0, \pm\frac{1}{\phi}, \pm\phi)$, $(\pm\frac{1}{\phi}, \pm\phi, 0)$, and $(\pm\phi, 0, \pm\frac{1}{\phi})$, where $\phi = (1 + \sqrt{5})/2$ is the "golden ratio". After transforming those points into spherical coordinates (r, θ, ϕ), the icosahedron is reconstructed from the basic CK triangle located at $(0, 0, R)$ by applying a custom rotation around the z-axis, followed by one of an angle θ around the y-axis and one of ϕ around the z-axis. This procedure was implemented using Mathematica©. In this way we generated the (x, y, z) coordinates of all capsomers on an icosahedron of edge $l = \sqrt{T}$ and radius $R = \frac{1}{4}\sqrt{10 + 2\sqrt{5}}l$, defined as the distance from the center to any vertex. Finally, the optimal polyhedral shell is obtained by varying the radius R to minimize the total energy, as shown in Section 8.4. Note that we keep the same icosahedral shape for all radii, but in each case the distances between capsomers are uniformly rescaled.

8.3 Mechanical characterization

One of the relevant mechanical properties of the capsid is the distribution of stresses of the capsomers. To characterize and evaluate the stress distribution we have used the stress tensor [20]. In particular, the *global* stress has been measured as the average of the viral stress tensor, $\sigma_{\alpha\beta}$, and can be decomposed in the kinetic and force contributions of each morphological unit i:

$$\Omega^d \sigma^d_{\alpha\beta} = \sum_i \Omega^d_i (\sigma^d_{\alpha\beta})_i = \sum_i \left[\sum_j \frac{1}{2} m_i v^\alpha_i v^\beta_j - \frac{1}{2} \sum_{i \neq j} \frac{dV(r_{ij})}{dr_{ij}} \frac{r^\alpha_{ij} r^\beta_{ij}}{r_{ij}} \right]. \tag{8.2}$$

Here d is the dimensionality, usually $2D$ or $3D$, and Ω^d is the *d-volume*, i.e., the surface in $2D$ or the volume of the shell in 3D. In the kinetic term (the first term), m_i and v^α_i are the mass and the α component of the velocity of capsomer i, respectively. In the force term, r_{ij} is the module of the vector that joins particles i and j, i.e., $\vec{r}_j - \vec{r}_i$, and r^α_{ij} is the α component of this vector. Finally, Ω^d_i corresponds to the d-volume per capsomer, which can be defined in different ways, all leading to similar qualitative results. Here we use the simplest choice where each capsomer has the same d-volume $\Omega^d_i = \Omega^d_{Tot}/N$. Using these definitions, we can now compute the local, $(\sigma^d_{\alpha\beta})_i$, and global, $\sigma^d_{\alpha\beta}$, mechanical stresses.

Giving the high relative strength of the interactions compared to the thermal energy, $\varepsilon_0 = 15 k_B T$, we will neglect the kinetic contribution of the stress tensor in Eq.(8.2). In a first approximation the interactions between capsomers determine the main properties of the system, whereas thermal fluctuations would be relevant only in more accurate studies, e.g., the role of soft modes in the activation of buckling [13, 21, 22]. For spherical shells we express the stress tensor, Eq.(8.2), in spherical coordinates, and due to the symmetry of viral capsids it is worth to recombine its components in two terms [20]: the lateral stress

$$(\sigma^{2D}_T)_i = -\frac{N}{2A} \sum_j \frac{dV(r_{ij})/dr_{ij}}{r_{ij}} \frac{(\vec{r}_{ij} \cdot \widehat{e_\theta})^2 + (\vec{r}_{ij} \cdot \widehat{e_\phi})^2}{2} \tag{8.3}$$

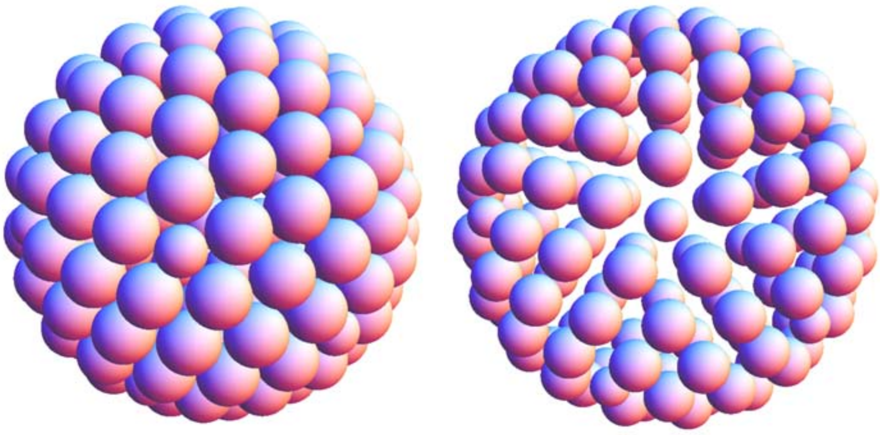
and the (45^o) shear stress

$$(\sigma^{2D}_{\theta\phi})_i = -\frac{N}{2A} \sum_j \frac{dV(r_{ij})/dr_{ij}}{r_{ij}} (\vec{r}_{ij} \cdot \widehat{e_\theta})(\vec{r}_{ij} \cdot \widehat{e_\phi}) \tag{8.4}$$

where A and V are the surface and volume of the shell.

For icosahedral shells the formulas are still valid but it is better to use a different set of basis vectors. In particular, for each face we replace ($\widehat{e_\theta}$,$\widehat{e_\phi}$) by two perpendicular vectors that are tangent to the surface of each equilateral triangle, and $\widehat{e_r}$ for a vector perpendicular to the face at all points. For both capsid shapes, Eq. (8.3) and (8.4) characterize the local stresses, and using them into Eq.(8.2) we obtain similar expressions for the global mechanics of capsids.

Finally, we will analyze the pressure p and the bulk modulus $K = -V\partial p/\partial V$, which measures the compressibility of the capsid. It is useful to rewrite both in



(a) (b)

Figure 8.2: (a) Optimal structure, corresponding to the minimum energy, obtained for R_0^{sph}, for a spherical $T = 13$ capsid. (b) Structure corresponding to $1.3R_0^{sph}$, where the energy is significantly larger than the optimal.

terms of energy variations as:

$$p = -\frac{\partial E}{\partial V} = -\frac{R}{3V}\frac{\partial E}{\partial R} \tag{8.5}$$

and

$$K = -V\left(\frac{\partial p}{\partial V}\right) = \frac{R^2}{9V}\left(\frac{\partial^2 E}{\partial R^2}\right) \tag{8.6}$$

where we have transformed volume derivatives into radial derivatives, using the fact that $V \sim R^3$. For all optimal structures the pressure is zero [20] and their compressibility will be labeled as K_0.

8.4 Energy of spherical versus polyhedral capsids

Given the constraint that the capsomers have to remain on the surface of a sphere or an icosahedron with perfect icosahedral order, the total energy only depends on the radius. Thus, the first step was to find the optimal radii, R_0^{sph} and R_0^{ico}, for the spherical and polyhedral capsids with T-numbers listed in Table 1.6. To that end, starting from the exact angular coordinates of the spherical and icosahedral shells described in section 8.2, we subjected the capsids to a radial deformation and calculated the associated energy, as shown in figures 8.2 and 8.3 for the $T = 13$ case.

For each shape and T -number we obtained a nearly parabolic behavior around the energy minima. Fig. 8.4 plots the optimal radii of the spherical and polyhedral shells as a function of the T-number. In both cases the radii increase as $R_0 \approx \sqrt{T}$,

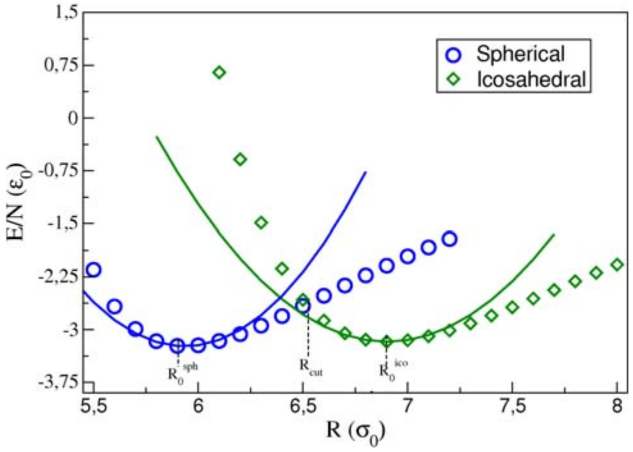


Figure 8.3: Energy per capsomer for the spherical and icosahedral $T = 13$ shells under radial deformation. The optimal radii of the sphere, R_0^{sph}, the icosahedron, R_0^{ico}, and the cut radius where both energies intersect, R_c, are also indicated. The lines show a parabolic fit $E = E_0 + \frac{1}{2}a_E(R - R_0)^2$ around the respective minima.

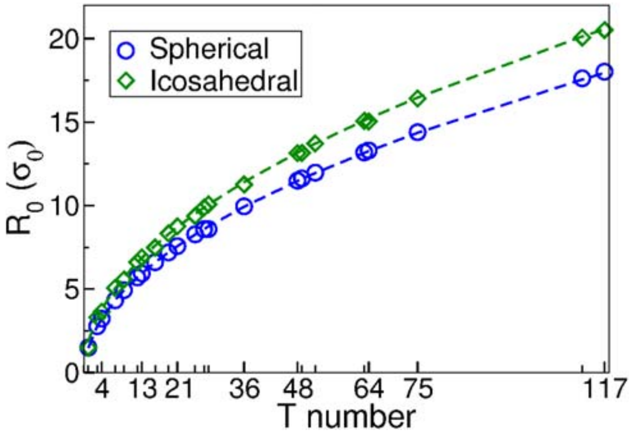
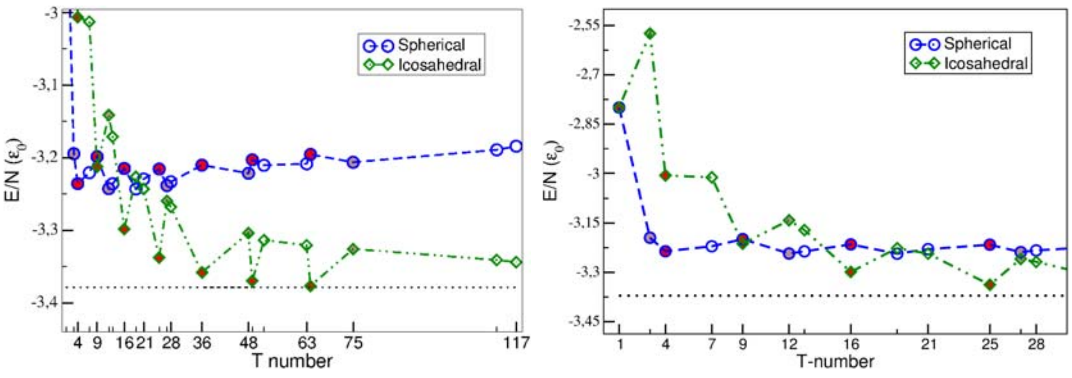


Figure 8.4: Spherical and icosahedral optimal radii as a function of the T-number. The lines are the theoretical predictions of the surface-covering model (see Appendix A), and the points are obtained by minimizing the energy of the different structures.

since the T-number is proportional to the shell surface which grows as R^2. Comparing the optimal radii, we systematically observe that the polyhedral shells are larger than the spherical capsids typically by $\approx 15\%$, when both are made out of capsomers of equal size and strenght of interactions. This is an expected result, since for a fixed radius a sphere has a bigger surface than an icosahedron, so the latter requires a larger radius to accommodate the same number of capsomers. It is remarkable that both optimal radii can be accurately estimated using a simple surface covering model described in Appendix A. This is a direct consequence of the prevalence of steric effects between capsomers, so the results obtained for the optimal radii are essentially independent of the details of the interaction model.



(a) (b)

Figure 8.5: Optimal energy per capsomer for the spherical (blue circles) and icosahedral (green diamonds) T shells. The T numbers corresponding to the $P = 1$ and $P = 3$ classes are highlighted in red and brown, respectively. Zoom highlighting the behavior corresponding to the T-numbers missing in the main figure.

Fig. 8.5 plots the energy per capsomer of the optimal spherical and polyhedral structures as a function of the T-number. This figure shows three different regions. For small structures corresponding to $T \leq 4$, the energy is dominated by the pentamers rather than by the arrangement of the few existing hexamers, leading to a behavior that deviates from that of the remaining shells of a given class. For those small capsids, the energy is less for the spherical shape than for the polyhedral shape (see Fig. 8.5(b)). The second region corresponds to big capsids ($T > 28$), where spherical shells have systematically a higher energy than icosahedral ones, in agreement with the predictions of continuum elasticity theory. The energetic frustration of the spherical shape comes from the pentameric disclinations that introduce a stretching energy penalty proportional to R^2, which is not present an the flat faces of in icosahedron.

Between these two limits there is a region where for some T-numbers the spherical shape has less energy per capsomer than the icosahedral shape, while for other T-numbers the opposite behavior occurs. This structural dependence of the optimal energies on the T-numbers is difficult to justify in the framework of continuum elasticity [12]. In fact, there seems to be a well-defined optimal shape that, seems to be determined by the P-class rather than being controlled specifically by the T-number. In particular, T-structures corresponding to the class $P = 1$ are local maxima in the energy landscape of spherical shells, but local minimal in the polyhedral one. This suggests that $P = 1$ capsids are not particularly satisfied in the spherical shape and will have a strong tendency to become faceted. On the other hand, $P = 3$ capsids show the opposite behavior of $P = 1$: they are local minima in the spherical landscape and local maxima in the faceted one. Structures from classes $P = 7$ and $P = 13$ have an intermediate behavior between that observed for $P = 1$ and $P = 3$ class. The origin of these energy differences comes from the specific distribution of capsomers for each class P, and becomes more evident in the analysis of the stress distribution described in the following sections.

8.5 Local distribution of stress

To understand better the relevance of the arrangement of capsomers on the tendency of different classes to adopt a spherical or faceted shell, we computed the local distribution of stresses at the level of capsomers for the different T-shells. Our analysis is based on a previous paper [20], where the mechanical properties of spherical capsids were studied in detail using a similar model. We will focus on the distribution of local lateral stresses, using the convention that a positive or negative stress means that a capsomer is being compressed or stretched, respectively. The distribution of stresses for the different T-shells can be naturally grouped in classes that show a similar pattern. To illustrate how the precise arrangement of capsomers influences the tendency of a capsid to remain spherical or to become faceted, we have analyzed series of T-numbers corresponding to the first four P-classes (see Table 1.6).

The simplest case is $P = 1$, where each pair of neighboring pentamers are connected by a straight line of hexamers (see Fig. 8.6). This structural feature clearly dictates the local stress pattern. In the spherical shells, positive stress concentrates on the lines connecting pentamers for $T > 4$, which delimits triangular areas with stretched hexamers. The absolute values of those stresses get larger as the the T-number of the shell increases. Compared with the rest of spherical P-shells (see below), the $P = 1$ class shows the highest local stretching and compression. This stress frustration, associated to the geometrical configuration of capsomers in spherical $P = 1$ capsids, is the underlying reason why they are local maxima in the optimal energy landscape of Fig. 8.5. Moreover, the stress distribution clearly highlights the energetic advantage of adopting a polyhedral shape, since the hexamer stretching will be relieved if the triangular regions between pentamers are flattened. In fact, the distribution of stress in the polyhedral $P = 1$ capsids (second column in Fig. 8.6) shows that the local lateral stresses at the triangular faces are close to zero and get smaller as the size of the shell increases. Furthermore, compared to the other classes, icosahedral $P = 1$ capsids have the lowest stress frustration, which justifies why they are local minima in the optimal energy landscape (see Fig. 8.5).

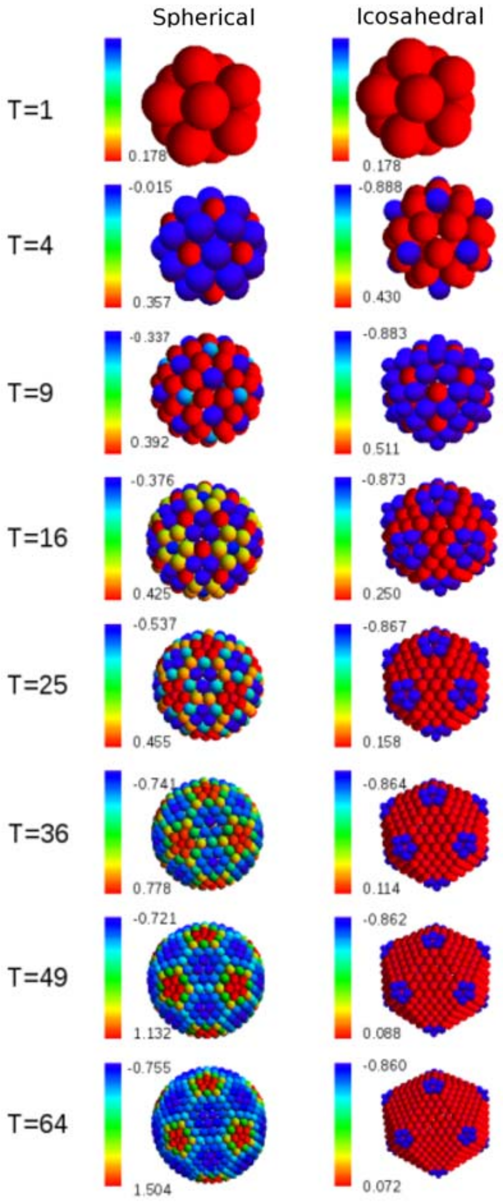


Figure 8.6: Lateral stress distribution, see Eq. (8.3), for $P = 1$ spherical and icosahedral shells. The color bar indicates the values of the local stresses, and it is rescaled for each structure using the most compressed (positive, red) and most stretched (negative, blue) capsomers.

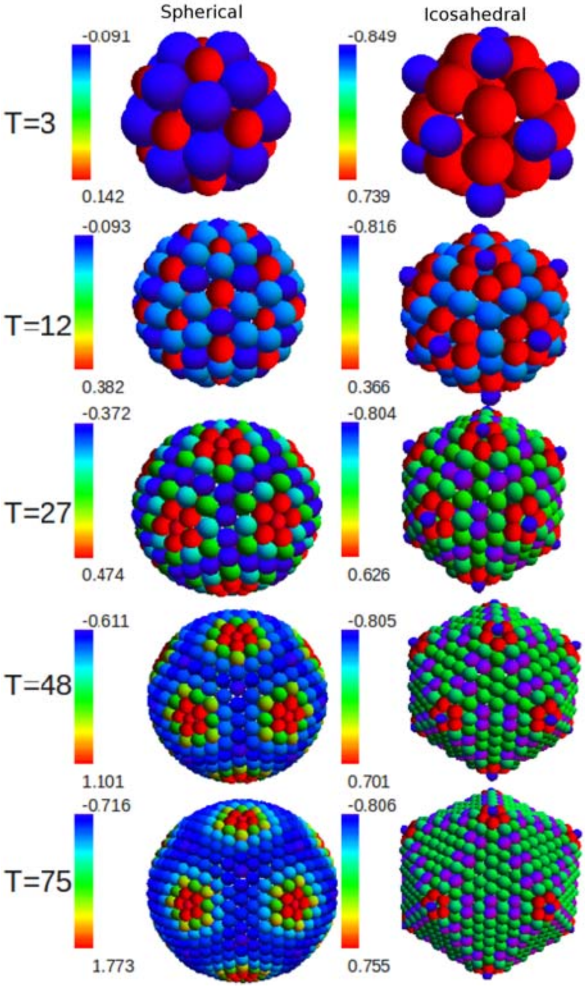


Figure 8.7: Lateral stress distribution for $P = 3$ spherical and icosahedral shells. The color code is described in Fig. 8.6.

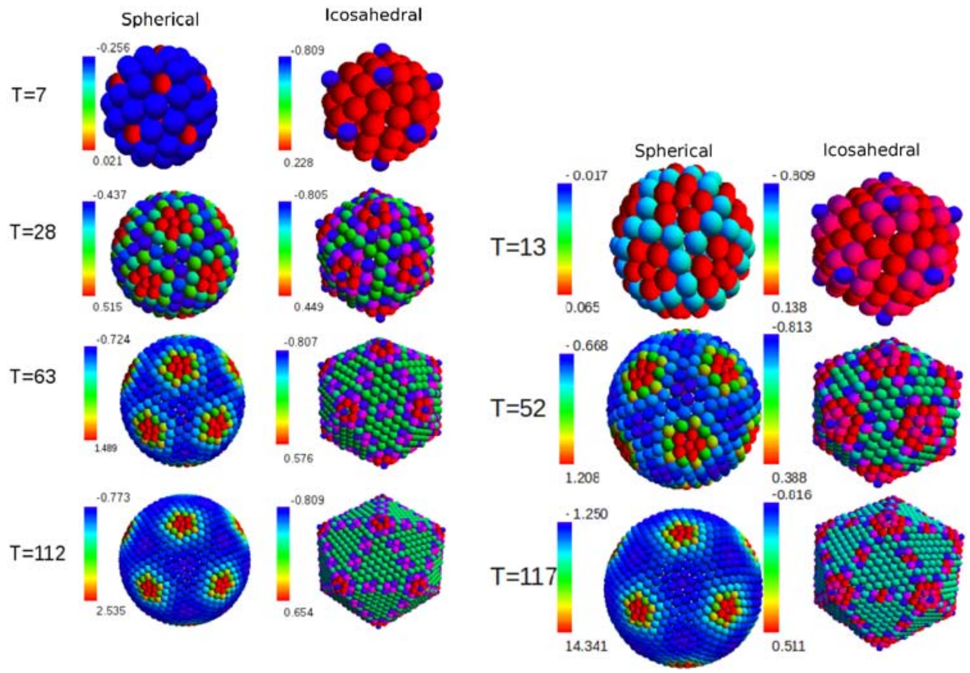
The other achiral class, $P = 3$, is characterized by zig-zag lines of hexamers connecting pentamers, instead of straight lines (see Fig. 8.7), which leads to an inversion of the stress behavior. In spherical $P = 3$ shells, compression concentrates on the pentamers, whereas the hexamers surrounding them are slightly stretched, leading to a dodecahedral pattern instead of the icosahedral one observed in $P = 1$ (see Fig. 8.7). In this case, nothing indicates any particular relief of the stress on the hexamers upon adopting an icosahedral shape. On the contrary, the stress pattern suggests that the shell will flatten along the lines connecting second neighbor pentamers, adopting a dodecahedral shape rather than the icosahedral one observed in $P = 1$ (see Fig. 8.7). It is worth to mention that such a dodecahedral faceting seems to be present in the reconstruction images of Chilo iridescent virus (CIV), a large $T = 147$ ($P = 3$, $f = 7$) virus [23]. In the spherical $P = 3$ shell, due to the symmetric capsomer arrangement, the local stress is smoothly shared among all capsomers and its values are low enough to produce local minima in the optimal energy landscape of Fig. 8.5. Contrarily, in the polyhedral shape, the $P = 3$ class has a higher stress frustration compared to the faceted $P = 1$, showing a less uniform pattern, where stress accumulates at the zig-zag lines along the edges. In fact, the icosahedron $P = 3$ construction is the most stressed of all polyhedral classes leading to local maxima in the optimal energy per capsomer (see Fig. 8.5).

The remaining P classes are chiral, i.e., for each shell we have two specular structures with different handness. The arrangement of hexamers along the line connecting neighboring pentamers is now skewed, and the resulting stress patterns show neither a clear advantage or disadvantage on adopting the shape of an icosahedron (see Fig. 8.8). Accordingly, the values of optimal energies and stresses are intermediate between those of $P = 1$ and $P = 3$. Therefore, classes with $h_0 \approx k_0$ and $h_0 \geqslant k_0$ will behave similarly to $P = 3$ and $P = 1$, respectively.

Finally, a common feature for all spherical classes is that positive stress accumulates on the pentameric disclinations as the shell gets bigger, in agreement with the predictions of continuum elasticity theory [12]. The accumulation of stress on the pentamers of big shells will eventually lead to a buckling transition towards a faceted shape, and provides a mechanical justification of why for all larger capsids the polyhedral shape seems to be energetically favored.

8.6 Buckling of spherical capsids upon expansion

The energy analysis for the different shapes under radial deformation performed in section 8.4 reveals the interesting tendency of spherical capsids to become icosahedral upon expansion. This is illustrated for instance in Fig. 8.3 for a $T = 13$ capsid. Upon expansion, the energy of the spherical shell increases and eventually crosses the compression branch of the polyhedral structure. This



(a) (b)

Figure 8.8: Lateral stress distribution for (a) $P = 7$ and (b) $P = 13$ spherical and icosahedral shells. The color code is described in Fig. 8.6.

crossing defines a transition radius R_c beyond which the energy of the expanded spherical shell is larger than that of the compressed faceted capsid, thus signaling the energetic onset of buckling for this structure. By comparing the radii R_c with the optimal radii R_0^{sph} for different T-numbers (see Fig. 8.9) we obtain a linear dependence that suggests that spherical capsids will undergo a buckling transition when they are subjected to a radial expansion of $\approx 7\%$ if the strength and range of the interactions is kept constant. This is consistent with the capsid expansions observed during the maturation process of many viruses, which lead to the faceting of the initially spherical procapsid [6].

8.7 Bulk modulus

From the exploration of the energy under radial deformation, the bulk modulus of the optimal structures, K_0, can also be extracted. The bulk modulus quantifies the resistance of a substance in response to a uniform radial deformation (compression or expansion). This quantity is defined in general as the ratio of the infinitesimal pressure increase to the resulting relative decrease of the volume $K = -V\frac{dP}{dV}$ (see eq. 8.6), where V is the volume and P the pressure.

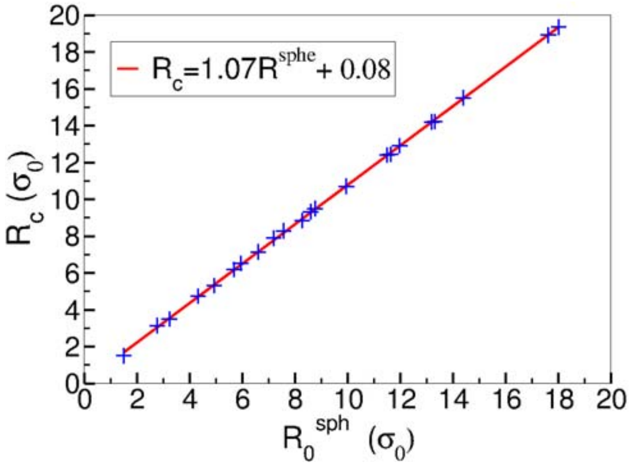


Figure 8.9: Spherical optimal radii R_0^{sph} as a function of the cut radius R_c. The line is a linear fit.

To get it, we first fit a parabola $E = E_0 + \frac{1}{2}a_E(R - R_0)^2$ around the energy minimum (see Fig. 8.3). Then, by substituting $a_E = (\frac{\partial^2 E}{\partial R^2})_{R_0}$ into Eq. 8.6, we determine K_0 for the different shapes and T-numbers (see Fig. 8.10). For both capsids shapes the bulk modulus decays inversely proportional to their corresponding optimal radii, i.e., $K_0 \approx 1/R_0$, so in general larger capsids will be more compressible. This is a a direct consequence of the geometrical prefactor $R^2/V \approx 1/R$ in Eq. 8.6 and the fact that the energy is proportional to R^2 under radial deformation. In addition, except for $T = 3$ and $T = 7$, the bulk modulus of faceted shells, K_0, is systematically larger than that of the spherical ones, K_0, by a factor ranging from $2-10\%$ for small capsids to almost 20% for bigger ones. Finally, there seems to be no special structural dependence in the bulk modulus of spherical shells. Contrarily, polyhedral $P = 1$ capsids are comparatively stiffer than those of any other class, see red points in Fig. 8.10.

8.8 Pressure and bursting

An important property of capsids is their tolerance to internal pressure before bursting [2, 24, 25]. For instance, bacteriophage $\phi 29$ can tolerate 30-60 atm [26] of internal pressure keeping its integrity and many bacteriophage use that internal pressure to inject the genetic material inside the cell [2, 3].

To compare the resistance of spherical versus polyhedral capsid shapes, we calculated the maximum pressure that will lead to lateral failure of the shells. For our simple Lennard-Jones potential, failure will take place when the separation between capsomers and the maximum force exceeds that at the inflection point, defined by $\frac{d^2 V_{LJ}}{dr_{ij}^2} = 0$. This inflection point is located at $r_{ij}^b = \left(\frac{13}{7}\right)^{1/6} \sigma_{hh}$ where we

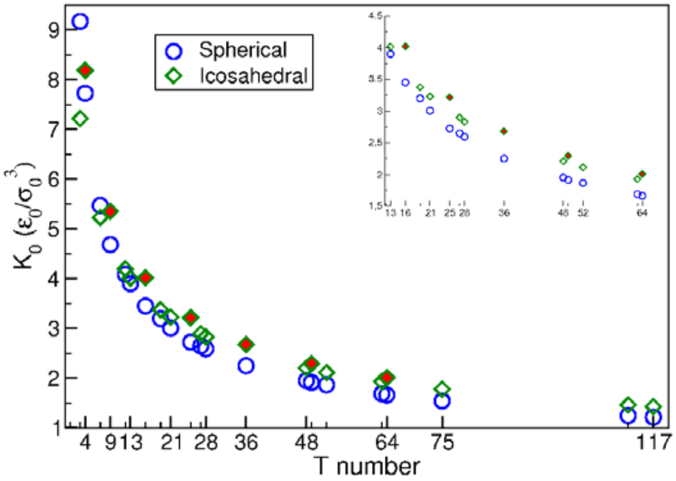


Figure 8.10: Bulk modulus for the spherical and icosahedral shells at the optimal radii. The inset shows a zoom in the region $13 < T < 64$ to highlight the fact that polyhedral $P = 1$ shells (colored in red) seem to be comparatively stiffer.

have used the distance between hexamers, since they tend to be most stretched subunits. At this inflection point, the Lennard-Jones force between two capsomers becomes

$$f^b = 12\frac{\epsilon_0}{\sigma}\left(\frac{7}{13}\right)^{13/6} - \frac{7}{13}^{7/6} \approx -2.69\frac{\epsilon_0}{\sigma} \tag{8.7}$$

Assuming now a flat 2D hexagonal lattice of capsomers one can compute the lateral stress at these condition, using eq. 8.3 yielding $\sigma_b^{2D} \simeq -1.1\frac{\epsilon_0}{\sigma^2}$. Accordingly, We defined the onset of bursting as the radius where the largest local stretching stress reaches the value $-1.1\frac{\epsilon_0}{\sigma^2}$. This is roughly the limit of resistance of a flat hexagonal lattice of LJ particles [30], and was also found to be a good approximation for the onset of bursting in the simulations of Ref. [20]. The bursting radii, R_b, are plotted in Fig. 8.11 and show a linear dependence on the optimal radii, which correspond to expansions of roughly $\approx 6\%$ and $\approx 9\%$ for spherical and polyhedral shells, respectively. Hence, polyhedral capsids resist larger relative expansions than their spherical counterparts. The associated internal pressures at the onset of bursting are plotted in Fig. 8.12. We observe that, at equal strength and range of interactions, polyhedral capsids systematically tolerate higher values of the bursting pressure, p_b, than their spherical counterparts, ranging from $\approx 10\%$ to $\approx 20\%$ more for the smallest T-numbers up to a $\approx 35\%$ for the largest viruses. This advantage is especially evident for the $P = 1$ capsids. However, in general the bursting pressure decays with the shell's size as $p_b \approx \frac{1}{R_b}$, because $V \approx R^3$ and under deformation $E \approx R^2$ (see eq. 8.5). Finally, it is worth mentioning that the use of scaled units in our study also makes it possible to compare the resistance of viruses with different strengths of interaction, by rescaling properly the pressure

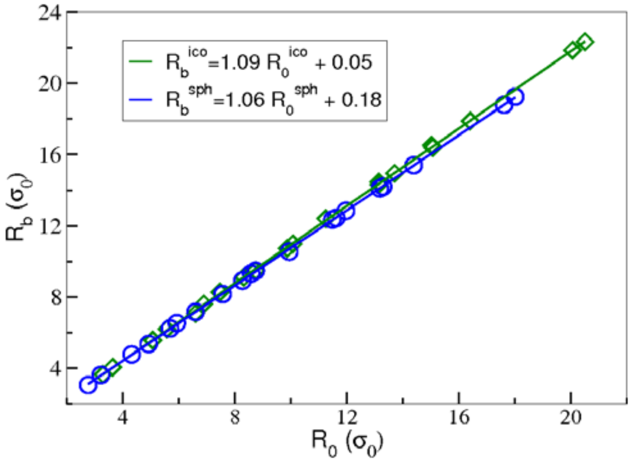


Figure 8.11: Bursting radius as a function of the optimal radii for the spherical and polyhedral shells. The lines represent linear fits to the data.

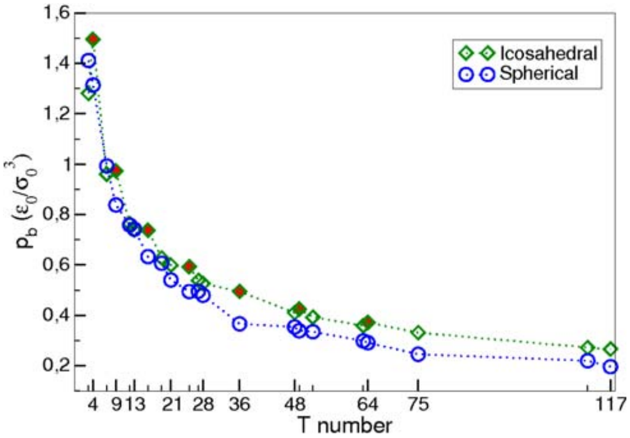


Figure 8.12: Bursting pressure as a function of the T-number. Icosahedral $P = 1$ shells are highlighted in red.

in terms of the corresponding binding energies and capsomer sizes.

8.9 Comparison with data from real viruses

The aim in this section is to contrast some of the predictions of our analysis with data from real viruses, and to discuss the potential biological implications of our results. In order to compare these predictions with known structures of viruses, we have used Chimera [28] to classify the shape and dimensions of all viruses with $T \geqslant 4$ listed in the Electron Microscopy Data Bank [29]. The procedure to evaluate the shape factor is described in detail in Appendix B.

The results obtained using this protocol are summarized in Table 8.1 and will

Table 8.1: T-number, radii and sphere factor of $T \geq 4$ viruses listed in the Electron Microscopy Data Bank [29].

Ref	Name	T number	Radius 5f Å	Form factor
1525	Thermusphage P23-77	28	380	0.61
1350	SH1	28	392	0.66
1662	Sputnik	27	402	0.08
1392	PhiKZ	27	695	0.01
1011	PRD1	25	345	0.12
1123	Bam35c	25	341	0.06
1490	Adeno	25	482	0.17
1085	PM2	21	304	0.22
5259	Herpes Simplex	16	681	0.08
1445	KSHV	16	625	0.07
1238	IBDV	13	349	0.21
1377	RDV	13	380	0.45
1207	Phi 6	13	265	1.00
5124	Rift Valley	12	498	0.36
1472	N4	9	367	0.12
5237	80alfa procapsid	7	253	0.79
1164	80alfa virion	7	302	0.26
1162	T7 procapsid	7	273	0.76
1164	T7 Virion	7	325	0.0
5187	Simian V. 40	7	212	0.65
1827	P22 procapsid	7	306	0.61
1826	P22 virion	7	343	0.04
5237	Phage Staphyl. aureus proc.	7	252	0.82
5236	Phage Staphyl. aureus virion	7	312	0.09
1691	Gifsy-2 procapsid	7	274	0.66
1694	Gifsy-2 virion	7	345	0.29
1334	K1E, K1-5	7	297	0.32
1339	Syn5	7	338	0.38
1400	Hep B	4	218	0.79
1608	NωV	4	166	0.68

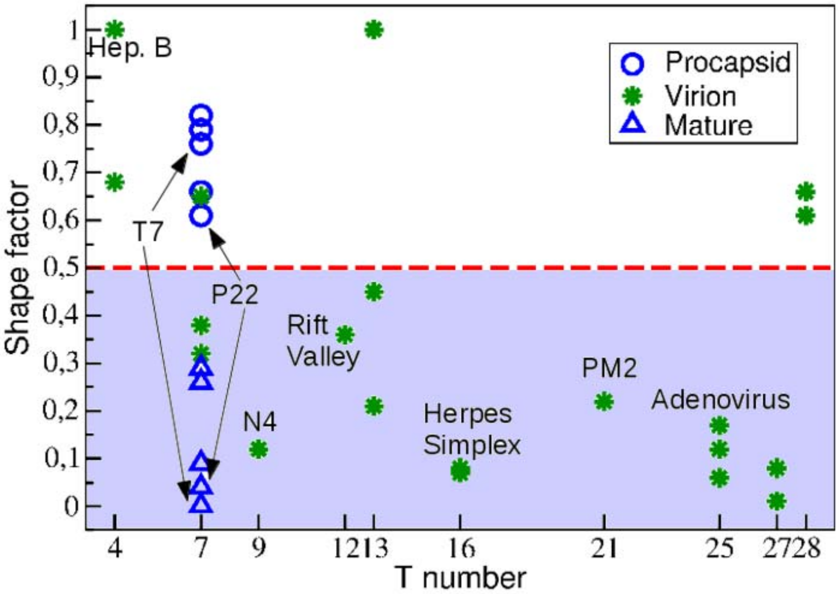


Figure 8.13: Shape factor as a function of the T-number for the different viruses listed in Table 8.1. For the cases where both the structures of the procapsid and the mature capsid are known, they are represented by open circles and triangles, respectively. The red line at 0.5 separates spherical ($sf > 0.5$) from icosahedral ($sf < 0.5$ blue region) structures.

be now used to discuss some of the predictions of our simple model. Two cautionary remarks have to be made regarding this comparison. First, the maturation process of a virus can be very complex, involving capsid expansion, cleavage of proteins, and even covalent reinforcement of the shell [6]. The predictions of our simple model are not expected to be applicable in this complex scenario, but rather focus on the simplest case of viruses with a maturation pathway that does not require auxiliary proteins, cleavage of the coat protein, nor cross-linking between subunits and where the interactions remain the some. Second, the known structures of large viruses are scarce and very often quite correlated, so the statistics of this classification may not be representative for all viruses. Nevertheless, overall the results mostly agree with the general predictions of the model.

In particular, if the strength of interactions and the size of the proteins does not change during maturation, our model predicts that the polyhedral capsids will be larger than the spherical proheads by ≈ 15% in radius, and that they become faceted upon a radial expansion larger than ≈ 7% in radius. Interestingly, for the few $T = 7$ viruses for which the structure of the procapsid and mature capsid are known, the ratio of the mature vs procapsid radii seems to agree with this prediction, although in some of these viruses the conformational changes in the capsid proteins that occur during maturation will restrict the applicability of our results. In all cases the procapsid is initially spherical (with a sphere factor $sf > 0.5$), and become faceted ($sf < 0.5$) after a maturation process that is accompanied by a significant expansion. Specifically, the ratio of the mature vs procapsid radii of 80alpha, T7, P22, Staphylococcus aureus bacteriophage, and Gifsy-2 viruses are 19%, 19%, 12%, 23% and 26% (see Table 8.1), respectively, which roughly coincide with the expected value of 17% for a $T = 7$ virus.

We stress that a simple surface-covering model accurately reproduces the radii obtained in the simulations (see Fig. 8.4), meaning that the result should be robust to different types of capsomer-capsomer interaction. However, if during virus maturation the interactions or the effective size of capsomers are strongly altered by conformational changes or chemical processes, important deviations could be expected. For instance, Nudaurelia capensis ω virus (NωV) is a $T = 4$ virus that during maturation undergoes a cleavage in the coat protein that leads to a final capsid 15% smaller in radius than the initial procapsid, i.e., the opposite behavior of the one expected in our study [32]. Another exception is CCMV that expands by more than 20% keeping its spherical shape by a change of pH and the removal of Ca2+ ions, which obviously will significantly modify the strength and range of the interactions. It is important to emphasize that, in some viruses, the maturation and the buckling transition is triggered by changes in the viral proteins and accordingly in their effective interactions; but this realistic scenario goes beyond the scope of the present work.

Another interesting consequence of our results is that for $T \leq 28$, the class P clearly dictates the energetic preference in adopting either a spherical or a polyhedral capsid. Fig. 8.14 plots the values of the sphere factor as a function of the T-number for all viruses listed in Table 8.1. Viruses are considered spherical if $sf > 0.5$ and icosahedral if $sf < 0.5$ (blue). We can note how structures from $P = 1$ class tend to be more faceted than structures from the $P = 3$ class that tend to be more spherical. In fact some large $T = 12$ dsRNA viruses like Rift Valley remain spherical even after maturation [35]. On the other hand, in viruses where the icosahedral shell is the most stable, like for $T = 25$, the capsid assembly might produce a polyhedral shell already as a procapsid, as has been observed in the $T = 25$ bacteriophage PRD1 [33].

As mentioned before, buckling in T-shells produces icosahedral capsids with radii ≈ 15% bigger than the spherical ones. However, the volume of the resulting icosahedron V_0^{ico} is ≈ 8% smaller than the spherical one V_0^{sph} (see Appendix A). This counterintuitive result, that occurs if no change in capsomer-size takes place, is coherent with the fact that both capsids have the same number of capsomers, i.e., the same surface, and the sphere is the shape that maximizes the volume/surface ratio. In addition, this result brings up an interesting question. Many viruses that undergo a maturation process with buckling end up storing dsDNA at high density [3], so the capsid must sustain an important internal pressure due to the confinement of this semi-flexible and electrostatically charged polymer. The ds-DNA density is correlated with the internal pressure [2], which means that the same amount of genetic material will generate a higher pressure in the icosahedral shell than in the spherical capsid. This might be a desirable effect in viruses that take advantage of this internal pressure to initiate the infection [2].

8.10 Conclusions

In this chapter we have analyzed the relevance of the T-number of viruses in the buckling phenomenon, using a simple physical model that captures the essential ingredients of capsomer-capsomer interactions and that successfully reproduces the equilibrium structures of viral capsids [17, 18, 20]. Despite the simplicity of the model, it is important to point out that the qualitative results obtained do not depend on the particular choice of the details of the intercapsomer potential, but are rather dictated purely by the geometrical arrangement of the capsomer in the T -shells. Obviously, the actual numerical values of properties like the local stresses, energies, pressures or bursting radius do depend on the details of the potential; but the trends and general conclusions that we summarize below seem to be insensitive to them and they are just determined by general physical and geometrical considerations.

We have found that, for small T-numbers, the tendency to buckle strongly depends on the class P. In particular, structures from the class $P = 1$, e.g., $T = 9$ or 16, are the most favorable to produce icosahedral shells, whereas capsids from the class $P = 3$ e.g., $T = 3$ or 12, are more stable as spheres. The only exception is $T = 4$ that, due to its large curvature and the large influence of pentamers, prefers to remain spherical. For the chiral classes $P > 3$ we obtain an intermediate behavior. Nevertheless, independently on P, for big capsids ($T > 28$) the icosahedral shell is always more stable than the spherical shell, in agreement with continuum elastic theory [12].

Furthermore, the analysis of the local lateral stress distribution unveils the microscopic explanation of the different tendency to buckle observed in different classes P. For spherical shells we generally observe highly squeezed pentameric zones and stretched hexamers, which is in agreement with the theory of disclinations in hexagonal lattices [12]. However, the class P dictates the capsomer arrangement and the resulting pattern of stress, which for spherical shells indicate the tendency of a structure to produce a polyhedral shell.

Interestingly, even in the cases where the spherical shape is more stable, viruses tend to undergo a buckling transition and to become polyhedral upon expansion, as it is often the case during virus maturation. In fact, several $T = 7$ viruses adopt spherical procapsids but end up becoming faceted after undergoing a buckling transition triggered by an expansion during virus maturation [6, 36].

More importantly, the choice of a polyhedral instead of spherical shape seems to have mechanical advantages that might play an important biological role. Faceted icosahedral shells have higher bulk moduli and tolerate larger expansions before bursting. In addition, they are able to withstand internal pressures $\approx 20\%$ higher than spherical capsids. All these enhanced properties could be advantageous for viruses that rely on a pressurized capsid to initiate the genetic material ejection [2]. However, since the maximum tolerable pressure is inversely proportional to the radius of the shell [20], this suggests that viruses could only take advantage of an internal pressurization mechanism for a specific range of low T's.

Finally, even though many viruses show a buckling transition during maturation, the pathways and processes involved can be very complex, including auxiliary proteins, cleavage or cross-linking. Obviously, our simple model cannot describe these pathways nor capture those complications. But our results could be helpful in understanding the biophysical advantages of undergoing a maturation and buckling process and adopting a faceted shape.

Appendix A

Surface Covering model

There is a simple way to estimate the optimal radius of a given shell based on a surface-covering assumption [37]. The idea is to assign an effective area to each capsomer and then compute the radius of the T-structure whose area can accommodate all the capsomers, assuming full coverage. In the spherical case an hexamer is represented by an hexagon of area $S_h = 2\sqrt{3}\sigma_0^2$ that circumscribes the capsomer of radius σ_0, and the pentamer by a pentagon with the same side as the hexagon, with a surface $S_p = \frac{5}{3} cotan(\frac{\pi}{5})\sigma_0^2$. For a T-shell the number of hexamers is $10(T-1)$ and there are always 12 pentamers. So the total surface is $S = 12S_p + 10(T-1)S_h$. Then, by equating the surface of the spherical shell, $S_{sph} = 4\pi R^2$ to the total area coming from the capsomers, we obtain the radius

$$R_{th}^{sph} = \sqrt{\frac{1}{4\pi}\left[10S_h T - 2(5S_h - 6S_p)\right]} = \sqrt{\frac{5\sqrt{3}T}{\pi}}\sigma_0 \tag{8.8}$$

For the icosahedral shell we proceed equivalently, but since the pentamers correspond to corners in the polyhedron, we take the effective area introduced in the CK construction, i.e., $S_p = \frac{5}{6}S_h$, corresponding to the removal of one of the triangles of the lattice. Taking into account that the surface of an icosahedron with radius R (measured from the center to one vertex) is $S_{ico} = \frac{40\sqrt{3}}{5+\sqrt{5}}R^2$ [38], we obtain:

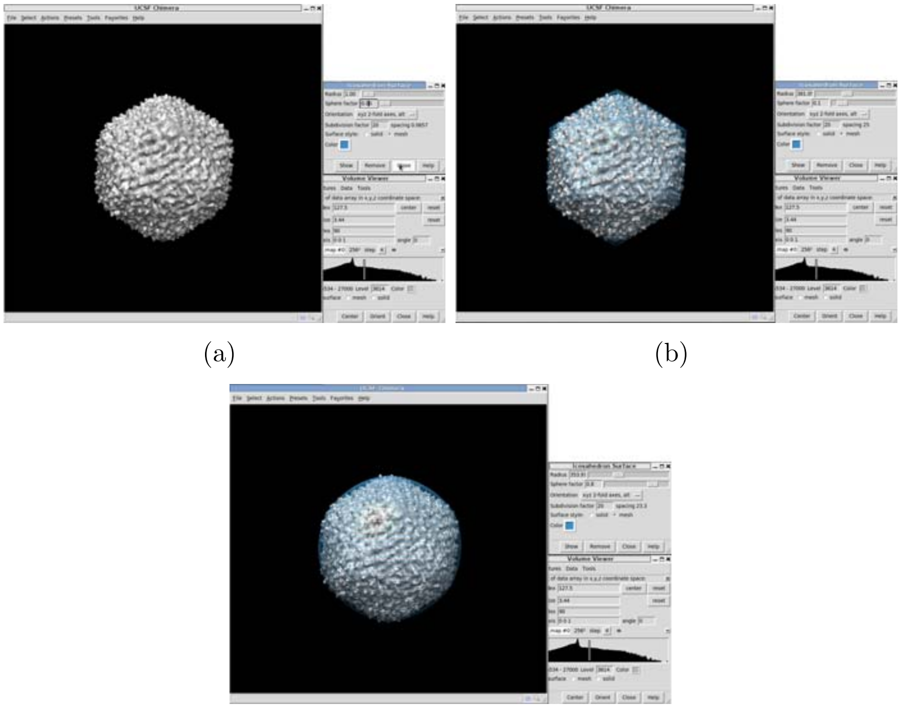
$$R_{th}^{ico} = \sqrt{\frac{5+\sqrt{5}}{2}T}\,\sigma_0 \tag{8.9}$$

The volume of the resulting icosahedron is $V = \frac{80(3+\sqrt{5})}{3\left(\sqrt{10+2\sqrt{5}}\right)^3}(R_{th}^{ico})^3$, which is $\sim 8\%$ smaller than that of a spherical capsid having the same surface area.

Appendix B

Determination of the shape factor of real viruses using Chimera

We have used Chimera [28] to classify the shape of all virus structures with $T \geqslant 4$ that are listed in the Electron Microscopy Data Bank [29]. For $T \leqslant 3$ the high curvature and the "decoration" of the capsid proteins makes difficult to characterize the shape. The structures are classified using the "sphere factor", sf, the Chimera term for the radial interpolation factor (0 to 1) between a flat faced icosahedron ($sf = 0$) and a sphere ($sf = 1$). The procedure used to evaluate this sphere factor for each structure is described below.



(a) (b)

(c)

Figure 8.14: (a) Electronic density map of bacteriophage PRD1 virus from Electron Microscopy Data Bank (grey). (b) Optimal icosahedral mesh with shape factor 0.1 and radius 381 $\mathring{A}$ showing a very good fit. (c) Icosahedral mesh with shape factor 0.8 and radius 353.9 $\mathring{A}$(blue) overlapped on the density map, showing that the fit is not optimal.

We first open the virus electronic density map with Chimera and place it at the center of the box. We then build an icosahedral surface with the same orientation of the virus, a similar radius and an arbitrary initial sphere factor $sf = 0$. We then convert this hypothetical structure into a volume map, try to fit both structures and measure the correlation of the fit. We then repeat the procedure in a loop that sweeps over different values of radii and of the sphere factor (between

0 and 1), and measures the correlation obtained in the fit. The final values of the radius and of the sphere factor for each structure are selected as the ones maximizing the correlation between the hypothetical icosahedron and the real virus.

References

[1] A. Cordova, M. Deserno, WM. Gelbart and A. Ben-Shaul, Osmotic shock and the strength of viral capsids, *Biophys. J.*, **85**, 70-74, 2003.

[2] WM. Gelbart and CM. Knobler, Pressurized viruses, *Science*, **323**, 1682-1683, 2009.

[3] A. Evilevitch, L. Lavelle and CM. Knobler, *et al.*, Osmotic pressure inhibition of DNA ejection from phage, *PNAS*, **100**, 16, 9292-9295, 2003.

[4] X. Agirrezabala, J. Martin-Benito, JR. Caston, R. Miranda, M. Valpuesta and JL. Carrascosa., Maturation of phage T7 involves structural modification of both shell and inner core components *EMBO J.* **24** 3820-3829, 2005.

[5] A. Ionel, JA. Velzquez-Muriel, D. Luque, A. Cuervo, JR. Caston, JM. Valpuesta, J. Martín-Benito and JL. Carrascosa, Molecular Rearrangements Involved in the Capsid Shell Maturation of Bacteriophage T7. *J. Biol. Chem.* **286** 234-242, 2010.

[6] AC. Steven, JB. Heymann, N. Cheng, BL. Trus and JF. Conway, Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr. Opin. Struc. Biol.* **15** 227-236, 2005.

[7] JE. Johnson, Virus particle maturation: insights into elegantly programmed nanomachines, *Curr. Opin. Struc. Bio.*, **20**, 2, 210-216, 2010.

[8] I. Gertsman, L. Gan, M. Guttman, K. Lee, JA. Speir, RL. Duda, RW. Hendrix, EA. Komives and JE. Johnson, An unexpected twist in viral capsid maturation. *Nature* **458** 7238, 2009.

[9] F. Tama and C. Brooks III, Diversity and Identity of Mechanical Properties of Icosahedral Viral Capsids Studied with Elastic Network, *J. Mol. Biol.*, **345** 299-314, 2005.

[10] LD. Landau and IM. Lifshitz 1975. Theory of Elasticity. Pergamon, New York.

[11] T. Nguyen, RF. Bruinsman and WM. Gelbart., Elasticity theory and shape transitions of viral shells. *Phys. Rev. E.* **72** 051923, 2005.

[12] J. Lidmar, L. Mirny and DR. Nelson, Virus shapes and buckling transitions in spherical shells. *Phys. Rev. E.* **68** 051910, 2003.

[13] M. Widom, J. Lidmar and DR. Nelson, Soft modes near the buckling transition of icosahedral shells. *Phys. Rev. E.* **76** 031911, 2007.

[14] TS. Baker, NH. Olson and SD. Fuller, Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microb. and Molec. Bio. Rev.* **63** 862-922, 1999.

[15] VS. Reddy, *et al.*, Energetics of quasiequivalence: computational analysis of protein-protein interactions in icosahedral viruses. *Biophys. J.* **74** 546-558, 1998.

[16] RH. Hardin, NJA. Sloane and WD. Smith Tables of spherical codes with icosahedral symmetry, published electronically at http://www.research.att.com/∼njas/icosahedral.codes/.

[17] R. Zandi, D. Reguera, RF. Bruinsma, WM. Gelbart and J. Rudnick, Origin of icosahedral symmetry in viruses. *Proc. Natl. Acad. Sci. USA* **101** 15556-15560, 2004.

[18] A. Luque, R. Zandi and D. Reguera, Optimal architectures of elongated viruses. *Proc. Natl. Acad. Sci. USA* **107** 5323-5328, 2010.

[19] DLD. Caspar and A. Klug, Physical principles in the construction of regular viruses, in Cold Spring Harbor Symp. Quant. Biol., **27**, 1-24, Cold Spring Harbor Laboratory Press, 1962.

[20] R. Zandi and D. Reguera, Mechanical properties of viral capsids. *Phys. Rev. E.* **72** 021917, 2005.

[21] ER. May, A. Aggarwal, WS. Klug and CL. Brooks III, Viral Capsid Equilibrium Dynamics Reveals Nonuniform Elastic Properties *Biophys. J.* **100**(11) L59-L61, 2011.

[22] WH. Roos, I. Gertsman, RE. May, CL. Brooks III, JE. Johnson, GJL. Wuite, Mechanics of bacteriophage maturation. *Proc. Natl. Acad. Sci. USA* **109** 2342-2347, 2012.

[23] X. Yan, NH. Olson, JL. Van Etten, M. Bergoin, MG. Rossmann and TS. Baker, Structure and assembly of large lipid-containing dsDNA viruses. *Nature structural biology* **7** 101-3, 2000.

[24] A. Siber, Buckling transition in icosahedral shells subjected to volume conservation constraint and pressure: Relations to virus maturation. *Phys. Rev. E.* **73(6)** 1-10, 2006.

[25] WS. Klug, RF. Bruinsma, JP. Michel, CM. Knobler, IL. Ivanovska, CF. Schmidt and GJ. Wuite, Failure of Viral Shells. *Phys. Rev. Lett.* **97** 1-4, 2006.

[26] M. Hernando-Pérez, R. Miranda, M. Aznar, JL. Carrascosa, I. T. Schaap, D. Reguera and PJ. de Pablo, Direct Measurement of Phage ϕ29 Stiffness Provides Evidence of Internal Pressure, *Small*, **8**, 15, 2366-2370, 2012.

[27] PK. Purohit, J. Kondev and R. Phillips, Mechanics of DNA packaging in viruses. *Proc. Natl. Acad. Sci. USA* **100** 3173-3178, 2003.

[28] EF. Pettersen, TD. Goddard, CC. Huang, GS. Couch, DM. Greenblatt, EC. Meng and TE. Ferrin, UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25** 1605, 2004.

[29] Electron Microscopy Data Bank (EMDB) http://www.ebi.ac.uk/pdbe/emdb/

[30] RLB. Selinger, Z-G. Wang and WM. Gelbart, Effect of temperature and small-scale defects on the strength of solids, *J. Chem. Phys.*, **95**, 9128, 1991.

[31] M. Aznar, A. Luque and D. Reguera, Relevance of capsid structure in the buckling and maturation of spherical viruses, *Phys Biol.* **9** 036003, 2012.

[32] B. Bothner, D. Taylor, B. Jun, KK. Lee, G. Siuzdak, CP. Schlutz, JE. Johnson, Maturation of a tetravirus capsid alters the dynamic properties and creates a metastable complex, *Virology*, **334**, 17-27, 2005.

[33] SJ. Butcher, DH. Bamford and SD. Fuller, DNA packaging orders the membrane of bacteriophage PRD1. *EMBO J.* **14** 6078-6086, 1995.

[34] KN. Parent, R. Khayat, LH. Tu, MM. Suhanovsky, JR. Cortines, CM. Teschke, JE. Johnson and TS. Baker, P22 coat protein structures reveal a novel mechanism for capsid maturation: stability without auxiliary proteins or chemical crosslinks. *Structure* **18** 390-401, 2010.

[35] MB, Sherman, A. Freiberg, MR. Holbrook and SJ. Watowich, Single-particle cryo-electron microscopy of Rift Valley fever virus. *Virology*, **387**, 1, 11-5, 2009.

[36] W. Jiang, Z. Li, Z. Zhang, ML. Baker, PR. Prevelige Jr. and W. Chiu, Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions, *Nature Struct. Biol.*, **10**, 131-135, 2003.

[37] A. Luque, Thesis. Structure, Mechanical Properties, and Self-Assembly of Viral Capsids, 2011.

[38] HSM. Coxeter, Introduction to geometry. Wiley; 2nd edition, 1989.

CHAPTER 9

Internal pressure in bacteriophage $\phi 29$

9.1 Introduction

The mechanisms used by a virus for genome delivery strongly depend on the kind of host [1]. Eukaryotic viruses, which infect animal and plant cells, are engulfed by the cell (endocytosis) and usually disassemble inside, releasing the genome. However bacteria present a complex outer envelope, and their viruses (bacteriophage) cannot use endocytosis. Instead the viral shell remains outside the host and the genetic material is translocated through the membrane helped in some cases by the binding of proteins [2] and/or the activity of polymerases inside the cytoplasm that pull it inwards [3]. However, experiments and theory [4] support the hypothesis that in some phages the initial stages of the genome injection into the host occurs by a spontaneous process triggered by the release of elastic and electrostatic energy stored by the dsDNA confined inside the capsid. This energy has been previously built up inside the phage during the DNA packing process.

In general, double-stranded DNA bacteriophages first assemble in an icosahedral prohead, which is later filled with DNA during the maturation process. Given the stiff and self-repulsive nature of dsDNA, its insertion inside the prohead requires the presence of a packing motor and the consumption of chemical energy in the form of ATP [5, 6]. At the end of the packing process, double-stranded (ds) DNA is confined to crystal densities [7], building a pressure within the phage with mainly a twofold origin. On one hand, dsDNA has to bend to fit inside a cavity (capsid) having a size similar to the dsDNA persistence length (50 nm). On the other hand, the spatial confinement reduces the neighbor DNA-DNA distances to a few $\AA$, and the electrostatic repulsion forces provoked by the DNA phosphates become crucial [8].

Although there seems to be no doubt that the DNA inside phages is in general pressurized, its role to help the entrance of DNA remains controversial [9]. There are several pieces of indirect experimental evidence for the pressurization of DNA inside phages. In particular, a variety of experiments have shown that by imposing an external osmotic pressure using osmolites, it is possible to counteract this internal pressure and to control the DNA ejection process in the phages λ [11] and T5 [12]. Furthermore, both theoretical [13, 14] and experimental [5, 15] results indicate that the packaging processes of the DNA inside phages requires forces about 50-100 pN. These forces can be translated into estimates of the pressurization of the genome inside the capsid, resulting in 50 to 100 atm. However, these techniques do not offer a direct measurement of the phage internal pressure.

In this chapter, we will describe a joint experimental and theoretical work providing a direct evidence and measurement of the internal pressure in a bacteriophage. In particular, we were able to determine the pressure inside bacteriophage $\phi 29$, combining AFM nanoindentation experiments on individual viruses with FE simulations and theoretical modeling. By comparing the effective spring constant of DNA full virions and DNA-devoid particles, we have been able to measure the contribution of the confined DNA to the stiffness of the shell and the subsequent evaluation of its internal pressure. Moreover, we also show that this pressure can be reversibly modified by the presence of counterions that reduce the electrostatic repulsion of the confined DNA.

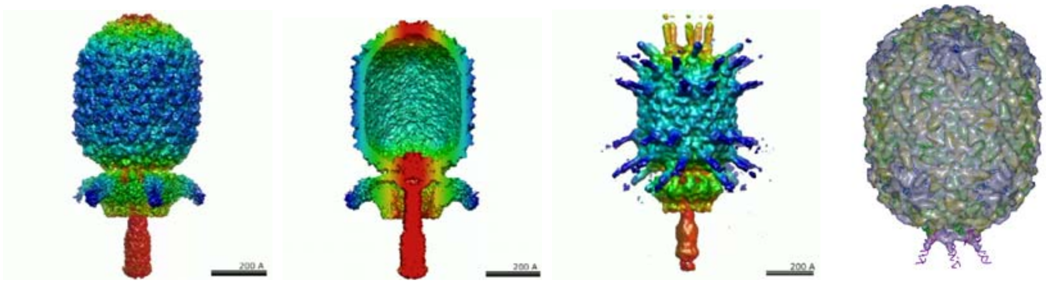
The chapter is organized as follows. In Section 9.2 we provide the relevant biological information on the structure and self-assembly of bacteriophage $\phi 29$. Section 9.3 is devoted to describe the AFM experiments that measure the mechanical properties of $\phi 29$. Section 9.4 introduces the finite element simulations used to model the previous AFM experiments and to describe the internal pressure. In section 9.5, we calculate the predictions for the pressure inside bacteriophage $\phi 29$ using the inverse spool model. Finally, in Section 9.6 we compare these predictions with the experimental results and discuss the biological implications on bacteriophage $\phi 29$.

9.2 Bacteriophage $\phi 29$

In this section we will summarize some relevant features of the viral cycle and structure of bacteriophage $\phi 29$. It is a virus from the family Podoviridae, that infects Bacillus subtilis and other related Bacillus bacterias.This double-stranded DNA bacteriophage is a prolate virus whose capsid dimensions are 54×42 nm and adopts an architecture $T_{end} = 3$ and $Q_{5F} = 5$ centered on a 5-fold axis [18, 19], see Fig. 9.1.

Bacteriophage $\phi 29$ is constructed from 235 gp8 subunits arranged in 11 pentameric plus 20 hexameric units forming icosahedral end caps, and 10 hexameric units forming the cylindrical equatorial region [20]. In one of the end caps, the central pentamer is replaced by the connector complex [21], which is a dodecameric assembly of protein gp10. The prolate icosahedra prohead is assembled by interaction of the connector protein (gp10), the scaffolding protein (gp7) and the major head protein (gp8). The absence of the connector or the scaffolding proteins (or mutants of these proteins) yields aberrant structures made of gp8 (open rounded shells, icosahedral capsids, tubular assemblies) [22] thus indicating that the built-in information of the major head protein is not sufficient to define the shape and size of the virus head, but rather it is the interaction of these components what direct the precise curvature and extension of the contacts to generate the shell architecture [19, 23]. An additional component of the head are fibers (made of protein gp8.5), which are dispensable for virus infectivity.

After the formation of the prohead, $\phi 29$ packages its 19.3 *kbp* dsDNA genome inside the capsid by means of a portal complex attached to the conector that hydrolyses ATP [19]. The maturation process starts when the DNA is packaged into



(a) (b) (c) (d)

Figure 9.1: Cryo Electron Microscopy 3D reconstruction of a ϕ29 particles. (a) External surface view of the prohead colored by cylindrical radius. (b) Cut of the prohead showing the empty interior (c) Mature virion with fibers taken from Ref. [16]. (d) Reconstruction of the fiberless bacteriophage ϕ29 mature virion, taken from Ref. [17].

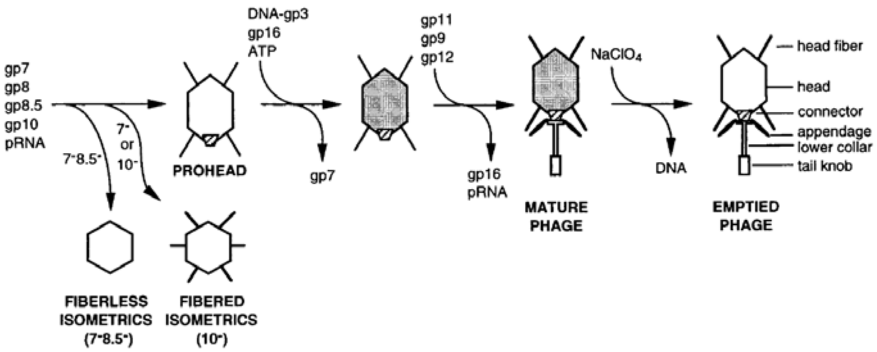


Figure 9.2: Assembly pathway of bacteriophage ϕ29 from the prohead to the mature virion. Taken from Ref. [20].

the prohead and it is accompanied by the release of the scaffolding proteins [24–26].

After the DNA incorporation the connector interacts with other tail components (gp11, gp12 and gp9) to secure the DNA (see Fig. 9.2) inside the head shell resulting in the mature virus. Interestingly, in the maturation process of ϕ29, there are no conformational changes in the capsid during DNA packaging. This bacteriophage, contrarily to other phages, does not show expansion or structural differences between the prohead and final virion beyond the presence of the tail and the fibers [24, 26, 27]. This final, mature virus particle is then ready for further infection cycles by attaching to host cells [26]. After attachment to the host, the entrance of the genome of phage ϕ29 apparently shows a two step push-pull mechanism [28]. During the first stage of genome injection, the push stage carries about 65% of the DNA through the tail into the host. Afterwards, during the pull stage, a variety of proteins of the cytoplasm pull the remaining DNA inwards. Since the push stage would be triggered by releasing the internal pressure of the phage, it is essential to determine the pressurization, if any, of the viral particles. This has been accomplished using Atomic Force Microscope nanoindentation experiments, as described in the following section.

9.3 AFM experiments

The Atomic Force Microscope (AFM) is a high-resolution microscope that allows the imaging and characterization of samples with nanoscale resolution. In recent years, it has been successfully applied to the study of the mechanical properties of individual viral capsids [10, 29, 30]. The AFM consists of a small tip (typically of about 15 nm radius) that is ultimately responsible for touching the sample (i.e, the capsid in our case). This tip is attached to a cantilever that can exert forces on the viral capsids [31]. The position of the tip is controlled using a piezoelectric device that can move over the sample in the three spacial directions by an applied electric voltage [32]. When a capsid is probed by the AFM tip, the cantilever suffers a deflection that is measured by a laser beam and translated into force units, see Fig. 9.3 [33]. Then using the AFM is possible to scan the surface and obtain several structural and mechanical properties of the viral capsids.

Experiments [30, 34] have shown intensively that as long as the viral shell is not indented beyond its thickness, viral particles deform linearly [29]. Hence, data from indentation experiments can be used to calculate the value of the effective spring constant k of the shell along the direction of the applied force, providing a quantitative description of the mechanical stiffness of the viral particle.

In order to know the role of the dsDNA in the origination of internal pressure in ϕ29, nanoindentation experiments were performed by the group of P.J. de Pablo at the Universidad Autonoma de Madrid on individual proheads, emptied mature capsids (both shells devoid of DNA) and virions (with DNA inside) in physiological condition. The samples were prepared as described in Appendix A, and the details of the AFM experiments are presented in Appendix B.

Fig. 9.4(a) shows single representative force versus indentation curves corresponding the prohead (red), virion (green), emptied virion (blue) and the experimental curve on the substrate (dashed) [33]. The results show a linear behavior of the force versus indentation, whose slope defines the effective spring constant k. Each particle is indented a few times in the linear regime, avoiding any damage, mechanical fatigue, or excessive tilting and the lineal portion of the force versus indentation curve is fitted to get the spring constant k.

The results corresponding to the slopes of 116 indentations carried out on 14 virions (green), 8 proheads (red), and 3 emptied virions (blue) performed at the

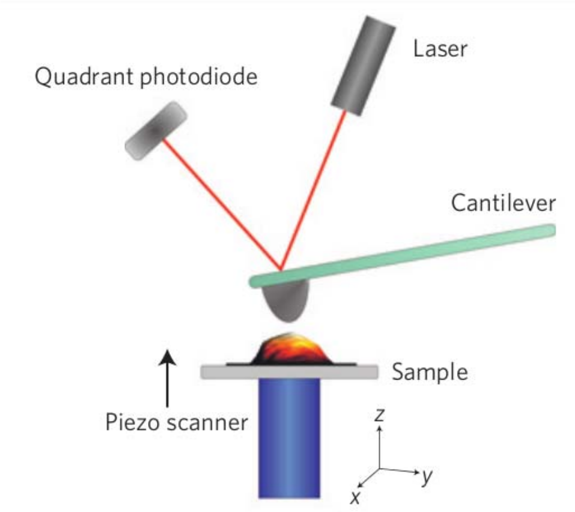
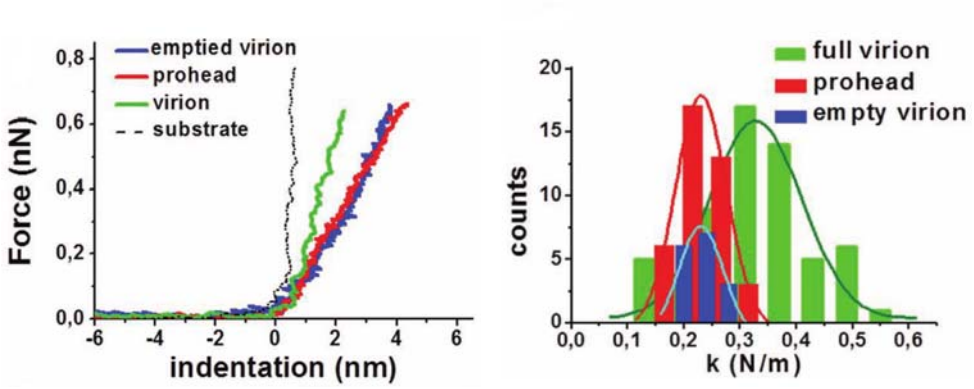


Figure 9.3: Schematic diagram of an AFM nanoindentation setup. A piezoelectric device controls the position of a cantilever that is progressively lowered on top of the sample. The deflection of the cantilever is detected by the reflected laser light on a photodiode, and it is used to calculate the force exerted on the sample. Image taken from Ref. [29].

same solution conditions and substrate are plotted as an histogram in Fig. 9.4(b). Gaussian fitting of the data results in spring constants of 0.23 ± 0.05, 0.33 ± 0.05 and 0.23 ± 0.04 N/m for proheads, virions and emptied virions, respectively. ϕ29 proheads and emptied virions, show a similar k in contrast with Murine leukemia virus [35] and HIV [36]. Our result can be interpreted in the light of the maturation process: unlike other phages such as T7 [37] or lambda [38], ϕ29 prohead does not experience relevant changes in the shell structure [26] that could modify its mechanical stiffness. Remarkably, ϕ29 DNA-containing virions exhibit larger value of k compared to empty virions of about 0.1 N/m that can be attributed to the presence of DNA inside.

The mechanical reinforcement of the virion associated to the presence of the genetic material may have two different origins. On one hand, the DNA-capsid interaction may confer a mechanical stiffening [39], in a similar way that beams buttresses the structure of a building. This is what happens, for instance, in the case of the Minute Virus of Mice [39], whose X-ray data [40] reveal strong interaction between the DNA and the capsid. However, the EM structure of the ϕ29 virion [41] does not show shell cavities with DNA, indicating a poor DNA-capsid interaction.

On the other hand, if the DNA is confined at high densities inside the capsid, it would generate an outwards force that will stiffen the shell [15]. The strategy to unravel the mechanical reinforcement mechanism in ϕ29 thus points at modify-



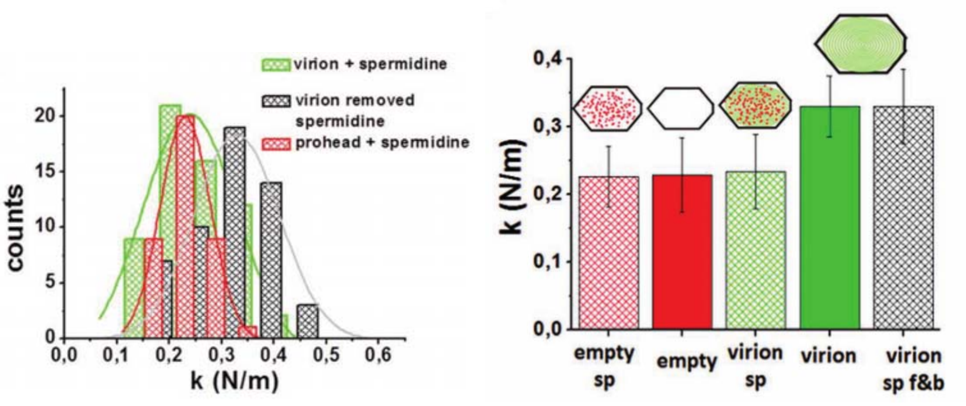
(a) (b)

Figure 9.4: Nanoindentation experiments. Indentation curves on samples: hard substrate (dashed), virion (green), the empty virion (blue), and prohead (red). b) Histograms of the spring constants.

ing the factors governing the energetics of confined DNA (i.e., DNA bending and DNA-DNA interactions) [42] to subsequently monitor the variations of the virus stiffness. We have chosen to focus on the leading energetic contribution governing DNA confinement, i.e., the DNA-DNA electrostatic repulsion [8].This repulsion originates from the fact that the negatively charged packaged DNA in phages such as lambda [38], T7 [43], ϵ15 [44], T4 [45] or ϕ29 [41] is usually structured in layers with interlayer distances ranging from 23 to 29 Å.

A simple way to reduce or eliminate DNA-DNA repulsion is to add small multivalent counterions that can diffuse through the capsid wall and screen the electrostatic interactions. In particular, spermidine (SP) is a trivalent polyamine commonly used to induce DNA condensates in solution [46] and thus we foresee the modulation of the DNA-DNA electrostatic repulsion in ϕ29 by adding SP. Thus, to understand the role of the DNA-DNA interactions in the reinforcement of the virion the nanoindentation experiments were repeated in the presence of $1mM$ of spermidine, as explained in Appendix A.

The Fig. 9.5 shows the results of this experiment obtained by performing 60 indentations on 12 particles. In the presence of 1 mM SP the virion spring constant decreases to $0.23 \pm 0.05N/m$, which is the same value as that of the emptied virion and the prohead (Fig. 9.5(b)). Control experiments by adding SP to prohead show that the shell itself is not affected by counterions (Fig. 9.5(a)), since the spring constant remains unaltered: $0.23 \pm 0.05N/m$. Moreover, it was found that the SP effect on the virions is reversible. The SP was removed by washing with buffer the AFM liquid chamber of the same viral particles previously measured in presence of SP. At the end of the process, it was verified (Fig. 9.4(a)) that the virion recovers its original value of elasticity: $0.33 \pm 0.05N/m$ (54 indentations on 8 particles).



(a) (b)

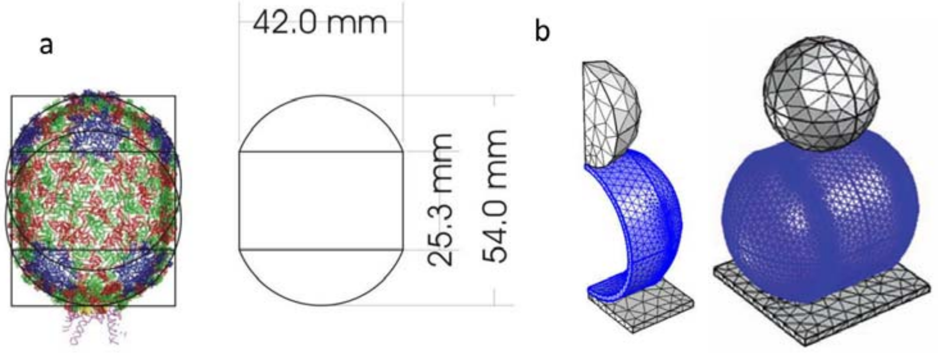
Figure 9.5: (a) Histograms of the spring constants obtained with and without spermidine. (b) Compilation of the the spring constants measured in the experiments: *empty SP* = empty particles incubated with spermidine; *empty* = empty shells; *virion SP* = virions incubated with spermidine; virion SP $f\&b$ = virions which have been incubated with spermidine, and later they have been washed out.

To summarize the results of the different experiments, Fig. 9.5(b) shows a compilation of the data, showing the SP effect on the different viral capsids. We can infer not only that the contribution of the DNA spring constant is about 0.1 N/m, but also that SP removes the mechanical role of DNA in a reversible manner.

AFM experiments only provide an effective measure of the mechanical stiffness of the capsid in terms of its spring constant. In order to translate the changes in stiffness measured in the experiments into a value of the pressure, we have performed a continuous elasticity analysis, using FE simulations, as described in the following section.

9.4 FEM simulations

We have used finite elements simulations method (FEM) [47] to reproduce the AFM experiment on the bacteriophage ϕ29 in order to link the changes in stiffness measured by AFM with the internal pressurization of the capsid. Finite Elements simulations of the AFM indentation of ϕ29 were performed using the program COMSOL Multiphysics 4.2a (Comsol, Stockholm, Sweden) as described in Chapter 5. In those simulations, the shape of the model was optimized to fit the geometry of ϕ29, as shown in Fig. 9.6(a). The capsid of ϕ29 was modeled as having a cylindrical body of length L and external radius R_{cyl} closed by two spherical caps of external radius $R = 22.5nm$. The centers of the spherical caps were placed $9nm$ apart, so that the total length of the virus is $45 + 9 = 54nm$, thus coinciding



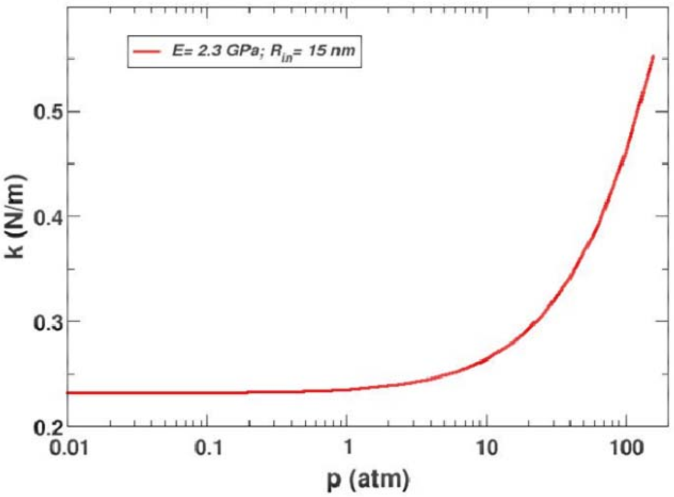
(a) (b)

Figure 9.6: Geometric representation used to generate the FE model optimized to fit the realistic shape and dimensions of $\phi 29$. (a) The caps are made from two spheres of external radius $22.5nm$ whose centers are placed $9nm$ apart. The cylindrical part has a radius of 21nm, a length $25.3nm$, and the thickness of the wall is $h = 1.6nm$. (b) FE model used in the simulations. The laid-down virus is indented by a spherical object on a surface.

with the reported experimental values [20]. The capsid wall consists of a homogeneous material with thickness $h = 1.6nm$, Young modulus E and Poisson ratio $\mu = 0.3$. This model capsid was placed on a hard flat substrate and indented by a spherical object with radius R_{in}. The internal surface of the capsid was subjected to a pressure p implemented as a boundary load. The contacts between the shell and the tip and the supporting surface during indentation were implemented with a contact penalty stiffness method according to the manufacturers manual. The model was simplified to a quarter by making use of symmetry planes, and meshed with over 25000 tetrahedral elements (see Fig. 9.6(b)). A parametric, non-linear solver was used to simulate the stepwise lowering of the tip onto the model. The spring constant was obtained from a linear fit of the force vs indentation, for small indentations. The radius of the indenter R_{in} and the Young's modulus E were varied to reproduce the experimental values, and to analyze its influence on the measured stiffness.

Fig. 9.7 shows the variation of the spring constant with pressure for $R_{in} = 15nm$ and a value of the Young modulus of $E = 2.3Gpa$ chosen to reproduce the experimental value of the empty (unpressurized) capsids. The spring constant reaches a value of $0.33N/m$ when the internal pressure is approximately $40atm$.

The precise value of the spring constant depends on the (unknown) elastic properties of the shell and, in principle, on the radius of the indenter. In order to ascertain their influence, we repeated the simulations for different values of the Young modulus in a reasonable range of $E = 1Gpa$ to $E = 10GPa$. Obviously, the values and behavior of the spring constant with pressure were different in each



(a)

Figure 9.7: Value of the spring constant of the shell k as a function of the internal pressure p obtained from the FEM simulations with $E = 2.3GPa$ and $R_{in} = 15nm$, chosen to reproduce the unpressurized spring constant k_0 measured in the AFM experiments.

case, as shown in Fig. 9.9. To analyze the influence of the radius of the tip, we redid the simulations using a smaller tip radius of $R_{in} = 5nm$. In addition, to analyze the validity of the thin shell approximation, commonly used in the literature, the simulations were repeated using a thin shell model with the same dimensions optimized to approximately fit to the center of the ϕ29 capsid wall. The model was made out of 5000 thin shell elements (so compression in the normal direction within the shell is ignored, and buckling is not accounted for), and indented in this case by a point force exerted at the middle of the cylindrical wall. The resulting indentation was measured as a function of the indentation force for different values of the internal pressure.

The results of all these different simulations are summarized in Fig. 9.8. The results for $E = 3.1GPa$ with different radius of the indenter, $R_{in} = 15nm$ (orange line) and $R_{in} = 5nm$ (black line) clearly show that this has a little influence on the measured spring constant. Moreover, comparing the results for $E = 3.1GPa$ (orange line) with the results obtained using a thin shell model indented by a point force (blue line), one can see that both the thin shell and the point force approximations are actually quite accurate to describe the indentation of a shell with the dimensions and thickness of ϕ29. Interestingly, if the spring constant k is scaled by its unpressurized value k_0, and the pressure is normalized as $\tau = \frac{pR}{k_0}$ we find that all data from these different simulations collapse into a single curve, see Fig. 9.9. This scaling is suggested by the interesting results in the thin shell

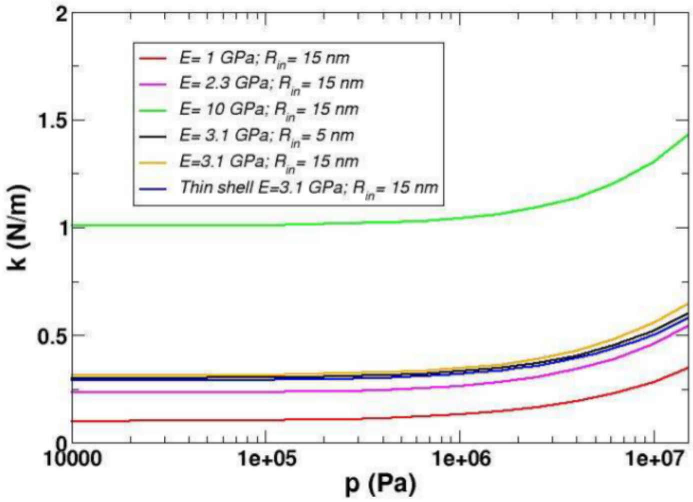


Figure 9.8: Value of the spring constant k versus the internal pressure from the FEM simulations(a) for different values of the Young modulus E and the indenter radius R_{in}. The blue line plots the results of a thin shell model with $E = 3.1GPa$ indented by a point force.

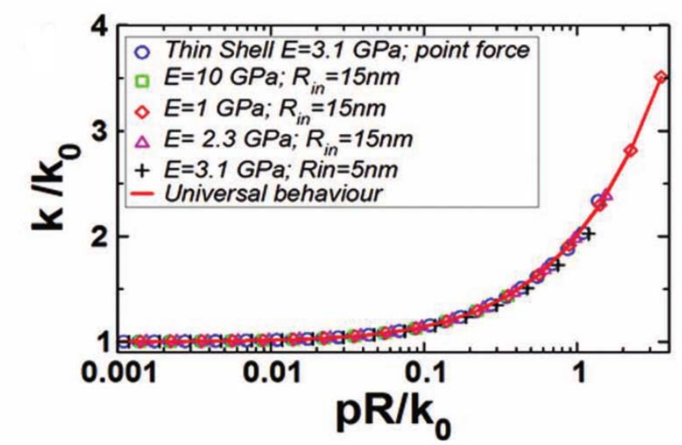


Figure 9.9: Value of the spring constant k versus the internal pressure from the FEM simulations scaled by the unpressurized value k_0, where the blue in circles are the results of a thin shell model indented by a point force the red line is a single universal line.

approximation for pressurized spherical and ellipsoidal shells found by Lazarus *et. al* [48] and Vella and co-workers [49, 50].

Fig. 9.9 plots the results of the effective spring constant scaled by its value for the unpressurized shell k_0, as a function of the dimensionless pressure $\tau = \frac{pR}{k_0}$, for different values of the Young modulus and the radius of the indenter. Remarkably, in this scaled form, all results collapse in a single universal curve that can be used to evaluate the pressure from the experiments, irrespective of the specific value of the mechanical properties of the capsid and the radius of the tip. From the experimental values of the spring constant of the virion ($k = 0.33 \pm 0.05 N/m$) and that of the empty (unpressurized) capsid ($k_0 = 0.23 \pm 0.05 N/m$), one gets $k/k_0 = 1.44 \pm 0.2$. That ratio corresponds in the master curve to a dimensionless pressure of $\tau = 0.36 \pm 0.2$, that yields a value of the pressure of $40 \pm 20 atm$. It is important to emphasize that the value of the pressure obtained from the AFM experiments and the FEM analysis is independent of the physical mechanism that is generating this pressure.

9.5 Estimation of the pressure from the inverse spool model

We now compare the obtained value for the pressure of bacteriophage ϕ29 with the theoretical predictions of the so-called “inverse spool model”.

There have been several models that evaluate the energetics of packed DNA [14, 51–53]. Here we will use the model introduced by Purohit et al [14] In this model the total free energy of confined DNA is assumed to consist of two main contributions: the elastic energy G_{bend}, due to the fact that DNA inside the capsid has to be bent at radius smaller than its persistence length (estimated in 50 nm); plus an interaction energy G_{int}, accounting mainly for the electrostatic interactions among neighboring DNA strands. Thus the total energy becomes:

$$G_{tot}(L, d_s) = G_{bend}(L) + G_{int}(L, d_s) \tag{9.1}$$

and depends on both the total length of the genetic material L, as well as on the separation between strands d_s.

The elastic contribution is calculated by assuming that the DNA inside the capsid is arranged in an inverse spool configuration where the strands are packed in a hexagonal array with a spacing d_s. With these assumptions, the elastic energy and the total DNA length become:

$$G_{bend}(L) = \frac{2\pi\xi_p k_B T}{\sqrt{3}d_s} \int_{R_{in}}^{R_{out}} \frac{N(R')}{R'} dR' \tag{9.2}$$

$$L = \frac{4\pi}{\sqrt{3}d_s} \int_{R_{in}}^{R_{out}} N(R') R' dR' \tag{9.3}$$

where $\xi_p = 50nm$ is the DNA persistence length, k_B is Boltzmann's constant, T is the temperature, $N(R')$ is the number of hoops of radius R' in the capsid,

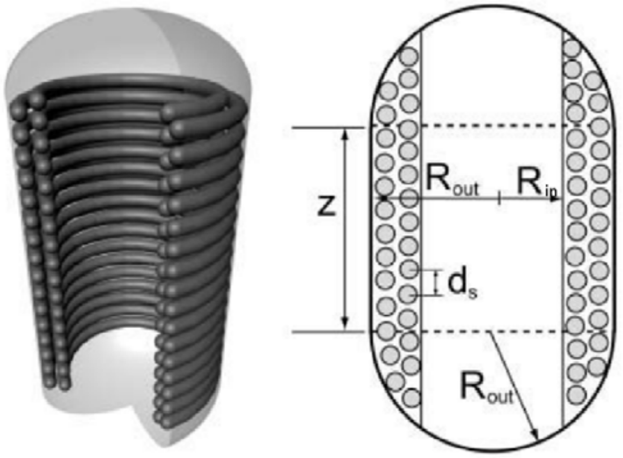


Figure 9.10: Inverse spool model, where $z = 9nm$, d_s is the separation between strands, $R_{out} = 22.5nm$ is the radius of the caps and $R_{in} = 19.4nm$ for the ϕ29 bacteriophage. Taken from Ref. [14]

and R and R_{out} are the inner and the outer radius of the inverse spool, the latter taken as the radius of the inner surface of the capsid (see Fig. 9.10).

An accurate and first principle calculation of the DNA-DNA interactions is a difficult and unarchived task, so commonly the interaction energy G_{int} is evaluated from osmotic stress experiments at the proper conditions. These experiments have shown that the electrostatic interaction between DNA strands has a different behavior depending on the valence of the added salt. For monovalent and divalent salts, the contribution is purely repulsive and the dependence of the osmotic pressure $\Pi^{rep}(d_s)$ with respect to the separation between strands d_s is given by:

$$\Pi^{rep}(d_s) = F_0 e^{-d_s/c} \tag{9.4}$$

where c and F_0 are parameters that characterize the decay length and strength of interactions, which depend on salt conditions. In the case of trivalent (like spermidine) and tetravalent salts the effective electrostatic interaction has an optimal distance between strands, which at smaller separations leads to a repulsive interaction whereas at higher separations can have an attractive regime. A convenient empirical expression for the osmotic pressure in this situation is

$$\Pi^{attr}(d_s) = F_0(e^{-(d_s-d_0)/c} - 1) \tag{9.5}$$

where d_0 is the optimal separation between strands (which for the case of spermidine has been measured to be $d_0 = 2.97nm$ [54]. From the osmotic pressure, the

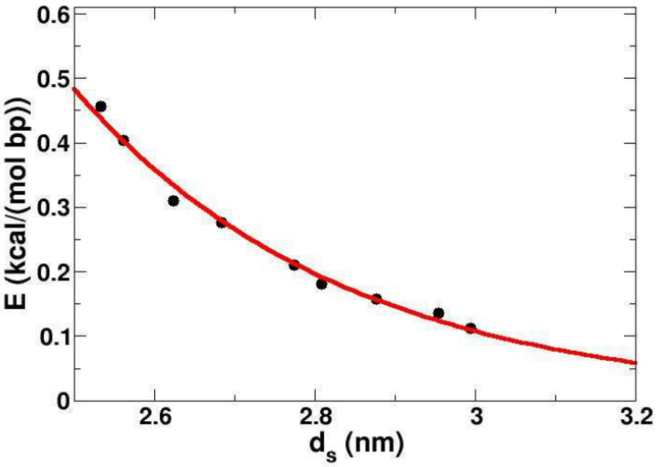


Figure 9.11: The free energy of interaction per basepair ($E = G_{int}l/L$, being $l = 0.34nm$) as a function of the DNA-DNA separation, d_s in TMS buffer ($100mMNaCl + 10mMMgCl_2$), as obtained from the osmotic stress experiments of Ref. [56]. The dots represent the experimental data and the line is a fit to eq. 9.7.

interaction energy can be evaluated as

$$G_{int}(L, d_s) = \sqrt{3}L \int_{\infty}^{d_s} \Pi(x)x dx \tag{9.6}$$

which yields

$$G_{int}^{rep}(L, d_s) = \sqrt{3}F_0 L(c^2 + cd_s)e^{-d_s/c} \tag{9.7}$$

in the purely repulsive regime and

$$G_{int}^{attr}(L, d_s) = \sqrt{3}F_0 L\left[(c^2 + cd_s)e^{-(d_s-d_0)/c} - (c^2 + cd_0) - \frac{1}{2}(d_0^2 - d_s^2)\right] \tag{9.8}$$

In the presence of tri- and tetra-valent counterions. It is important to emphasize that these interaction free energies G_{int} obtained from osmotic stress experiments not only account for the electrostatics of the DNA and counterions, but also for entropic and hydration contributions [55].

The values of the constants c and F_0 governing the electrostatic interactions for the different buffer conditions used in the present experiments have been obtained by fitting eqs. 9.4 and 9.5 to the osmotic stress experiments reported by Rau et al. for TMS buffer [56] and for a buffer containing spermidine [54], respectively.

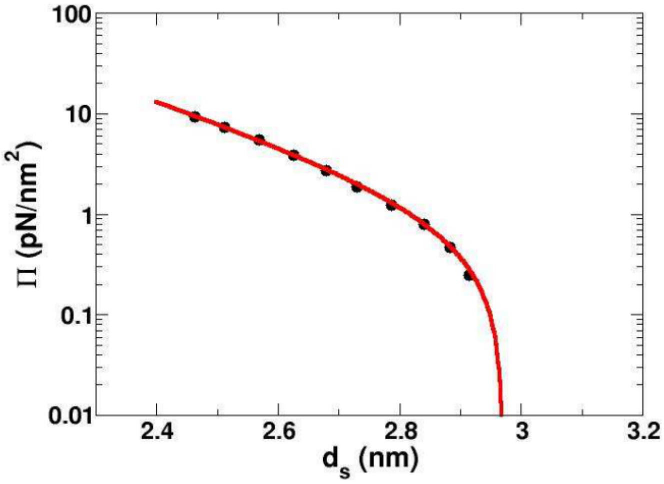
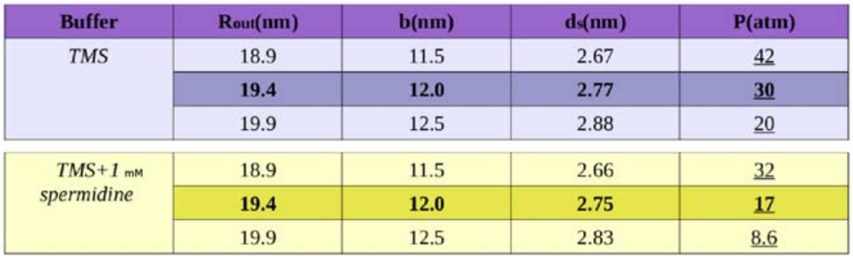


Figure 9.12: Osmotic stress as a function of the d_s in the presence of spermidine (attractive regime), as obtained in the experiments of Todd et al. [54](points). The line is a fit to eq. 9.5.

For the TMS buffer, the values obtained from the fit are: $F_0 = (3 \pm 1)10^4 pN/nm^2$ and $c = 0.30 \pm 0.01nm$. The data as well as the resulting fit are plotted in Fig. 9.11. In the case of the TMS buffer $+1mM$ spermidine, the values obtained from the fit are (see Fig. 9.12): $F_0 = 0.93 \pm 0.05pN/nm^2$ and $c = 0.210 \pm 0.005nm$, using an equilibrium separation of $d_0 = 2.97nm$ [54].

Knowing the contributions of the bending and interaction terms, the optimal energy and DNA spacing are then obtained by minimizing the total energy $G_{tot}(L, d_s)$ with respect to d_s for a given total genome length L. The internal pressure is finally obtained from it by its standard thermodynamic definition as $P = -\left(\frac{\partial G_{tot}}{\partial V}\right)$.

The capsid of ϕ29 is not spherical but elongated. In order to evaluate the pressure, we have assumed that it is an spherocylinder of internal radius $R_{out} = 19.4nm$, shell thickness $h = 1.6nm$, and length of the cylindrical body $b = 12nm$, which are the dimensions reported by Tao et al [20]. In Purohit et al [55] it was assumed that in elongated viruses the DNA is bent in an inverse spool configuration coiled around the cylinder axis. However, it seems more natural and energetically favorable from the elastic point of view (and it is also partially supported by the simulations [61]) to assume that it is coiled parallel to the cylinder axis. The advantage of this configuration is that it saves the bending energy associated to the unbent DNA in the cylindrical portion of the virus. With this configuration, the packaged length and the bending energy become



Buffer	Rout(nm)	b(nm)	ds(nm)	P(atm)
TMS	18.9	11.5	2.67	42
	19.4	**12.0**	**2.77**	**30**
	19.9	12.5	2.88	20
TMS+1 mM spermidine	18.9	11.5	2.66	32
	19.4	**12.0**	**2.75**	**17**
	19.9	12.5	2.83	8.6

Table 9.1 Values of the pressure P and the d_s obtained from the model in a TMS and TMS+1mM SP buffer using different values of the inner radius of the capsid R_{out}, the length of the cylindrical body b [33] and genome length $L = 19.3$ kbp.

$$L = \frac{8\pi}{3\sqrt{3}d_s^2}(R_{out}^2 - R_{int}^2)^{3/2} - \frac{4h}{\sqrt{3}d_s^2}\left(R_{in}\sqrt{R_{out}^2 - R_{in}^2} + R_{out}^2\left(sinh^{-1}\left(\frac{R_{in}}{R_{out}}\right) - \frac{\pi}{2}\right)\right)$$

$$G_{bend}(L) = -\frac{4\pi\xi_p k_B T}{\sqrt{3}d_s}\left(\sqrt{R_{out}^2 - R_{in}^2} - R_{out}log\left(\frac{R_{out} + \sqrt{R_{out}^2 - R_{in}^2}}{R_{int}}\right)\right) \quad (9.9)$$

Using this model, we have estimated the pressures and the optimal separation between strands at the two different buffer conditions used in the experiments. It turns out that the results are quite sensitive to the dimensions of the capsid used, specially to the radius. Table 9.1 and Fig. 9.14 show the values obtained for the TMS buffer and in the presence of 1mM of spermidine, using the internal dimensions of the capsid reported by Tao et al, and assuming a variation in them of $\pm 0.5nm$, which is similar to the resolution. For the internal dimensions of the capsid reported by Tao et al. [20] (i.e., considering ϕ29 a spherocylindrical shell of internal radius $R_{in} = 19.4nm$, length of the cylindrical body $b = 12nm$, and thickness $h = 1.6nm$), the results yield a pressure of approximately 30 ± 12 atm for the TMS buffer and 17 ± 12 atm for 1mm SP. The error bars have been estimated by considering a reasonable accuracy of the dimensions of the capsid of 0.5nm. These results are somehow consistent with the pressures measured by AFM.

It is important to note that the results do not differ much from the case in which the prolate ϕ29 is modeled as an effective spherical virus of the same volume as the prolate capsid. Moreover, although the fact that the dsDNA is compacted in an inverse spool configuration is still controversial [56], the estimates of the pressures are not very sensitive to the configuration, since they are mostly controlled by the electrostatic contributions [57] and the packing density of the DNA. In fact, a simple estimate of the strand separation can be obtained by equating the total volume available inside the capsid V with that of the hexagonally closed packed DNA (assuming that it occupies the whole volume), yielding $d_s =$

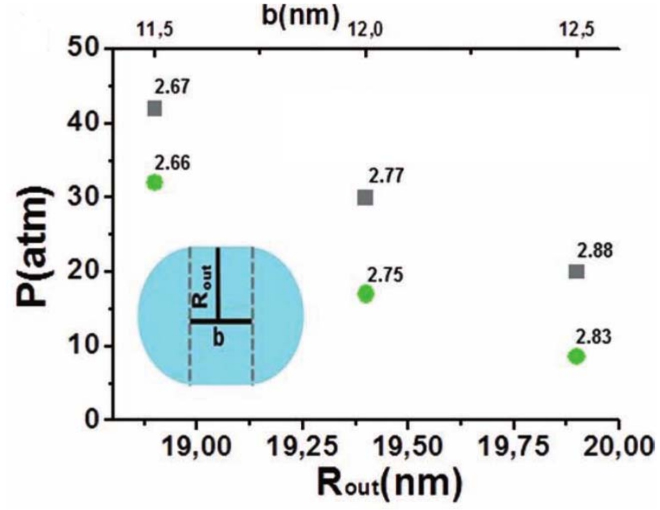


Figure 9.14: Different values of the pressure calculated with the Purohit model as a function of the radius of the cap R_{out} an its axial length b for TMS buffer (grey squares) and for TMS+1mM spermidine buffer (green circles). Numbers on top of the points are the DNA interlayer distance in nanometers (d_s). Taken from Ref. [14].

$\sqrt{\frac{2V}{\sqrt{3}}}\frac{1}{\sqrt{L}}$. Using this approximated value d_s in eqs. 9.4 and 9.5 one gets an accurate approximation of the capsid pressure due to the confined DNA. This suggests that the use of the inverse spool model is not critical in this case.

9.6 The pressure of ϕ29

In this section we are going to compare and discuss the experimentally measured value against the theoretical predictions.

From the AFM experiments and the FE analysis, we have obtained a value of the pressure in full ϕ29 particle of 40 ± 20 atm. This value is compatible with the 30 ± 12 atm estimated theoretically from the inverse spool model.

In the presence of spermidine, the packaging model predicts still a slight pressurization of $17 \pm 12 atm$, that using the universal curve of Fig. 9.9 will translate in an expected spring constant of $k_{SP} = 0.28 \pm 0.05 N/m$. This value is also compatible with the experimental results, within the experimental error bars. Therefore, our indentation experiments evaluate the pressure of the DNA inside the capsid, thus reflecting the trends established by the predictions of the electrostatic theoretical model.

Finally, the value of the pressure measured in the AFM experiments (40 atm)

is smaller than the estimate of 60 atm based on packaging force measurements using optical tweezers [5, 51, 55]. We attribute this discrepancy to the buffer used in those experiments that contained half the concentration of monovalent and divalent counterions (5 $mmMgCl_2$ and 50 $mmNaCl$), leading to larger repulsion between DNA strands, and accordingly to larger internal pressures.

9.7 Conclusions

In this chapter, we have seen how the combination of experiments and simulations has provided a direct way to determine the genome pressure in bacteriophage $\phi29$, much in the same way that pressure gauges evaluate the pressure in an inflated tire.

From this analysis, we have also obtained many interesting conclusions of biological relevance. On one hand we have shown that the spring constant of the emptied $\phi29$ virions and procapsids is the same, revealing a poor mechanical role of the scaffolding protein, if any. On the other hand, we have also proven that the presence of confined DNA stiffens the $\phi29$ phage in about $0.1N/m$ with respect to the emptied virion or capsid. By using purely mechanical models, we have inferred from the experiments that an internal pressure of $40 \pm 20atm$ is required for generating the observed stiffness variation between capsids and virions, irrespective to its physical origin. Interestingly, the value of this pressure is similar to that predicted by the DNA packaging model.

We have also established that the trivalent agent SP is able to eliminate reversibly the mechanical stiffening provided by DNA, since it screens the DNA-DNA repulsion in a reversible way. The results of this work constitute the first direct evaluation of the pressure built up by the confined DNA in bacteriophage $\phi29$.

The genome in other dsDNA bacteriophage as λ [38, 58, 59], T7 [37, 43], $\epsilon15$ [44] or T4 [45] is structured in layers with interlayer distances ranging from 23 to 29 $\AA$, similar to those of bacteriophage $\phi29$ (24 $\AA$). Thus, similar experiments and the same theoretical analysis can be used to determinate the pressure for other phages and the results will be crucial to understand the controversial role of pressure in the DNA entrance into the host cells [4].

Appendix A

Preparation of the $\phi 29$ samples.

To produce emptied $\phi 29$ particles, virions were incubated in $100 mM$ EDTA at $65^o C$ during 40 minutes. After dialysis of the sample against TMS to remove the EDTA, it was further incubated with DNase for 30 min at $37^o C$ to remove the DNA released from the virions. To evaluate the effect of the trivalent ion spermidine on the spring constant and DNA-DNA interaction the stock solution of virion $\phi 29$ was diluted 50 times in TMS buffer and then 5 times in TMS buffer with spermidine to reach $1 mM$ concentration. The drop was left about 30 minutes on HOPG and rinsed four times with $1 mM$ spermidine in TMS. As an additional control to test the direct effect of spermidine in the decrease of the interaction DNA-capsid, the following experiment was performed: the virion was diluted in TMS+$1 mM$ spermidine. A 25 μl aliquot was left on the HOPG surface 30 minutes. Then, the sample was rinsed 4 times with TMS-Spermidine buffer to reach progressively the buffer standard conditions (0.75 mM, 0.5 mM, 0.25 mM and 0 mM spermidine, respectively). Each buffer change was left about 20 minutes to stabilize.

Appendix B

AFM experiments on viral particles.

Stocks of $\phi 29$ viral particles were stored in TMS buffer (10 mM $MgCl_2$, 50 mM Tris and 100 mM NaCl, pH 7.8). A drop of 20 μl stock solution of viral particles was incubated on HOPG (ZYA quality NTMDT), for 30 min and washed with buffer. The tip was prewetted with 20 μl of buffer. The AFM (Nanotec Electrónica S.L., Madrid, Spain) was operated in jumping mode in liquid [32] using rectangular cantilevers RC800PSA, and Biolevers (BLRC150VB) (Olympus, Tokyo, Japan) with nominal spring constants of 0.05 N/m and 0.03 N/m, respectively. Cantilevers spring constants were routinely calibrated by using the Sader's method [31].

In the experiments, viral particles were immobilized on a flat surface for their AFM inspection [60]. Afterwards the adsorption geometry of the $\phi 29$ virions, the emptied virus and the proheads were identified by comparing AFM and geometrically tip dilated [62] electron microscopy (EM) data. The comparison between $\phi 29$ proheads and emptied virions does not show further structural differences beyond the tail and the fibers [26]. The absence of the tail allows proheads to adsorb on the substrate in twofold geometries: laid down on the side and upright [30, 34]. Fig. 9.15 shows a typical topography AFM micrograph of a laid down prohead particle which compares fairly well with the tip dilated [62, 63] EM data [17]. However, virions were constrained to attach laid down on the side because of the tail, resulting in a single and clean adsorption geometry.

Fig. 9.15(b) shows a typical AFM image of a virion with evident features of the tail and the collar which again matches with the expected tip-dilated model of the EM data [41]. Furthermore, the measurements indicate that the capsid is resting parallel to the surface in a geometry that is stable enough for our experiments. Emptied virions particles are similar to the full virion presented in Fig. 9.15 and they did not lose the tail during the process.

In order to perform nanoindentations, single viral particles were deformed by the tip by carrying out single FZs (force versus z-piezo displacement) experiments most likely right at the top of the virus: the shell was zoomed in continuously by reducing the x-y scanning size until the bump of the very top is under the whole piezo scan. Afterwards the FZ was executed at the top of the particle, probably with a few nm of uncertainty mainly provoked by the thermal drift, and the intrinsic nonlinearity and creep of the piezo. Still, this method has been proved to be robust enough to establish electrical contact with carbon nanotubes which are even smaller than viral particles. During the first stages of indentation the viral particles show a linear deformation [64] which provides the spring constant of the virus k_v (if it is considered like a spring in series with the cantilever) as $k_v = k_c \frac{S_g}{S_v - S_g}$, where k_c is the spring constant of the cantilever, S_g (nm/V) the

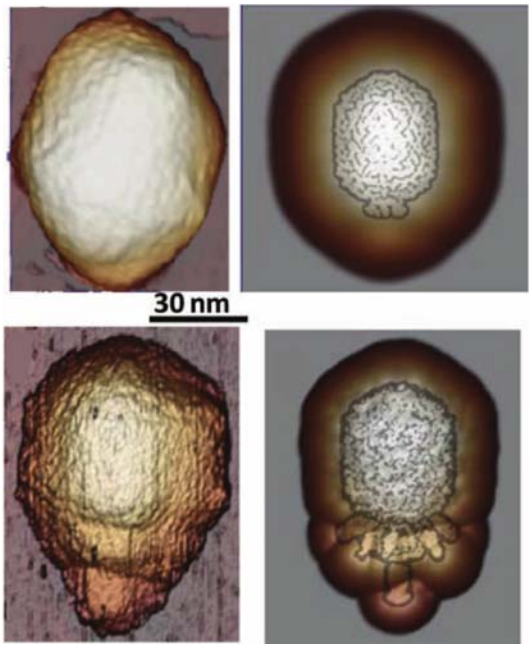


Figure 9.15: (Top row) AFM images of a ϕ29 prohead adsorbed on the surface and the geometrical dilation image of the prohead EM data. (Botton row) AFM images of a virion or emptied virion, which compares fairly well with the geometrically dilated EM data taken from Ref. [41].

slope of the cantilever deflection on the glass (here the substrate was considered as non deformable) and S_v (nm/V) the slope of the cantilever deflection on the virus. Hence, the viral particle is indented with a few separate sets of FZs of about 5 indentations in each one. Following each FZ set, an image of the virus is taken to confirm its integrity, as well as to know its position in order to correct for any drift if needed to perform the next FZs set. The maximum force applied during each FZ never exceeded $\approx 300pN$ to prevent the damage [65], collapse [64], buckling [66], or non-linear deformations [67] of the shells. The FZ speed is about 60 nm/s [67]. Even if the shell integrity is maintained, in our experiments only particles showing stable spring constants along the FZs sets were considered to avoid particle mobility effects that often occur when the particle is loosely bound to the surface. Images were processed using the WSxM software.

References

[1] J. Carter, V. Saunders, Virology: Principles and Applications, Wiley, New Jersey 2007.

[2] W. M. Gelbart, R. Zandi, D. Reguera, J. Rudnick, What drives the translocation of stiff chains?, *PNAS*, **100**, 8649, 2003.

[3] C. M. Knobler, W. M. Gelbart, Physical Chemistry of DNA Viruses, *Annu. Rev. Phys. Chem.*, **60**, 367, 2009.

[4] W.M. Gelbart and C.M. Knobler, Pressurized viruses, *Science*, **323**, 1682-1683, 2009.

[5] D.E. Smith, S.J. Tans, S.B. Smith, S. Grimes, D. L. Anderson, and C. Bustamante, The bacteriophage $\phi 29$ portal motor can package DNA against a large internal force, *Nature*, **413**, 748-752, 2001.

[6] Y.R. Chemla, K. Aathavan, J. Michaelis, S. Grimes, P.J. Jardine, D.L. Anderson, and C. Bustamante, Mechanism of force generation of a viral DNA packaging motor, *Cell*, **122**, 683-692, 2005.

[7] W.C. Earnshaw, S.C. Harrison, DNA arrangement in isometric phage heads, *Nature*, **268**, 598, 1977.

[8] D.C. Rau, B. Lee, V.A. Parsegian, Measurement of the repulsive force between poly-electrolyte molecules in ionic solution-hydration forces between parallel DNA doble helices, *PNAS*, **81**, 2621, 1984.

[9] D. Panja and I.J. Molineux, Dynamics of bacteriophage genome ejection *in vitro* and *in vivo*, *Phys. Biol.* , **7**, 045006, 2010.

[10] MG. Mateu, Structure and Physics of Viruses, Springer, London, 2013.

[11] A. Evilevitch, L. Lavelle, C. M. Knobler, E. Raspaud, and W. M. Gelbart, Osmotic pressure inhibition of DNA ejection from phage , *PNAS*, **100**, 9292-9295, 2003.

[12] A. Leforestier, S. Brasiles, M. de Frutos, E. Raspaud, L. Letellier, P. Tavares, F. Livolant, Bacteriophage T5 DNA Ejection under Pressure, *J. Mol. Biol.*, **384**, 730, 2008.

[13] S.Tzlil, J.T. Kindt, W.M. Gelbart, A. Ben-Shaul, Forces and pressures in DNA packaging and release from viral capsids, *Biophys. J.*, **84**, 1616, 2003.

[14] P.K. Purohit, M.M. Inamdar, P.D. Grayson, T.M. Squires, J. Kondev, R. Phillips, Forces during Bacteriophage DNA Packaging and Ejection, *Biophys. J.*,**88**, 851,2005.

[15] J.P. Rickgauer, D.N. Fuller, S. Grimes, P.J. Jardine, D.L. Anderson, D.E. Smith, Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage $\phi 29$, *Biophys. J.*, **94**, 159, 2008.

[16] Electron Microscopy Data Bank (EMDB) http://www.ebi.ac.uk/pdbe/emdb/

[17] M.C. Morais, K.H. Choi, J.S. Koti, P.R. Chipman, D.L. Anderson, and M.G. Rossmann, Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of $\phi 29$, *Mol. Cell*, **18**, 149-159, 2005.

[18] A. Luque, R. Zandi, and D. Reguera, Optimal architectures of elongated viruses, *PNAS* **107**, 53235328, 2010.

[19] P.X. Guo, S. Erickson, W. Xu, N. Olson, T.S. Baker, and D. Anderson, Regulation of the phage-$\phi 29$ prohead shape and size by the portal vertex, *Virology* , **183**, 366-373, 1991.

[20] Y. Tao, N.H. Olson, W.Xu, D.L. Anderson, M.G. Rossmann, and T.S. Baker, Assembly of a tailed bacterial virus and its genome release studied in three dimensions, *Cell*, **95**, 431-7, 1998.

[21] B. Ibarra, J.R. Caston, O. Llorca, M. Valle, J.M. Valpuesta, and J.L. Carrascosa, Topology of the components of the DNA packaging machinery in the phage $\phi 29$ prohead, *J. Mol. Biol.*, **298**, 807-815, 2000.

[22] A. Camacho, F. Jimenez, J. Delatorre, J. L. Carrascosa, R. P. Mellado, C. Vasquez, E. Vinuela, and M. Salas, Assembly of bacillus-subtilis phage-phi-29. 1. Mutants in cistrons coding for structural proteins , *Eur. J. Biochem.*, **73**, 39-55, 1977.

[23] K.H. Choi, M.C. Morais, D.L. Anderson, and M.G. Rossmann, Determinants of bacteriophage $\phi 29$ head morphology, *Structure*, **14**, 1723-1727, 2006.

[24] A.C. Steven, J.B. Heymann, N. Cheng, B.L. Trus, and J.F. Conway, Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity, *Curr. Opin. Struc. Biol.*, **15**, 227-36, 2005.

[25] T. Dokland, Scaffolding proteins and their role in viral assembly, *Cell. Mol. Life Sci.*, **56**, 580-603, 1999.

[26] Y. Xiang, M.C. Morais, A.J. Battisti, S. Grimes, P.J. Jardine, D.L. Anderson, and M.G. Rossmann, Structural changes of bacteriophage $\phi 29$ upon DNA packaging and release, *EMBO J.*, **25**, 5229-5239, 2006.

[27] J.E. Johnson, Virus particle maturation: insights into elegantly programmed nanomachines, *Curr. Opin. Struc. Biol.*, **20**, 210-216, 2010.

[28] V. Gonzalez-Huici, M. Salas, and J.M. Hermoso, The push-pull mechanism of bacteriophage ϕ29 DNA injection, *Mol. Microbiol.*, **52**, 529-540, 2004.

[29] W. H. Roos, R. Bruinsma, G. J. L. Wuite, Physical virology, *Nat. Phys.*, **6**, 733, 2010.

[30] I. L. Ivanovska, P. J. C. Pablo, B. Ibarra, G. Sgalari, F. C. MacKintosh, J. L. Carrascosa, C. F. Schmidt, G. J. L. Wuite, Bacteriophage capsids: Tough nanoshells with complex elastic properties, *PNAS*, **101**, 7600, 2004.

[31] J.E. Sader, J.W.M. Chon, and P. Mulvaney, Calibration of rectangular atomic force microscope cantilevers, *Rev. Sci. Instrum.*, **70**, 3967-3969, 1999.

[32] F. Moreno-Herrero, P.J. de Pablo, M. Alvarez, J. Colchero, J. Gomez-Hertero, and A. M. Baro, Jumping mode scanning force microscopy: a suitable technique for imaging DNA in liquids, *Appl. Surf. Sci.*, **210**, 22, 2003.

[33] M. Hernando-Pérez, R. Miranda, M. Aznar, JL. Carrascosa, I. T. Schaap, D. Reguera and PJ. de Pablo, Direct Measurement of Phage ϕ29 Stiffness Provides Evidence of Internal Pressure, *Small*, **8**, 15, 2366-2370, 2012.

[34] C. Carrasco, A. Luque, M. Hernando-Perez, R. Miranda, J.L. Carrascosa, P.A. Serena, M. de Ridder, A. Raman, J. Gomez-Herrero, I.A.T. Schaap, D. Reguera, P.J. de Pablo, Built-In Mechanical Stress in Viral Shells, *Biophys. J.*, **100**, 1100, 2011.

[35] N. Kol, M. Gladnikoff, D. Barlam, R. Z. Shneck, A. Rein, I. Rousso, Mechanical properties of murine leukemia virus particles: Effect of maturation, *Biophys. J.*, **91**, 767, 2006.

[36] N. Kol, Y. Shi, M. Tsvitov, D. Barlam, R. Z. Shneck, M. S. Kay, I. Rousso, A stiffness switch in human immunodeficiency virus, *Biophys. J.*, **92**, 1777, 2007.

[37] J.L. Carrascosa, A. Ionel, J.A. Velazquez-Muriel, D. Luque, A. Cuervo, J.R. Caston, J.M. Valpuesta, J. Martin-Benito, Molecular Rearrangements Involved in the Capsid Shell Maturation of Bacteriophage, *J. Biol. Chem.*, **286**, 234, 2011.

[38] J.E. Johnson, G.C. Lander, A. Evilevitch, M. Jeembaeva, C.S. Potter, B. Carragher, Bacteriophage Lambda Stabilization by Auxiliary Protein gpD: Timing, Location, and Mechanism of Attachment Determined by Cryo-EM, *Structure*, **16**, 1399, 2008.

[39] C. Carrasco, A. Carreira, I.A.T. Schaap, P.A. Serena, M.G., Mateu, and P.J. de Pablo, DNA-mediated anisotropic mechanical reinforcement of a virus , *PNAS.* , **103**, 13706-11, 2006.

[40] J. Tsao, M. S. Chapman, M. Agbandje, W. Keller, K. Smith, H. Wu, M. Luo, T. J. Smith, M. G. Rossmann, R. W. Compans, C. R. Parrish, The 3-Dimensional Structure of Canine Parvovirus and Its Functional Implications, *Science*, **251**, 1456, 1991.

[41] JH. Tang, N. Olson, P.J. Jardine, S. Girimes, D.L. Anderson, T.S. Baker, DNA Poised for Release in Bacteriophage ϕ29, *Structure*, **16**, 935, 2008.

[42] H.G. Garcia, P. Grayson, L. Han, M. Inamdar, J. Kondev, P.C. Nelson, R. Phillips, J. Widom, P.A. Wiggins, Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity, *Biopolymers*, **85**, 115, 2007.

[43] J.L. Carrascosa, X. Agirrezabala, J.A. Velazquez-Muriel, P. Gomez-Puertas, S.H.W. Scheres, J.M. Carazo, Quasi-atomic model of bacteriophage T7 procapsid shell: Insights into the structure and evolution of a basic fold, *Structure*, **15**, 461, 2007.

[44] W. Jiang, J. Chang, J. Jakana, P. Weigele, J. King, W. Chiu, Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus, *Nature*, **439**, 612, 2006.

[45] A. Fokine, P. R. Chipman, P. G. Leiman, V. V. Mesyanzhinov, V. B. Rao, M. G. Rossmann, Molecular architecture of the prolate head of bacteriophage T4, *PNAS*, **101**, 6003, 2004.

[46] N.V. Hud, K.H. Downing, *PNAS*, Cryoelectron microscopy of lambda phage DNA condensates in vitreous ice: The fine structure of DNA toroids, **98**, 14925, 2001.

[47] J.P. Wolf and C. Song, Finite-element modelling of unbounded media, *Wiley*, England, 1996.

[48] A. Lazarus, HCB. Florijin and PM. Reis, Geometry-induced rigidity in non-spherical pressurized elastic shells, *Phys. Rev. Lett.*, **109**, 144301, 2012.

[49] D. Vella, A. Ajdari, A. Vaziri and A.Boudaoud, The indentation of pressurized elastic shells: from polymeric capsules to yeast cells, *J. R. Soc. Interface*, **9** 448-455, 2012

[50] D. Vella, A. Ajdari, A. Vaziri and A.Boudaoud, Indentation of ellipsoidal and cylindrical elastic shells, *Phys. Rev. Lett.*, **109**, 144302, 2012.

[51] J. Kindt, S. Tzlil, A. Ben-Shaul, W. M. Gelbart, DNA packaging and ejection forces in bacteriophage, *PNAS*, **98**, 13671, 2001.

[52] SC. Riemer, VA. Bloomfield, Packaging of DNA in Bacteriophage heads, *Biopolymers*, **17**, 3, 785-794, 1978.

[53] T. Odijk, Statics and dynamics of condensed DNA within phages and globules, *Philosophical Transactions of the Royal Society of London*, **362**, 1820, 1497-1517, 2004.

[54] B.A. Todd, V.A. Parsegian, A. Shirahata, T.J. Thomas, and D.C. Rau, Attractive Forces between Cation Condensed DNA Double Helices *Biophys. J.*, **94**, 4775, 2008.

[55] P. K. Purohit, J. Kondev, R. Phillips, Mechanics of DNA packaging in viruses, *PNAS*, **100**, 3173, 2003.

[56] S.C. Harvey, and A.S. Petrov, Structural and Thermodynamic Principles of Viral Packaging, *Structure*,**15**, 21, 2007.

[57] A. Cordova, M. Deserno, WM. Gelbart and A. Ben-Shaul, Osmotic shock and the strength of viral capsids, *Biophys. J.*, **85**, 70-74, 2003.

[58] P. Grayson, A. Evilevitch, M.M. Inamdar, P.K. Purohit, W.M. Gelbart, C.M. Knobler, and R. Phillips, The effect of genome length on ejection forces in bacteriophage lambda, *Virology* , **348**, 430-6, 2006.

[59] S. Koster, A. Evilevitch, M. Jeembaeva, and D.A. Weitz, Influence of internal capsid pressure on viral infection by phage lambda, *Biophys. J.*, **97**, 1525-1529, 2009.

[60] W.H. Roos, M.M. Gibbons, A. Arkhipov, C. Uetrecht, N.R. Watts, P.T. Wingfield, A.C. Steven, A.J.R. Heck, K. Schulten, W.S. Klug, G.J.L. Wuite, Squeezing Protein Shells: How Continuum Elastic Models, Molecular Dynamics Simulations, and Experiments Coalesce at the Nanoscale, *Biophys. J.*, **99**, 1175, 2010.

[61] S.C. Harvey, and A.S. Petrov, Packaging Double-Helical DNA into Viral Capsids: Structures, Forces, and Energetics, *Biophys. J.*, **95**, 497, 2008.

[62] J. S. Villarrubia, Algorithms for scanned probe microscope image simulation, surface reconstruction, and tip estimation, *J. Res. Natl. Inst. Standards Technol.*, **102**, 425, 1997.

[63] I. Horcas, R. Fernandez, J. M. Gomez-Rodriguez, J. Colchero, J. Gomez-Herrero, and A. M. Baro, Wsxm: A software for scanning probe microscopy and a tool for nanotechnology, *Rev. Sci. Instrum.*, **78**, 013705, 2007.

[64] P.J. de Pablo, I.A.T. Schaap, F.C. MacKintosh, and C.F. Schmidt, Deformation and collapse of microtubules on the nanometer scale, *Phys. Rev. Lett.*, **91**, 98101, 2003.

[65] W.S. Klug, R.F. Bruinsma, J.P. Michel, C.M. Knobler, I.L. Ivanovska, C.F. Schmidt, and G.J.L. Wuite, Failure of viral shells, *Phys. Rev. Lett.*, **97**, 228101, 2006.

[66] G.A. Vliegenthart and G. Gompper, Mechanical deformation of spherical viruses with icosahedral symmetry, *Biophys. J.*, **91**, 834-41, 2006.

[67] M. Zink and H. Grubmuller, Mechanical properties of the icosahedral shell of southern bean mosaic virus: A molecular dynamics study, *Biophys. J.*, **96**, 1350-1363, 2009.

CHAPTER 10

Virtual Atomic Force Microscopy

10.1 Introduction

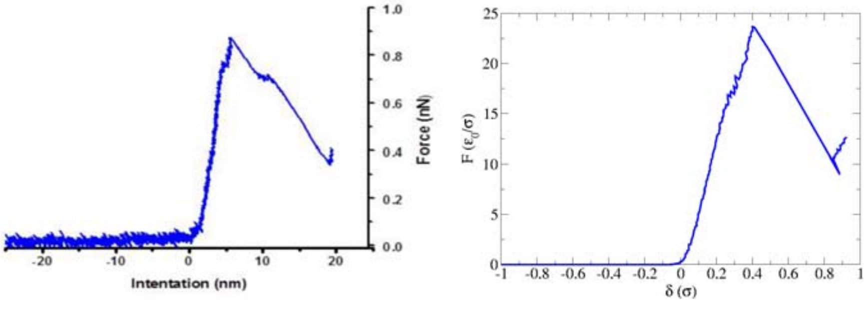
As mentioned in previous chapters, the most important role of the viral capsid is to protect the genetic material of the virus. During the extra cellular phase viruses face large changes in the environmental conditions [1], that could induce extreme changes in temperature, pH, osmotic shocks related to a sudden change of salt concentration, or dehydration. In addition, as shown in Chapter 9 the capsid of many dsDNA viruses has to withstand up to tens of atmospheres of pressure built up during the packaging of the viral genome at high densities [2–4]. Moreover, some viruses called extremophiles live in extreme conditions of salinity, radiation or temperature [5, 6]. Accordingly, it is very important that the capsid keeps its integrity under environmental changes and survives mechanical stresses in its cycle life.

As a consequence viruses have developed amazing mechanical properties. Notable examples are bacteriophage $\phi 29$ that, as shown in the previous chapter, is capable of withstanding about 40 to 60 atms of pressure inside [7]; or Cowpea Chlorotic Mottle Virus (CCMV) that is able to maintain its shape and structure in very broad pH ranges and after deformations of more than 30% without breaking [8].

In recent years, there has been a lot of interest in characterizing the mechanical properties of viruses. These studies have been possible by the development and application of different experimental techniques that have been able to extract precise mechanical properties of individual viruses. In particular optical tweezers were very important in the measurement of the force associated to the packing the genome [7]. But, arguably the most important and popular technique has been the application of atomic force microscope (AFM) to obtain the mechanical resistance and topography of single capsids [9, 10, 12]. Specially, AFM nanoindentation experiments have determined different mechanical properties such as the effective Young's modulus, breaking face, fatigue or even internal pressures [8, 11, 12].

AFM experiments are a very powerful tool to access interesting mechanical information of viral capsids, but sometimes it is very hard to get a good interpretation of the result and to relate this information to the biological cycle. In this context, theoretical modeling and simulations are required to interpret and understand better experimental AFM results [13].

The goal of this chapter is precisely to develop a "virtual AFM", i.e. a simulation mimicking the standard setup and experimental protocol of AFM nanoindentation experiments. With the combination of experiments and simulation, we are in a position to get more complete information about the mechanical response of viral capsid and how it will be influenced by changes in the environmental conditions.



(a) (b)

Figure 10.1: (a) A typical force versus indentation curve obtained in AFM experiments for T7 proheads in the 3-fold direction. The initial linear response ends abruptly by a sudden drop in this case associated to the fracture of shell. (b) A similar curve obtained in our simulations for $T = 7$ shell.

This chapter is organized as follows. In Section 10.2 we will briefly summarize the basics of AFM nanoindentation experiments and the main results that have been obtained in the study of viral capsids. Section 10.3 reviews the essential of continuum elasticity theory that is customarily used in the interpretation of AFM experiments. Section 10.4 will be devoted to the description of the "virtual AFM", the BD simulations that we have implemented to mimic the experimental set up. We next will show how the elastic response of the virus changes with the simulations parameters. After that, in Section 10.5, we will focus our simulation in the study of bacteriophage T7 and the comparison with AFM experiments. Finally, in Section 10.6 we summarize our findings.

10.2 Atomic Force Microscopy on Virus

The application of AFM has become a revolution in the study of the mechanical properties of viruses. The AFM device permits an exquisite control of the height, position, and force in the study of individual viruses, to acquire nanometric-resolution images, and also to perform a physical characterization of the rigidity of the sample.

The AFM is a high-resolution imaging technique that allows to explore single capsids. The basic set up of an AFM, consist of a flat horizontal surface, that supports the sample and a small tip at the end of a cantilever, which touches the sample. The position of the tip is controlled using a piezolectric device that can move over the sample in the three directions by an applied electric voltage. With this set-up when the tip touches the sample, the cantilever suffers a deflection that is measured by a laser and a photodiode. From the deflection of the cantilever, one can get an accurate topographic image of the sample with nm resolution [14].

The deflection of the tip can be translated into a force with a resolution of about 10 pN, using the elastic constant of the cantilever, which can be determined using, for instance, Sader's method [15]. AFM can also be used in liquids, thus facilitating the study of biological samples in their natural environment.

As described in the previous chapter, in a typical nanoindentation experiment, the sample is immobilized on a flat hard substrate. The cantilever is then lowered progressively, recording the force and deflection as a function of the vertical position of the tip [12].

For viruses it is possible to distinguish two deformation regimes, corresponding to small and big indentations, see Fig.10.1. For small indentations a linear force-response can be observed that allows to determine the particle's spring constant. For bigger indentations one can observe mechanical failure of the capsid often associated to the presence of strong discontinuities that arise in the force-indentation curve, see Fig.10.1.

AFM nanoindentation experiments of viruses are providing an important set of data about their mechanical properties. The use of this method on viruses such as the phages ϕ29 [16] and λ [17], and Cowpea Chlorotic Mottle Virus (CCMV) [8], all nonlipid viruses, has revealed that they are mechanically robust structures and that they have interesting elastic properties. In particular, bacteriophages have been found to have high effective Young's modulus, comparable to that of hard plastics [18].

The AFM is a helpful tool also in order to estimate the internal pressure of viruses, and the role of the genome in this pressure, as shown in the previous chapter. Nanoindentation experiments comparing empty and full capsids have been performed on various viruses such as CCMV [8], MVM [18], phage λ [17] and HSV1 [19, 20]. Results on CCMV, MVM and phage λ reported stiffening and/or an increase in breaking force of the full capsid compared to an empty capsid. Interestingly, no difference in mechanical properties was observed for empty and DNA filled nuclear HSV1 capsids. These results triggered questions on how the presence of the genome modifies the mechanical properties of capsids.

In addition, there are different studies about the changes in the capsid through the maturation process, where indentation experiments are preformed on capsids before and after this process. Studies for moloney murine leukemia virus (MLV) [21] and human immunodeficiency virus (HIV) [22] show that their proheads are stiffer than their mature capsidds. AFM experiments for bacteriophage λ and HK97 show that Young's modulus and breaking force increase during maturation [23].

The mechanical information obtained by AFM experiments has traditionally been interpreted in most cases by continuum elasticity theory, as described in the following section.

10.3 Elastic theory of indentation experiments

Continuum elasticity theory provides a macroscopic description of viruses. In the thin shell approximation, viral capsids can be described in terms of stretching and bending deformations. The elastic response of the capsid in indentation experiments depends on the competition between these two contributions and on the strength of the applied force (or depth of the applied indentation). One can distinguish two deformation regimes, corresponding to small forces of indentation or for larger forces.

For small deformations the behavior of the force versus indentation is linear and reversible. In these conditions, the mechanical behavior of the cantilever can be approximated as an ideal spring whose elasticity is given by Hooke's law: $F = -k_c x$, where F is the restoring force, which is equal to the force of tip-sample interaction, k_c is the spring constant of the cantilever, and x is the deflection of the cantilever. The deformable sample can be considered like a spring in series with the cantilever. Using this approximation the effective elastic constant of the capsid is given by $k_{cap} = k_c \frac{S_g}{S_v - S_g}$, where S_g (nm/V) is the slope of the cantilever deflection on the glass (the substrate is considered as non-deformable) and S_v (nm/V) the slope of the cantilever deflection on the virus. In the small deformations regime, k_{cap} is calculated from a lineal fit of the force versus the indentation curve. The k_{cap} is related to the effective 3D Young modulus E as:

$$k_{cap} = \frac{2}{\sqrt{3(1-\mu^2)}} \frac{Eh^2}{R} \tag{10.1}$$

where μ is the Poisson ratio, R is the radius of the capsid, and h its thickness. The previous relation is exact for a thin spherical shell indented by a point force. For a thick shell, the same formula is used, replacing the prefactor $\frac{2}{\sqrt{3(1-\mu^2)}}$ by an effective coefficient which is close to 1.

For indentations δ larger than the shell thickness, inverse buckling occurs [24]. In this regime, the force is no longer linear, but rather goes as $F \approx \sqrt{\delta}$. This inverse buckling is expected to happen when $\delta > h$ and when the force exceeds a critical value $F_{inv} \approx \kappa/R$. Applying even larger deformations eventually causes irreversible changes in the shell structure associated to bond ruptures. Rupturing studies can thus provide helpful information about the molecular interactions between capsomers.

Some of the approximations involved in thin shell theory can be removed by the use of finite elements simulations to solve the complex 3D continuum elasticity equations. Using FEM, it has been possible to analyze the effect of the thickness of the capsid, to extract phenomenologically the effective Young's modulus of the capsid, and even to account for the effect of inhomogeneities using refined meshes constructed from the atomic maps of some viruses [25].

This continuum mechanics description of viruses works well as a first approximation and has provided very useful information of the elastic properties of viruses. However, a continuum description cannot capture the discrete nature of subunits forming the capsid, which, as we saw in Chapter 8 and in this chapter, are very important for viruses. In addition, recent AFM experiments on bacteriophage T7 show an anisotropic stiffness in the mechanical response of T7 capsids [26, 27], that cannot be described with a continuum theory. Therefore, accounting for the discrete and inhomogeneous nature of the capsid is crucial to properly understand its stability and mechanical response. That is the reason why we developed our "virtual AFM" described in the next section.

10.4 Virtual AFM

As we mentioned earlier, extracting the relevant information from AFM nanoindentation experiments is difficult. Thus, it is often necessary to combine experimental information with simulations in order to interpret correctly results. Different kinds of simulations, accounting for the discreteness of the capsid at diverse resolution levels, have been implemented in the literature [23, 28]. In particular, coarse-grained models have been implemented, to study the maturation pathway of viruses of different sizes and structures [29, 30]. Elastic network models have also been used to study the buckling process of viruses [28].

In this chapter, we have used our coarse-grained model described in Chapter 3, to implement a "virtual AFM". The main motivation for the implementation of the "virtual AFM" was to mimic and interpret recent experimental results on bacteriophage $T7$. However, the model can be used to study the mechanical properties of any virus of arbitrary T-number, thus being potentially very useful to get information on the stability and resistance of viruses.

10.4.1 Simulation set-up

In order to reproduce the nanoindentation experiments, we wanted to implement a simulation mimicking the typical AFM setup. Since we are interested in the response of the viral capsid and the dynamics of the capsomers during the

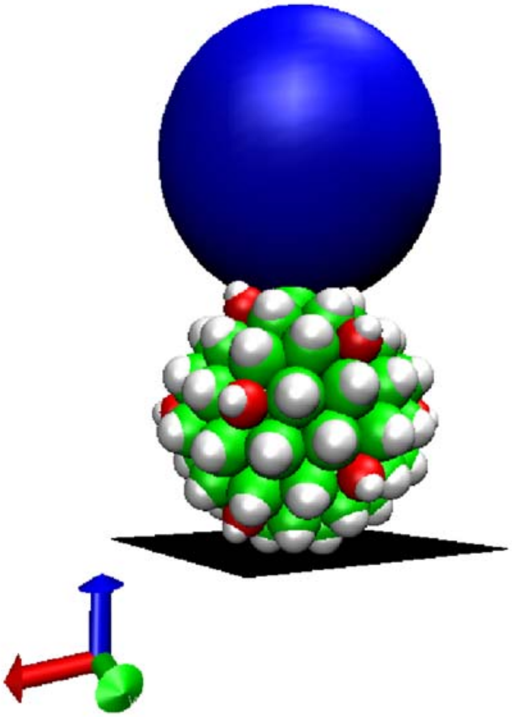


Figure 10.2: Snapshot of the indentation of a $T = 7$ coarse-grained shell model along the 2-fold direction. The blue sphere represents the AFM tip. The viral capsid is modeled at the level of capsomers, where hexamers and pentamers are represented by green and red spheres, respectively. The black rectangle illustrates the hard substrate.

indentation process, we decided to use an algorithm based on Brownian Dynamics simulations, as described in Chapter 5.

The simulation setup is shown in Fig.10.2. In these simulations we have added a surface at a position $z = 0$; a cantilever at the top, modeled as a sphere, which can only move in the z direction, and a preformed capsid between them to mimic the experiments. The capsomers, depicted by the red and green spheres in Fig.10.2, interact by pairs following the coarse-grained model described in Chapter 3. They also interact with the tip and with the surface. In the initial configuration, the capsid is placed sligthly above the surface and the position and orientation of all capsomers correspond to the optimal one for the parameters and T-number used.

The interaction between the capsomers and the sphere of radius d_c that represents the AFM tip (depicted at Fig.10.2 by a blue sphere in the top) is modeled by the following purely repulsive harmonic potential:

$$V_c = \frac{1}{2} k_{repul} \left(r_{ic} - (d_c + \sigma/2) \right)^2 \quad if \quad r_{ic} < d_c + \sigma/2 \tag{10.2}$$

where k_{rep} is an effective repulsive spring constant, r_{ic} the distance between the center of the cantilever and capsomer i, d_c the radius of the cantilever and σ is σ_p or σ_h for pentamer or hexamers, respectively.

To mimic the deflection of the cantilever caused by the interaction with the capsid, its z position is updated using the following overdamped Euler equation :

$$z_c = \frac{\Delta t}{\eta_c} (F_c - k_c(z_c - h)) + \xi \sqrt{2 D_c \Delta t} \tag{10.3}$$

where $\Delta t = 10^{-4}$ is the time step used in the simulations,
$F_c = -\sum_i k_{repul} \left(r_{ic} - (d_c + \sigma/2) \right)$ is the total force on the cantilever due to the interaction with all capsomers, k_c is the spring constant of the cantilever η_c is the friction of the cantilever, D_c is its diffusion coefficient, and ξ is a Gaussian noise with zero mean and variance 1. As in a real AFM, the height of the cantilever is controlled and fixed to a value h, that is lowered stepwise at intervals of 10^5 time steps.

Finally, the substrate is modeled as a rigid surface (Fig.10.2, black surface), with a purely repulsive harmonic interaction with stiffness k_{subs}. To account for the potential effects of adsorption with the substrate, as often occurs in experiments, a binding energy of adsorption E_{ads} is also considered. This adsorption contribution is descrited by the interaction potential

$$V_{sub}(r) = E_{ads} \left[\left(\frac{\sigma_i/2}{|z|} \right)^{12} - 2 \left(\frac{\sigma_i/2}{|z|} \right)^{6} \right] \tag{10.4}$$

where $\sigma_i/2$ is the radius of capsomer i (that can be a pentamer (σ_{pp}) or an hexamer (σ_{hh})) and $|z|$ is the distance between capsomer i and the substrate.

In experiments the capsid must be fixed in the correct symmetry on the surface. To prevent the rotation and sliding of the viruses in the simulations, the position of some capsomers is fixed, depending on the orientation of the virus as discussed in Appendix A.

In all simulations we have fixed some parameters: the radius of the cantilever $d_c = 3\sigma$, the spring constant of the cantilever $k_c = 40\epsilon_0/\sigma^2$, and the strength of the tip-capsomer repulsion to $k_{rep} = 80\epsilon_0/\sigma^2$ or $k_{rep} = 160\epsilon_0/\sigma^2$ (the behavior for both values is the same).

As in experiments, the simulation procedure involves the progressive lowering of the tip. In our simulations, the z position of the cantilever is given by h and it is lowered at steps of 0.005σ. We measure the average force on the cantilever sampled every 100 steps. We lower the tip typically after 10^6 steps, measuring. The simulation runs for a total of $3x10^8$ steps corresponding to a total indentation of 1.5σ. In some simulations, we also have simulated the retraction of the tip, to check the reversibility, using the same number of steps and increments.

With the "virtual AFM" it is possible the study of different capsid structures. But we will focus on the $T = 7$ structure as a particular example in order to compare the simulation results with AFM nanoindentation experiments of bacteriophage $T7$ which were performed by the group of P.J. de Pablo at the Universidad Autónoma de Madrid.

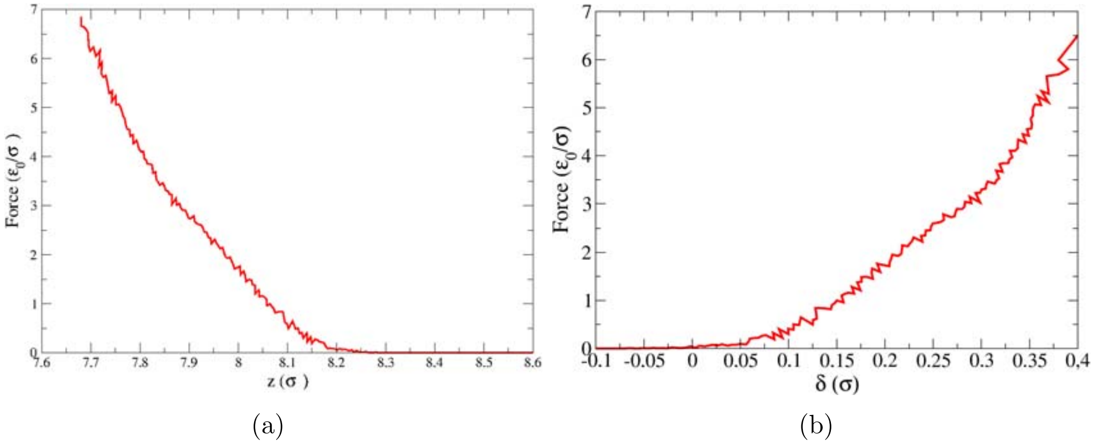
As we have mentioned previously, viral capsids might have different mechanical response depending on the absorption geometry, typically corresponding to indentations along the 5,3,2-fold symmetry axes. In order to study this anisotropy we have rotated the optimal $T = 7$ structures in different ways and fixed the positions of some capsomers, as described in Appendix A.

The CK $T = 7$ capsid is a chiral structure, that could be built with two different chiralities: laevo and dextro, corresponding to (h, k) vector $(2, 1)$ or $(1, 2)$, as discussed in Chapter 1. Bacteriophage T7 has a $T = 7$ laevo structure. Thus, we have used this in our simulations. In any case, we have also performed simulations for a $T = 7$ dextro shell, finding no significant difference in the results.

10.4.2 Nanoindentation curves

Fig.10.3 shows three repetitions of a typical nanoindentation curve obtained for 2-fold oriented $T = 7$ capsid including also the retraction of the tip. In Fig.10.3(a) we have plotted the force F versus the z-position of the cantilever. As in experiments, from the F versus z curves we can calculate the indentation as $\delta = z + F/k_c$, where z is the position of the cantilever.The indentation curves are also shifted in z in such a way that the tip-sample contact start at $\delta = 0$. The resulting force versus δ curves are plotted in Fig.10.3(b).

For small indentations, the F vs δ curves show a lineal and reversible behavior corresponding to elastic response (see black curve in Fig.10.3(b)). At large indentations, see Fig.10.4, the linear regime end up by an abrupt decline of the force, corresponding in this case to the sinking of the top pentamer, that gets inside the capsid, leaving a small hole in the position of the lost pentamer. A second abrupt decline in the force is observed when tip breaks the top part of the capsid generating a large hole. This cracking of the capsid is not reversible, as shown by the



(a) (b)

Figure 10.3: Force versus indentation in the 2-fold direction for a $T = 7$ structure with all capsomers having the same size size, $\alpha = 0.1$, $k_c = 40$, $k_{repul} = 80$ and $T = 0.05$.

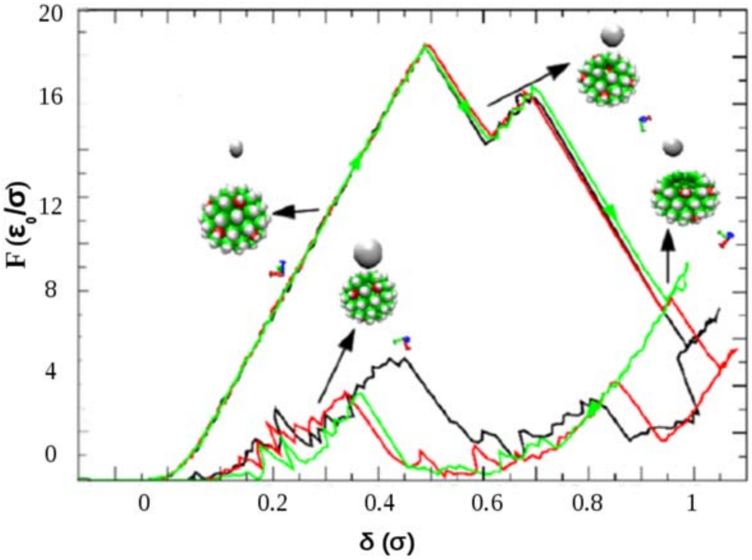


Figure 10.4: Force versus indentation in the 5-fold for a $T = 7$ structure with all capsomers having the same size, $\alpha = 0.1$, $\nu = 1.796$, $k_c = 40$, $\nu = 1.796$, $k_{repul} = 160$ and $T = 0.05$.

different routes followed by the retraction curves of the tip. The final structure after retraction sometimes is partially healed as shown in the inset figure of a 71 structure (missing one pentamer) with a small hole and pentamers and hexamers reorganized.

It is worth emphasizing that the indentation curves obtained using our "virtual AFM" are remarkably and qualitatively nearly identical to the experimental

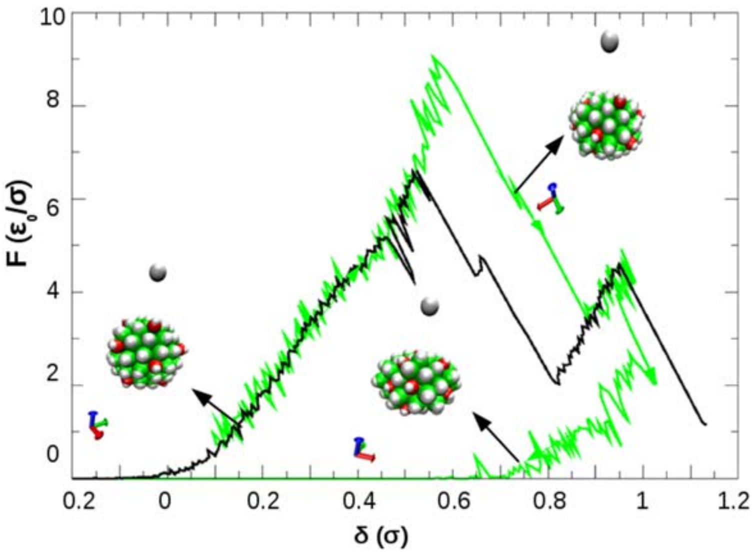


Figure 10.5: Force versus indentation in the 3-fold direction for a $T = 7$ structure with all capsomers having the same size, $\alpha = 0.1$, $\nu = 1.796$, $k_c = 40$, $k_{repul} = 80$, $T = 0.05$ and $E_{ads} = 0.1$.

curves measured with the real AFM, as illustrated in the comparison between the simulated indentation curves (Fig.10.4 and Fig.10.5) and the experimental ones, see Fig.10.1.

The indentation curves obtained for a $T = 7$ capsid along the 3-fold and 2-fold axis are qualitatively similar to the 5-fold one, see for instance Fig. 10.5 for the 3-fold. The figures show an initial linear regime ending by one or several abrupt drops in the force that are indicative of the loss of capsomers, the partial buckling, or the breakage of the capsid, as will be discussed in more detail in section 10.4.3. Note that the critical force associated to the sudden drop of the indentation curves changes in different repetitions of the same simulation. This is indicative of the stochastic, temperature-activated nature of this phenomenon.

We analyzed the influence on the mechanical response of the different parameters of the simulation, such as the bending stiffness, related to the parameter α (see Chapter 3), the temperature, the angle ν related to the spontaneous curvature, the size of the pentamer compared to the size of hexamers, the adsorption energy of the surface E_{ads}, and the relative strength of the binding energy between hexamers and pentamers. The results are briefly summarized in the following subsections and in Table 10.2 of Appendix B.

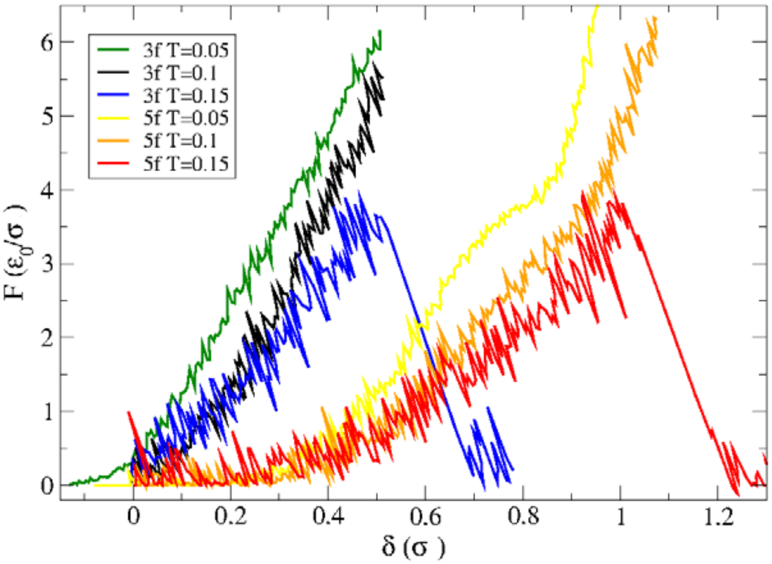


Figure 10.6: Force versus indentation in the 3-fold and 5-fold orientations for a $T = 7$ structure with all capsomers having the same size, $\nu = 1.796$, $k_c = 40$, $k_{repul} = 80$, $E_{ads} = 0.1$ and $\alpha = 0.7$ for different temperatures.

A. Influence of the temperature T

Fig.10.6 plots the indentation curves for a $T = 7$ structure indented along the 5-fold and 3-fold axes simulated at different values of the temperature T. The main influence of T is that, as we increase it, the structure tolerates smaller indentations and the curve becomes noisier. That is the expected behavior, since increasing the temperature entrances thermal fluctuations and renders the structure more unstable. The initial slope of the indentation curves is not significantly changed. The influence of the temperature is qualitatively similar for all orientations.

B. Influence of the adsorption energy E_{ads}

We have analyzed the influence of adsorption strength on the indentation curves. Fig.10.7 shows the results for $E_{ads} = 0.1, 0.5$ and 1.0, that indicate a negligible influence for small E_{ads}. At intermediate values of E_{ads}, the indentation of the capsid eventually leads to the flattering of its lower surface that gets adsorbed on the substrate. This is indicated by the plateau in the force in the curve corresponding to $E_{ads} = 0.5$ in Fig. 10.7. As E_{ads} becomes bigger, the capsid gets strongly adsorbed on the substrate right from the start of the simulation, as indicated by the fact that the initial contact between the tip and the capsid occurs at lower values of z. At very large E_{ads}, the shell eventually flattns on top of the substrate. To avoid a strong effect of the substrate on the indentation curves, we have fixed $E_{ads} = 0.1$ in all the simulations.

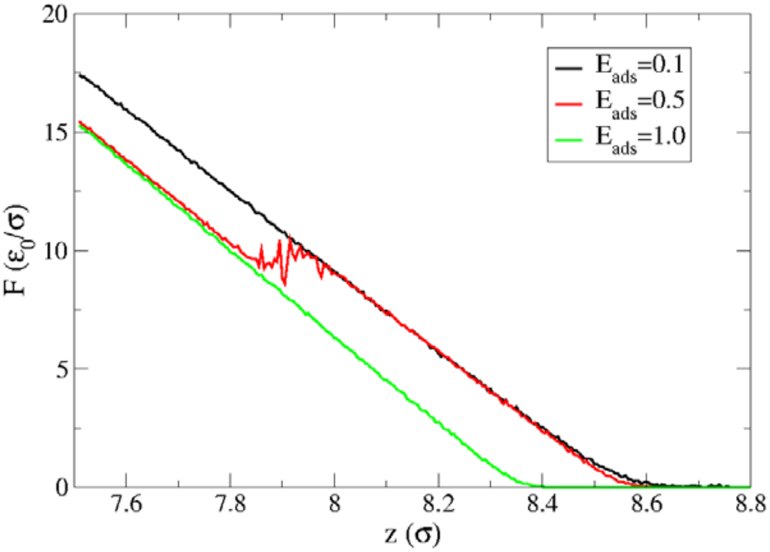


Figure 10.7: Force versus height of the cantilever in the 5-fold direction for a $T = 7$ structure using all capsomers of the same size, $\nu = 1.796$, $k_c = 40$, $k_{repul} = 80$, $T = 0.05$, $\alpha = 0.7$, and different E_{ads}.

C. Influence of ν

Fig.10.8 represents the influence of ν on the indentation curve in the 5-fold orientation. For the optimal angle ν_{opt}, the shells resist larger indentations and effective spring constant is comparatively small. In all cases where ν is different from the optimum the structure is either stretched (for $\nu < \nu_0$) or compressed (for $\nu > \nu_{opt}$), specially when all capsomers have the same size. As a consequence, the effective spring constant is higher and the structure in every direction tolerates smaller forces and indentations, breaking earlier than structures with the optimum angle. The general behavior in terms of the preference angle ν is the same for all orientations.

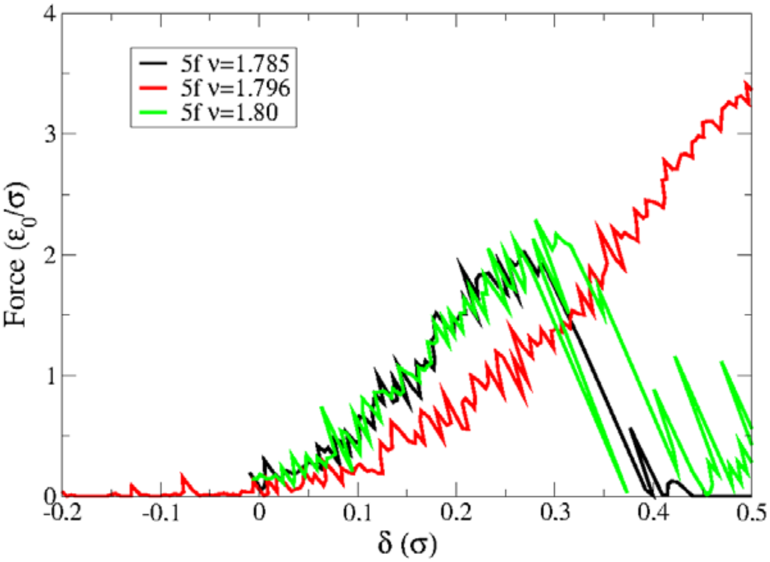


Figure 10.8: Force versus indentation in the 5-fold orientation for a $T = 7$ structure using with all capsomers having the same size, $T = 0.05$, $k_c = 40$, $k_{repul} = 80$, $E_{ads} = 0.1$ and $\alpha = 0.7$ for different ν.

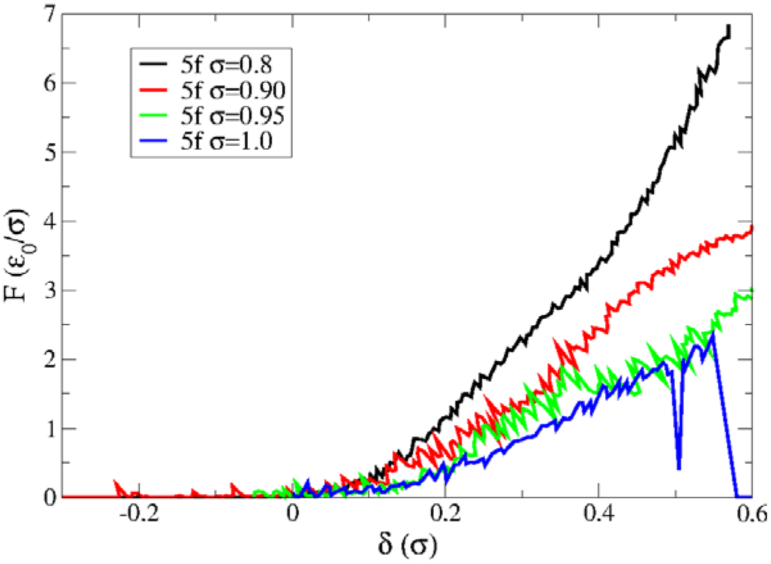


Figure 10.9: Force versus indentation in the 5-fold orientation for a $T = 7$ structure using different relative pentamer sizes, $k_c = 40$, $k_{repul} = 80$, $E_{ads} = 0.1$, $\nu = 1.796$ and $\alpha = 0.7$.

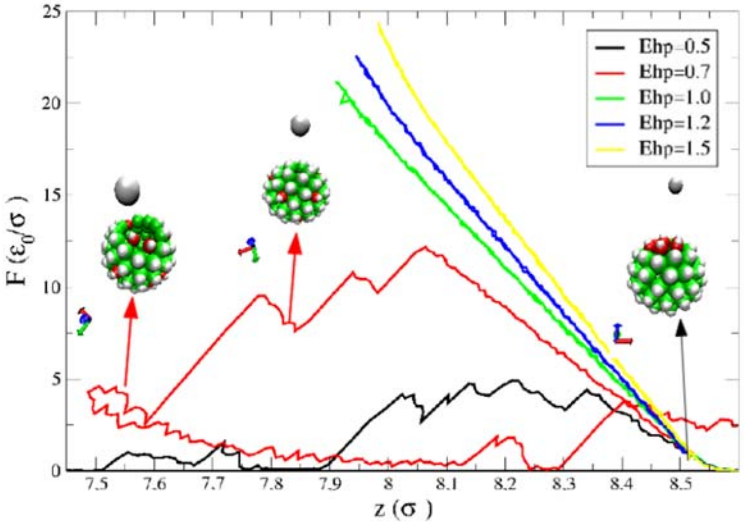


Figure 10.10: Force over the cantilever in the 5-fold direction versus the cantilever position for a $T = 7$ structure with all capsomers having the same size using $\nu = 1.796$, $\alpha = 0.7$, $k_c = 40$, $T = 0.05$ and $k_{repul} = 80$,for different E_{hp} values.

D. Influence of the relative size of pentamers

We also studied the influence of the relative size of the pentamers, see Fig. 10.9. To that end we implemented simulations with the follows sizes, $\sigma_{pp} = 0.8\sigma_{hh}$, $\sigma_{pp} = 0.9\sigma_{hh}$, $\sigma_{pp} = 0.95\sigma_{hh}$ and $\sigma_{pp} = \sigma_{hh}$. The case where all subunits have the same size ($\sigma_{pp} = \sigma_{hh}$) is more frustrated than structures with pentamers smaller than hexamers and accumulates larger stresses. That results in a structure that is softer in all directions and tolerates smaller forces and indentations, see also Appendix B.

E. Influence of E_{hp}

Now we want to study the role of different binding energies between capsomers. For a $T = 7$ structure where pentamers are not in direct contact with other pentamers, the relevant parameter to change is the pentamer-hexamer binding energy (E_{hp}) compared to the hexamer-hexamer one $E_{hh} = \epsilon_0$. We have changed E_{hp} in the range $E_{hh} = 0.5, 0.7, 1.0, 1.2$ and 1.5 in reduced units of ϵ_0. We have implemented simulations for different values of α and size of pentamers for the three folds. In all the cases we have obtained a similar behavior, which is illustrated in Fig.10.10 for a 5-fold example. When $E_{hp} > E_{hh}$, the shell deforms linearly and does not break in the range of indentations explored. The larger the value of E_{hp}, the higher is the slope, implying a stiffer structure. As E_{hp} is reduced below E_{hh}, the pentamers are less stable and the shell breaks at smaller indentations the lower the value of E_{hp}. In the force curves, see Fig.10.10, the first drop corresponds to

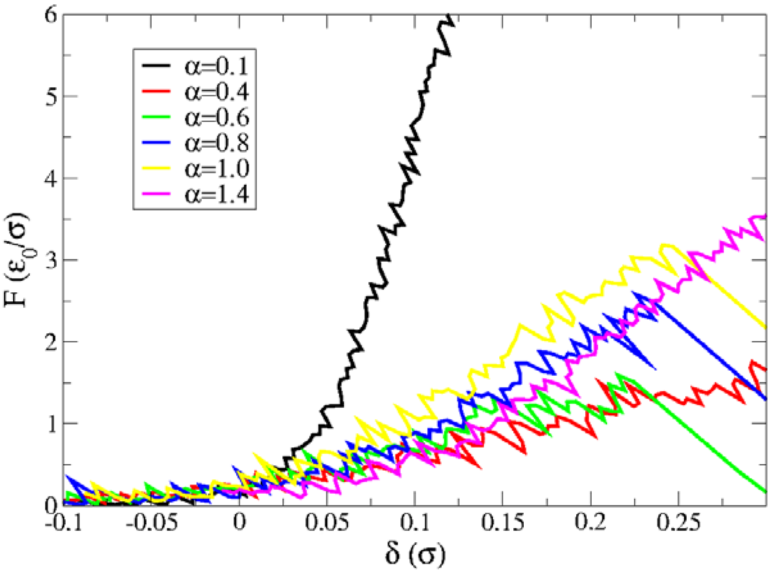


Figure 10.11: Force versus indentation in the 5-fold orientation for a $T = 7$ structure using all capsomer of the same size, $\nu = 1.796$, $k_c = 40$, $k_{repul} = 80$, $T = 0.05$ and $E_{ads} = 0.1$, for different values of α.

the sinking of the bottom pentamer inside the capsid. The second drop marks the entrance of the top capsomer under the tip inside the capsid. The third decline in the force is when the capsid cannot resist more indentation and breaks down. After the breaking and the retraction of the tip the pentamers and hexamers are reorganized and cluster together, since the stronger interaction between hexamers favours their clustering.

F. Influence of α

As we saw in Chapter 3 the parameter α is related to the bending stiffness. Simulations show the following the influence of the parameter α on the indentation along the 2-fold and 3-fold axis. As α decreases, the structure becomes stiffer and tolerates larger indentations. However, the 5-fold behavior under changes in the bending stiffness is different. Fig.10.11 shows the influence of α in the indentation of a 5-fold $T = 7$ structure with one type of capsomer. In the regions $\alpha \in [0.1, 0.5]$ and $\alpha \in [1.0, 1.4]$ the structure becomes softer when α is increased as happens for the 2 and 3-fold cases, see red, black and violet lines in Fig.10.11. In the region $\alpha \in [0.5, 1.0]$ the structure surprisingly becomes stiffer when α increases, see blue and yellow lines in Fig.10.11.

This is related to the occurrence of a buckling transition in the structures as discussed in Section 6.3.3. The values of the spring constant, maximum forces and indentations in each case are listed in Appendix B.

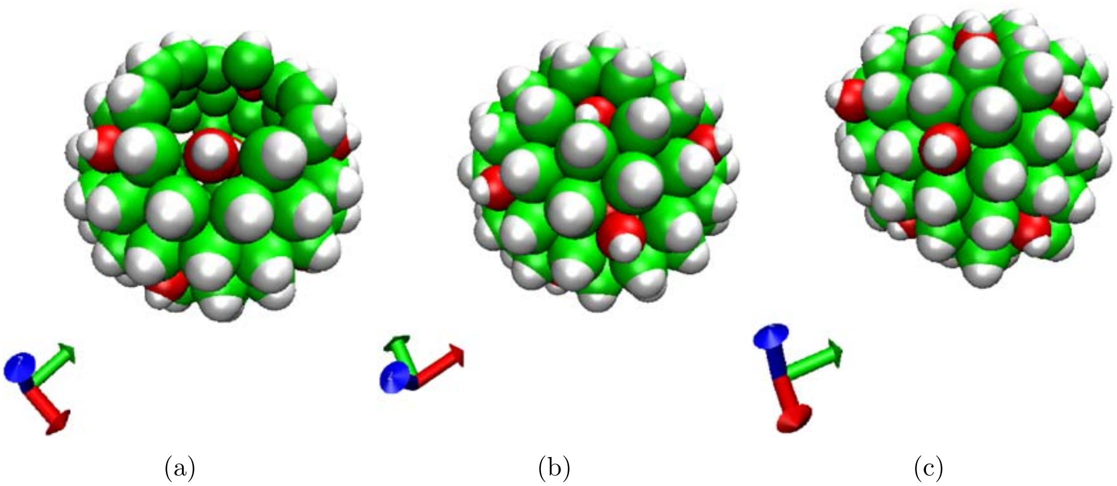

(a) (b) (c)

Figure 10.12: Snapshot corresponding to a drop in force for (a) an example of breaking for large indentations along the 3-fold direction and two examples of buckling: (b) in 5-fold direction, due to the entrance of the top pentamer and (c) in the 3-fold direction, associated to the faceting of the shell.

10.4.3 Buckling and breaking of viral capsid

One of the most representative characteristics of both experimental and simulated indentation curves are abrupt drops in the force, that occur after the initial linear regime. These force drops have been traditionally attributed to breaking, see Fig.10.1(a). However, in the simulations we see that many different physical phenomena can happen with the same signature corresponding to a drop. The advantage of the simulations is that we can correlate the observed drops in the force with snapshots obtained from the simulations. In that way, we have observed that the sudden drops in the force correspond, very often in an indistinguishable way, to breaking, buckling, rotation and/or sliding. We have obtained rotation and sliding in the case when pentamers were not fixed in the simulations. However, in the case where pentamers are fixed we could see breaking, buckling and the reorganization of capsomers, see Fig.10.12.

It is important to distinguish between breaking and buckling events. *Breaking* is defined as the disruption of the capsid involving rupture of intercapsomer bonds. This is typically an irreversible phenomenon that is associated to the presence of a crack or hole in the capsid (see Fig.10.12(a)), as can be observed also in AFM experiments. The breaking is important for the life cycle of virus, since it will expose the genome. In nano applications it is indispensable to prevent the unwanted breaking of the shell for nanoencapsulation.

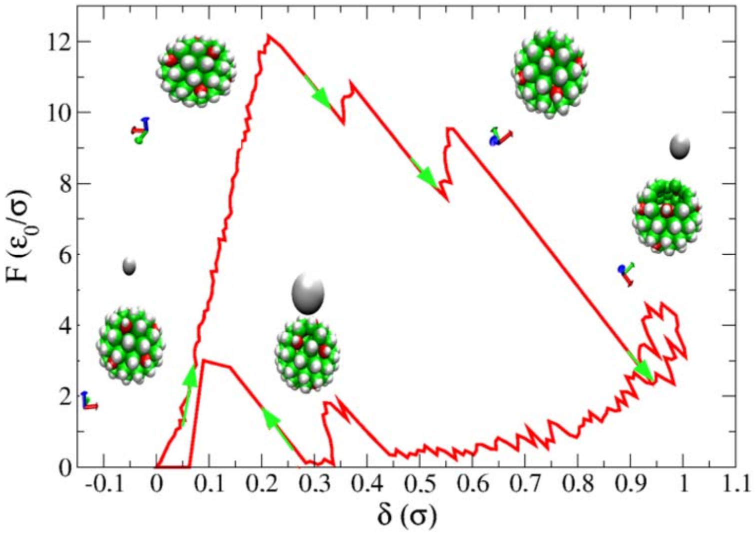


Figure 10.13: Force over the cantilver in the 5-fold direction versus the indentation for a $T = 7$ structure using $k_c = 40$, $k_{repul} = 80$, and $E_{hp} = 0.7$, showing different examples of mechanism, which it corresponds with falls in the force.

In the *buckling process* there is a collective change in the position of several capsomers, leading to a faceting of the structure, without cracks, see Fig.10.12(b),(c). If the indentation stops just after a buckling event and the tip is retracted we could observe very often that this buckling is reversible.

The goal of this section is to characterize a bit better these phenomena and how they depend on different factors, such as bending stiffness, spontaneous curvature or binding interaction. This information is potentially useful to tune the mechanical resistance of a capsid for different applications or to promote its disruption.

A good example of different drops in the force under indentation is plotted in Fig.10.13. This figure show a first a buckling process. The second drop corresponding to the collapse of the bottom pentamer, that moves at the same z-position that the hexamers around it and the third drop is when the structure actually breaks.

In the experimental force curves the mechanism involved in the drop is indistinguishable in most cases. It is usual to interpret always a drop in the force as a breaking event. But with the help of our simulation we can see that this is not always correct, and that the mechanism depends of the parameters. When the spontaneous curvature is not the optimum one or the bending stiffness is small we have obtained more often breaking. Buckling appears normally for large values of the bending stiffness.

In the study of the influence of the mechanical response of the viral capsid with the different axes of symmetry we have characterized the system in terms of the maximum indentation, δ^{max}, and the maximum “breaking force” F^{max}. The maximum breaking force is calculated as the average force associated to the first sudden drop in the indentation curve. The results for the different combination of parameters are summarized in Table 10.2 of Appendix B.

In general, δ^{max} and F^{max} are similar for the 2-fold and 3-fold axis. On the other hand, in the 5-fold direction, depending on the simulation parameters, the capsid tolerates a bit more or less indentation.

The δ^{max} and F^{max} for different absorption symmetries changes mainly with α, E_{hp}, and with the size of pentamers, see Table 10.2. For small values of α, for example $\alpha = 0.1$ the maximum indentation for 2-fold and 3-fold are around the same value with the 5-fold breaking earlier. In the case of large α the behavior is different. For example changing from $\alpha = 0.1$ to $\alpha = 1.0$ we have obtained $(\delta^{max}, F^{max}) = (0.6\sigma, 8\epsilon_0/\sigma)$ for 2,3-fold while for the 5-fold we obtained $(\delta^{max}, F^{max}) = (0.7\sigma, 6\epsilon_0\sigma)$. So the 5-fold orientation resists larger indentations but smaller forces.

In terms of the relative size of pentamers when all subunits have nearly the same size the structure tolerate less indentation because it accumulates more stress, especially in the 5-fold orientation. δ^{max} and F^{max} become progressive smaller as the size of the pentamers is increased, see Table 10.1, violet row. We have obtained a similar behavior upon changes in E_{hp} than in the pentamers size. For large values of E_{hp} the structure tolerates more indentation and the maximum force is bigger. As an example the values of (δ^{max}, F^{max}) for $E_{hp} = 1.0$ and $E_{hp} = 1.5$, for the 2,3,5-fold direction are listed in the orange rows of the Table 10.1.

10.4.4 Polyhedral vs spherical shells

As discussed in Chapter 8, bacteriophage T7 as well as many other viruses, change their shape from spherical to polyhedral during the maturation of their capsids [27]. This change in viral shape is present in the maturation of others viruses such as λ or HK97 [23, 31] and it is indispensable to become infective.

The change in shape might alter the mechanical response of the shell and could bring some mechanical advantage. Previous works in viruses like λ and HK97, which suffer a transition between spherical and polyhedral shapes during their maturation, show mechanical changes. In particular, λ mature virion is more resistant than its prohead [23, 32] or HK97 virion is stiffer than its prohead, but the prohead tolerates larger deformations [23].

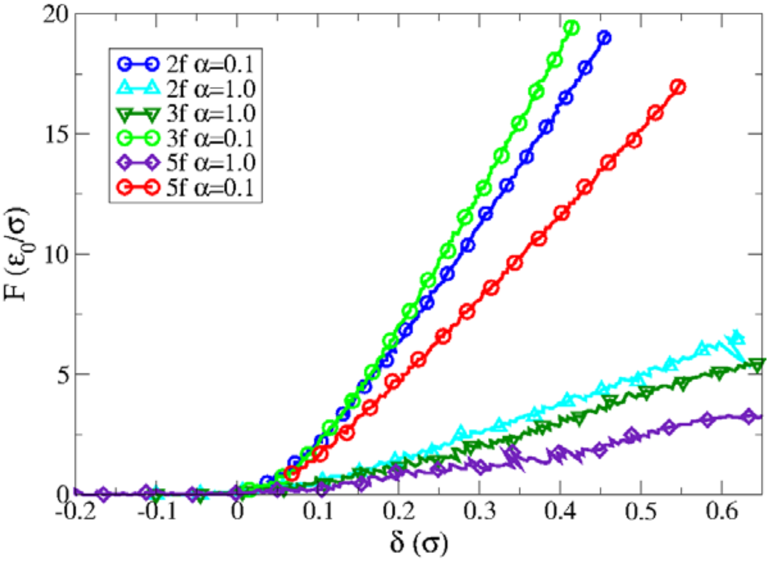


Figure 10.14: Indentation curves for a $T = 7$ structure for $k_c = 40$, $k_{repul} = 80$ $E_{hp} = 1.4$ and $\sigma_{pp} = \sigma_{hh}$ for the different orientations for a spherical capsid with $\alpha = 0.1$ (circles) and for a polyhedral shells with $\alpha = 1.0$ (triangles).

In this section we will compare the mechanical response of our model $T = 7$ structure in its spherical and polyhedral shapes. Using our model and the "virtual AFM" it is possible to obtain qualitative information about the mechanical changes upon maturation for $T = 7$ structures. This information is important not only for bacteriophage T7, but also for many other $T = 7$ structures, which undergo this maturation shape transition during their assembly [23].

In our model, the change in shape is obtained by changing the α parameter related to the bending stiffness (see Chapter 6). In particular we saw in Chapter 6 that $T = 7$ structures with one type of subunit exhibit a transition between the spherical shape and the polyhedral shape at $\alpha \approx 0.6$. The structure become progressively more and more faceted as α is increased. We have chosen the values of $\alpha = 0.1$ to represent the spherical prohead and $\alpha = 1.0$ for the mature capsid because for this value the structure is noticeable faceted, but stable enough to sustain AFM indentations.

As an example, we will analyze the case of bacteriophage T7 with one type of morphological subunit and $E_{hp} = 1.4E_{hh}$. The value $E_{hp} = 1.4E_{hh}$ has been chosen to reproduce the relative strength of the hexamer-hexamer and hexamer-pentamer interactions reported from the atomic structure in Ref. [26]. In particular, the hexamer-pentamer relative binding energy is calculated as the ratio between the binding energy of contact (A-G) plus contact (F-G), see Fig.10.15. The hexamer-hexamer binding energy is calculated as the average contact energies of the three hexamers: $E_{hh} = \frac{1}{3}(E_{12} + E_{13} + E_{23}) = 3.17$. Accordingly, $\frac{E_{hp}}{E_{hh}} = \frac{4.60}{3.17} = 1.4$.

Fig.10.14 shows the force versus indentation curves obtained for the model "prohead" (circles) and "mature capsid" (triangles) along the 2-fold, 3-fold and 5-fold symmetry axes. In general the spherical shape, for all folds, is stiffer than the polyhedral shape.

Interestingly, we have calculated the spring constants for both shapes and observed a change in their relative ordering. Specifically, we obtained $k_3^{sp} = 55.6, k_2^{sp} = 48.2, k_5^{sp} = 34.2 \pm 0.5\epsilon_0\sigma_0^2$. Thus, the 3-fold is the stiffest and the 5-fold is the softest. In contrast, for the polyhedral shape we obtained $k_2^{ico} = 11.6, k_3^{ico} = 9.9, k_5^{ico} = 5.8 \pm 0.5\epsilon_0\sigma_0^2$, and thus the softest fold is the 5-fold, and the stiffer is the 2-fold, yielding an order of constants $k_2 > k_3 > k_5$.

The spring constants obtained in experiments are similar for the prohead and virion in all symmetry axes, contrarily to our simulations. This discrepancy is mainly due to the fact that the increase in α required to obtain the polyhedral shape also modifies significantly the global stiffness of the shell. Thus, in its present form, our model cannot capture realistically the complicated changes associated to viral maturation.However, the order of the spring constant for our polyhedral shell is the same as that found in nanoindentation experiments for the T7 as described in the following section.

10.5 Anisotropic elasticity of T7

Recent experiments by the grousp of P.J. de Pablo and J.L. Carrascosa have analyzed the mechanical properties of bacteriophage T7 using AFM and chemical stress [33]. Bacteriophage T7 belongs to the Podoviridae bacteriophage family, characterized by viruses with an icosahedral capsid and a short non-contractile tail. T7 first assembles into the prohead, which is a shell built by 415 copies of the structural protein gp10A arranged in a $T = 7$ icosahedral lattice. The protein are distributed in clusters of 6 (hexons) and 5 (pentons) proteins. The T7 capsid maturation process involves the packaging of 40 kb of double-stranded (ds) DNA into the prohead, the release of the scaffold, the expansion of the shell, and the incorporation of the tail proteins in the vertex of the connector [33]. The maturation process of bacteriophage T7 involves an increase of the diameter of the particles, from $51nm$ to $60nm$, and a thinning of the shell from 4.1 to $2.3nm$ without increase the stiffness [33]. These drastic changes are based on the extensive reorganization of the interactions among the shell protein subunits, both within the capsomers and between them. This maturation process increases the structural stability of bacteriophage T7 against chemical treatments. This increased endurance would provide bacteriophages with the capability of standing in the hostile conditions during their extracellular cycle, where they have to face different temperatures, humidity and general environmental hazards.

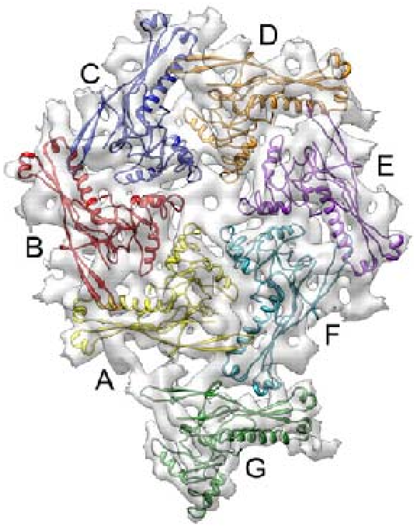


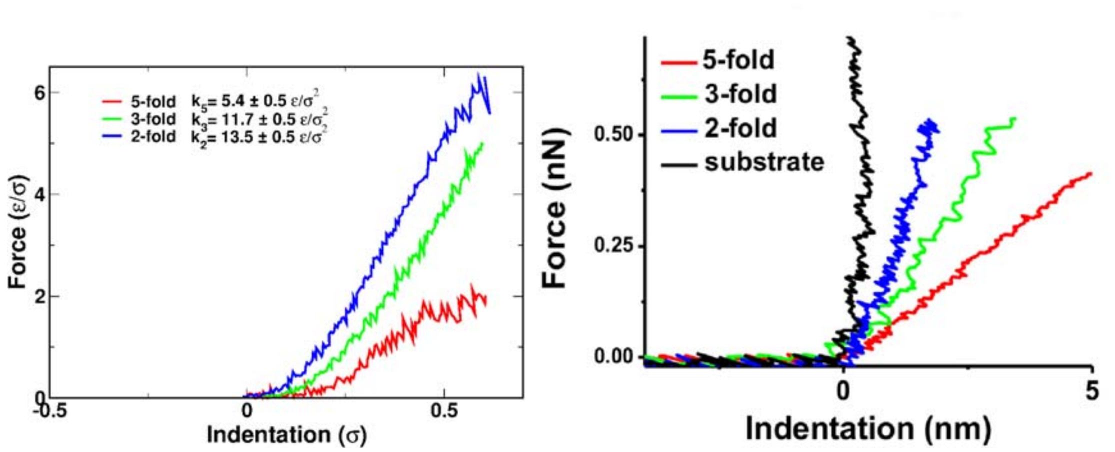
Figure 10.15: Pseudo-atomic model of the mature T7 shell protein, where the unit contains an entire hexamer (subunits A-F) and one pentamer subunit (subunit G) [26].

The initial mechanical characterization consisted on nanoindentation experiments on both proheads and capsids keeping the maximum force low enough to avoid nonlinearities and steps. From the linear regime, the effective spring constant as a function of the different geometrical orientations, including non-specific orientation particles, was extracted, and the results are summarized in Fig.10.16(b).

The spring constants, measured for procapsids were $k_{3p} = 0.24 \pm 0.04 N/m$ and $k_{5p} = 0.10 \pm 0.02 N/m$. On the other hand, experiments on empty mature capsids show capsid spring constants of $k_{2c} = 0.40 \pm 0.02 N/m$, $k_{3c} = 0.21 \pm 0.03 N/m$ and $k_{5c} = 0.11 \pm 0.02 N/m$. These results not only demonstrate an anisotropic elasticity for both T7 proheads and capsids, but also that the T7 bacteriophage virus rigidity does not increase upon maturation.

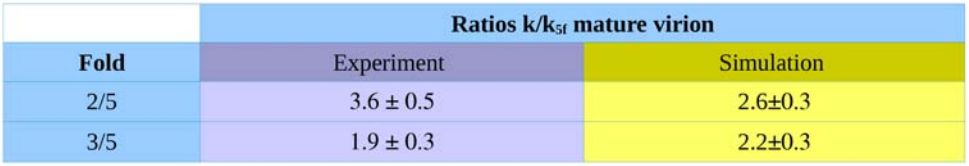
Interestingly, we have found that the anisotropic elasticity obtained from AFM indentation experiments of T7 viral capsids does not follow the anisotropy expected from continuum homogenous models [17]. Whereas the experimental results show the spring constant of T7 capsids for each icosahedral symmetry axis as $k_{5c} < k_{3c} < k_{2c}$, FEM simulations of the stiffness performed in thin shell continuum models establishes $k_5 < k_2 < k_3$ along with the increased thickness [34]. In addition, Finite Element simulations of a thin or thick icosahedral shell with the same dimensions as T7 also predicts an incorrect order of the spring constants. It is clear that it is necessary to use an inhomogeneous model to characterize correctly the mechanics of this virus.

To corroborate that discrete effects dominate the mechanical response of the



(a) (b)

Figure 10.16: (a) Force exerted by the tip as a function of indentation for a $T = 7$ structure using $k_c = 40$, $k_{repul} = 80$, $\alpha = 0.7$, $\sigma_{pp} = \sigma_{hh}$ and $E_{hp} = 1$, when the model capsid is indented along the different orientations: 5-fold (in red), 3-fold (in green) and 2-fold (in blue). The slope of a linear fit of each curve yields the values of the effective spring constants indicated in the legend. (b) Experimental nanoindentation curves performed on procapsids adsorbed on 5-fold (red), 3-fold (green) and 2-fold (blue) symmetry axes and on the substrate (black).



	Ratios k/k$_{5f}$ mature virion	
Fold	Experiment	Simulation
2/5	3.6 ± 0.5	2.6±0.3
3/5	1.9 ± 0.3	2.2±0.3

Table 10.1 Comparison between the spring constants measured for the different orientations in the AFM experiments for T7 mature virion, and for our simulations with a $T = 7$ shell, with $k_c = 40$, $k_{repul} = 80$, $\nu = 1.796$, $\alpha = 0.7$ and $\sigma_{pp} = \sigma_{hh}$.

T7 capsid and based on previous model [35], we have also performed AFM simulations for our discrete model of the bacteriophage T7 capsid, coarse-grained at the level of the morphological units. Fig.10.16(a) shows the resulting indentations curves. The spring constants measured in the simulations were obtained by a linear fit, and the results are listed in Table 10.1. Remarkably, the results of our discrete model reproduce the same order of the effective spring constants found in the experiments (i.e. $k_{5c} < k_{3c} < k_{2c}$, see Fig.10.16). This suggests that the distinct stiffnesses for the different orientations found in the T7 capsid are associated to the discrete nature of its structural components and to their particular arrangement in the final structure.

10.6 Conclusions

In this chapter we have presented the implementation of a “Virtual AFM”, a BD simulation using our coarse-grained model, designed to mimic the standard setup of AFM nanoindentation experiments. We have seen that our “Virtual AFM” is a very useful tool to investigate the mechanical properties of viruses, and how they change with the relevant physical parameters. With this tool it is also possible to interpret real experiments too, providing a simple way to account for the influence of the discrete structure of the capsid, and obtaining very useful information about the physical ingredients involved in the mechanical response of viral capsids.

We have used our “virtual AFM” to interpret experiments performed on bacteriophage T7 by the group of P.J de Pablo. Our simulations, explain the unusual anisotropic stiffness found in these experiments, and reproduce qualitatively the experimental results. We have also analyzed how the stiffness depends on different physical parameters, and on the shape of the virus. In particular, we have compared the mechanical response of spherical proheads versus faceted mature capsids. In this chapter we have focused on the $T = 7$ structure, but our model can be useful for many future studies of indentation, buckling and breaking of different T-number viruses.

Appendix A

Rotations and translations to put the structure in the different symmetry axes

In order to perform nanoidentations along 2-fold, 3-fold and 5-fold axes of symmetry we must rotate and translate the structure to locate it with the proper orientation between the tip and the sample. As our initial configuration, we have used the exact coordinates of the icosahedral spherical code for $T = 7$ levogiro corresponding to the solution of the maximum volume problem from [36].

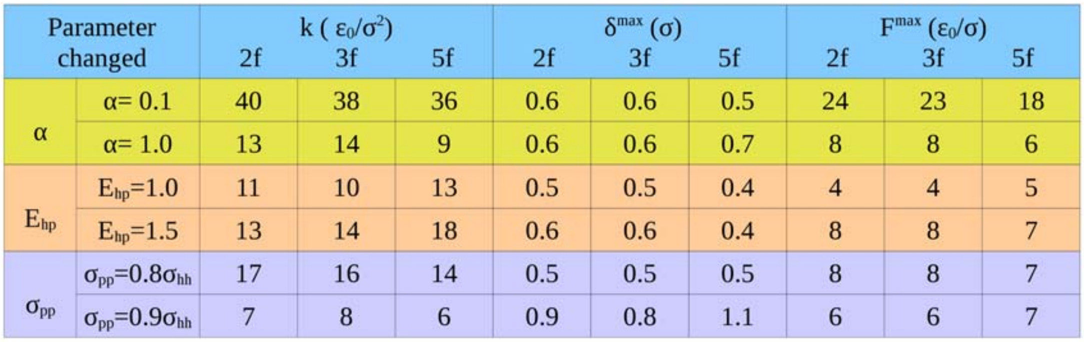
In this icosahedral spherical code, the structure has the proper orientation for a 2-fold. To convert it into a 5-fold, we first do a rotation to put one of the 5-fold axis coinciding with the z-axis. We then do a rotation around the z-axis to place one of the other pentamers at $y = 0$. Analogously, to get a 3-fold oriented shell, we first do a rotation to place the $(1, 1, 1)$ direction on the z-axis, followed by another rotation around the z axis to place one of the pentamers at $x = 0$. In all cases, the structure is translated along the z-axis so that the bottom capsomer is placed at a distance $\sigma/2$ over the xy-plane representing the substrate.

To avoid rotations or sliding of the structure under indentation, we have fixed some of the pentamers of the structure. For indentations along the 5-fold symmetry we have fixed 4 pentamers: the two pentamers located on the z-axis can only move in the z direction and the other two pentamers having initially $y = 0$, cannot move in the y-direction. For the 3-fold, we have fixed 4 pentamers: two pentamers have x, y fixed and can only move along z; and another 2 pentamers are fixed at $x = 0$. Finally, for the 2-fold direction we have fixed 8 pentamers: 4 of them have their x-coordinate fixed at $x = 0$; another 3 are fixed at $y = 0$, and of the pentamers closest to the bottom is fixed at $x = 0$ and $y = 0$. Fixing the capsomers in this way, the structures cannot rotate or slide, but the structure can be deformed under the influence of the cantilever.

Appendix B

Summary of the results for different parameters

We have performed different nanoindentation simulations for all symmetry axes. In these simulations we have calculated the maximum force and maximum indentation before the breaking of the structure, on the slope of the initial linear regime of indentation. The results are compiled in the following table:



Parameter changed		k (ε_0/σ^2)			δ^{max} (σ)			F^{max} (ε_0/σ)		
		2f	3f	5f	2f	3f	5f	2f	3f	5f
α	α= 0.1	40	38	36	0.6	0.6	0.5	24	23	18
	α= 1.0	13	14	9	0.6	0.6	0.7	8	8	6
E_{hp}	E_{hp}=1.0	11	10	13	0.5	0.5	0.4	4	4	5
	E_{hp}=1.5	13	14	18	0.6	0.6	0.4	8	8	7
σ_{pp}	σ_{pp}=0.8σ_{hh}	17	16	14	0.5	0.5	0.5	8	8	7
	σ_{pp}=0.9σ_{hh}	7	8	6	0.9	0.8	1.1	6	6	7

Table 10.2 Effective spring constant, maximum force and indentation for the 3 different symmetry axis as a function of different parameters. We have fixed in all simulations $k_c = 40$, $k_{repul} = 80$, $E_{ads} = 0.1$, and $T = 0.05$. For different values of α (yellow) we have fixed $E_{hp} = 1.0$ and all subunits have the same size. For different values of E_{hp} (orange) we have fixed $\alpha = 0.7$ and all subunits have the same size. Finally, when we changed the relative pentamer size, we have fixed $E_{hp} = 1.0$ and $\alpha = 0.7$ (violet).

References

[1] M. F. Moody, Geometry of phage head construction, *J. Mol. Biol.*, **293**, 401-33, 1999.

[2] A. Cordova, M. Deserno, W. M. Gelbart, and A. Ben-Shaul, Osmotic shock and the strength of viral capsids, *Biophys. J.*, **85**, 70-74, 2003.

[3] J. Kindt, S. Tzlil, A. Ben-Shaul, and W. M. Gelbart, DNA packaging and ejection forces in bacteriophage, *PNAS*, **98**, 13671-13674, 2001.

[4] P. K. Purohit, M. M. Inamdar, P. D. Grayson, T. M. Squires, J. Kondev, and R. Phillips, Forces during bacteriophage DNA packaging and ejection, *Biophys. J.*, **88**, 851-866, 2005.

[5] G. Rice, K. Stedman, J. Snyder, B. Wiedenheft, D. Willits, S. Brumfield, T. Mc-Dermott, and M. J. Young, Viruses from extreme thermal environments, *PNAS*, **98**, 13341-5, 2001.

[6] M. Prigent, M. Leroy, F. Confalonieri, M. Dutertre, and M. S. DuBow, A diversity of bacteriophage forms and genomes can be isolated from the surface sands of the Sahara Desert, *Extremophiles*, **9**, 289-296, 2005.

[7] D. E. Smith, S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson and C. Bustamante, The bacteriophage ϕ29 portal motor can package DNA against a large internal force, *Nature*, **413**, 748 (2001).

[8] JP. Michel, IL. Ivanovska, MM. Gibbons, WS. Klug, CM. Knobler, GJL. Wuite, and CF. Schmidt, Nanoindentation studies of full and empty viral capsids and the effects of capsid protein mutations on elasticity and strength, *PNAS*, **103**, 6184-6189, 2006.

[9] W. Klug, R. Bruinsma, J.-P. Michel, C. Knobler, I. Ivanovska, C. Schmidt, and G. Wuite, Failure of viral shells, *Phys. Rev. Lett.*, **97**, 228101, 2006.

[10] G. A. Vliegenthart and G. Gompper, Mechanical deformation of spherical viruses with icosahedral symmetry, *Biophys. J.*, **91**, 834-41, 2006.

[11] M. Zink and H. Grubmuller, Mechanical properties of the icosahedral shell of southern bean mosaic virus: A molecular dynamics study, *Biophys. J.*, **96**, 1350-1363, 2009.

[12] I. A. T. Schaap, C. Carrasco, P. J. de Pablo, F. C. MacKintosh, and C. F. Schmidt, Elastic response, buckling, and instability of microtubules under radial indentation, *Biophys. J.*, **91**, 1521-1531, 2006.

[13] WH. Roos, MM. Gibbons, A. Arkhipov, C. Uetrecht, NR. Watts, PT. Wingfield, AC. Steven, AJR Heck, K. Schulten, WS. Klug, and GJL. Wuite, Squeezing protein shells: how continuum elastic models, molecular dynamics simulations, and experiments coalesce at the nanoscale, *Biophys. J.*, **99**, 1175-1181, 2010.

[14] PJ. de Pablo, Introduction to Atomic Force Microscopy, *Single Molecule Analysis: Methods and Protocols*, **783**, 197-212, 2011.

[15] A. Ahadi, J. Colomo, and A. Evilevitch, Three-Dimensional Simulation of Nanoindentation Response of Viral Capsids. Shape and Size Effects, *J. Phys. Chem. B*, **113**, 11, 3370-3378, 2009.

[16] IL. Ivanovska, PJ, de Pablo, B. Ibarra, G. Sgalari, FC. MacKintosh, JL. Carrascosa, CF. Schmidt, and GJL. Wuite, Bacteriophage capsids: Tough nanoshells with complex elastic properties, *PNAS*, **101**, 7600-7605, 2004.

[17] I. Ivanovska, G. Wuite, B. Jonsson, and A. Evilevitch, Internal DNA pressure modifies stability of WT phage, *PNAS*, **104**, 9603-9608, 2007.

[18] D. Martinez-Martin, C. Carrasco, M. Hernando-Perez, PJ. de Pablo, J. Gomez-Herrero, R. Perez, MG. Mateu, JL. Carrascosa, D. Kiracofe, J. Melcher, and A. Raman, Resolving Structure and Mechanical Properties at the Nanoscale of Viruses with Frequency Modulation Atomic Force Microscopy, *PLOS ONE*, **7**, 1, 2012.

[19] I. Liashkovich, W. Hafezi, J. E. Kuhn, H. Oberleithner, A. Kramer and V. Shahin, Exceptional mechanical and structural stability of HSV-1 unveiled with fluid atomic force microscopy *J. Cell Sci.*, **121**, 2287-2292, 2008.

[20] W. H. Roos, K. Radtke, E. Kniesmeijer, H. Geertsema, B. Sodeik and G. J. Wuite, Scaffold expulsion and genome packaging trigger stabilization of herpes simplex virus capsids, *PNAS*, **106**, 9673-9678, 2009.

[21] N. Kol, M. Gladnikoff and D. Barlam *et al.*, Mechanical properties of murine leukemia virus particles: Effect of maturation, *Biophys. J.*, **91**, 767-774, 2006.

[22] N. Kol, Y. Shi and M. Tsvitov *et al.*, A stiffness switch in human immunodeficiency virus, *Biophys. J.*, **92**, 1777-1783, 2007.

[23] WH. Roos, I. Gertsman, ER. May, CL. Brooks III, JE. Johnson and GJL. Wuite, Mechanics of bacteriophage maturation, *PNAS*, **2012**, 109, 2342-2347, 2012.

[24] M. Buenemann and P. Lenz, Mechanical limits of viral capsids, *PNAS*, **104**, 9925-9930, 2007.

[25] MM. Gibbons and WS. Klug, Influence of nonuniform geometry on nanoindentation of viral capsids, *Biophys. J.*, **95**, 8 3640-3649, 2008.

[26] A. Ionel, JA. Velázquez-Muriel, D. Luque, A. Cuervo, JR. Castón, JM. Valpuesta, J. Martín-Benito and JL. Carrascosa, Molecular Rearrangements Involved in the Capsid Shell Maturation of Bacteriophage T7, *J. Biol. Chem.*, **286**, 234-242, 2010.

[27] X. Agirrezabala, J. Martin-Benito, JR. Caston, R. Miranda, M. Valpuesta and JL. Carrascosa, Maturation of phage T7 involves structural modification of both shell and inner core components *EMBO J.* **24** 3820-3829, 2005.

[28] GA. Vliegenthart and G. Gompper, Mechanical deformation of spherical viruses with icosahedral symmetry, *Biophys. J.*, **91**, 3, 834-841, 2006.

[29] ER. May, J. Feng, and CL. Brooks III, Exploring the Symmetry and Mechanism of Virus Capsid Maturation Via an Ensemble of Pathways, *Biophys. J.*, **102**, 3, 606-612, 2012.

[30] ER. May and CL Brooks III, On the Morphology of Viral Capsids: Elastic Properties and Buckling Transitions, *J. Phys. Chem. B*, **116**, 29, 8604-8609, 2012.

[31] AC. Steven, JB. Heymann, N. Cheng, BL. Trus and JF. Conway, Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity, *Curr. Opin. Struc. Biol.*, **15**, 227-236, 2005.

[32] MG. Mateu, Mechanical properties of viruses analyzed by atomic force microscopy: A virological perspective, *Virus Research*, **168**, 1-22, 2012.

[33] X. Agirrezabala, *et al.*, Maturation of phage T7 involves structural modification of both shell and inner core components, *Embo Journal*, **24** 3820-3829, 2005.

[34] M. Hernando-Pérez, M. Aznar, E. Pascual-Vega, A. Ionel , JR. Castón, A. Luque, JL. Carrascosa, D. Reguera and P J. de Pablo, The interplay between chemical stability and mechanical fragility of viral protein cages, *submitted PNAS*.

[35] A. Ionel, *et al.*, Molecular Rearrangements Involved in the Capsid Shell Maturation of Bacteriophage, J. Biol. Chem., **286**, 1, 234-242, 2011.

[36] Hardin RH, Sloane NJA and Smith WD. Tables of spherical codes with icosahedral symmetry, published electronically at http://www.research.att.com/∼njas/icosahedral.codes/.

Part IV

Conclusions and perspectives

Conclusions and Perspectives

In this thesis we have studied the general physical principles that are involved in several important processes in the life cycle of viruses. In particular, we have focused on the viral self-assembly process and the mechanical properties of the resulting capsid. We have developed simple, but physically inspired, coarse-grained models implemented into different simulations techniques, to get some insight into these phenomena. Now we will present the main conclusions of this thesis and the most relevant perspectives for future work.

The first part of the thesis was mainly devoted to review important aspects of the physical modeling of viruses. Specifically, in the first chapter of the thesis, the geometrical principles involved in the construction of spherical and elongated virus were revised. The importance of these geometrical ideas is that they set the basis for the identification and classification of viruses. Moreover, we have seen that the architecture and symmetry of the capsid plays an important role in the mechanical properties of the virus.

The second chapter was devoted to outline how classical nucleation theory (CNT) can be adapted to explain the kinetics of viral capsid assembly. The most salient features of this theory are the presence of a free energy barrier and the requirement of a critical size and concentration to trigger the self-assembly of capsids. The theory offers quantitative predictions that facilitate the understanding of the process and could guide the interpretation of experiments and simulations.

One of the most important results of this thesis is the development of a physically-inspired coarse grained model based on a simple interaction potential between the structural units of the capsid. The parameters of this model are directly related to physically relevant properties that can be measured experimentally, such the preferred angle of interaction between capsid proteins or the bending stiffness. The advantage of the model with respect to previous models is that it does not require artificial constraints or a non-trivial tuning of multiple interaction parameters to get the right structures.

In the continuum limit, the model reproduces the standard elastic response of a thin shell. Using this analogy we explored the interplay between bending and stretching interactions in determining the optimal size of a spherical capsid. We found that in general there are two potentially optimal radius for a spherical capsid: one that minimizes the stretching and another one that minimizes the bending. However, in most cases the size of the shell is a compromise solution between these two limiting values.

With our coarse-grained model, we have studied the stability and selection of different shapes of viral capsids, the kinetics of capsid formation, and its mechanical response upon indentation, obtaining general predictions that could be tested experimentally. In particular, Part II of the thesis was devoted to the study of the self-assembly of viruses.

In this context, we have found that the preferred angle and the local bending stiffness are the main physical parameters controlling the stability and selection of a particular structure by a virus. We have characterized the stability phase diagram of the smallest T-number structures and the snub cube using Monte Carlo simulations. In addition, we have studied the tolerance of the different shells to changes in physical parameters related to ambient conditions. We have obtained not only the optimal spherical structures of capsids, but also elongated, faceted, and decapsidated structures that are also present for many real viruses. A proper understanding of the ingredients that control the *in vitro* assembly of viruses is essential to get capsids with well-defined size and structure that could be used for promising applications in medicine or bionanotechnology. Linking the parameters of the model with the biophysical environmental conditions will open the door to a better control of the structure of artificial viral shells *in vitro*.

We then analyzed in Chapter 7 the kinetics of self-assembly and the intermediate steps in the formation of a viral shell, that are experimentally inaccessible due to the activated nature of the process. We found that the conditions that ensure the integrity of a virus are not necessarily the same ones that favour its formation *in vitro* from a solution of capsomers. We then constructed an assembly phase diagram with the outcome of assembly simulations as a function of the relevant physical parameters. This phase diagram shows the strong competence between structures that could justify the polymorphism of some viruses reconstituted *in vitro*, and the dramatic influence of the strength of interactions and density on the process. We have also analyzed and characterized accurately the kinetics of assembly of $T = 3$ shells. All features of the assembly point out to nucleation as the physical mechanisms controlling shell formation, and we were able to determine the critical cluster sizes, the height of the nucleation barrier and the rate of capsid formation directly from our assembly simulations, at different conditions. The results agree at least qualitatively with the predictions of CNT and open the door to a proper understanding and control of viral assembly.

Part III of the thesis focused on different mechanical properties of viruses determined combining theory, simulations with our coarse grained model, and Atomic Force Microscopy experiments.

In particular, in Chapter 8 we determined how the capsid shape and the buckling transition depend on the triangulation number T and the icosahedral class P of the virus structure, by using a simplified version of the model constrained on a spherical or icosahedral template. We found that, for small shells, capsids with $P = 1$ are most likely to produce polyhedral shapes that minimize their energy and accumulated stress, whereas viruses with $P = 3$ prefer to remain spherical. For big capsids, all shells are more stable adopting an icosahedral shape, in agreement with continuum elastic theory. Moreover, spherical viruses tend to become polyhedral under expansion, in consonance with virus maturation. The resulting icosahedral shell is mechanically stiffer, tolerates larger expansions and withstands higher internal pressures before failing. This could explain why some dsDNA viruses, which rely on the pressurization of their genetic material to facilitate the infection, undergo a buckling transition. The predictions of this simple model agree surprisingly well with the known structural data of viruses.

In Chapter 9, we show how the combination of theory and nanoindentation experiments using AFM could be very useful in the study of the mechanical properties of viruses. In particular, a crude continuum description and Finite element simulations were used to determine the internal pressure of $\phi 29$. The presence of confined and densely packed double-stranded DNA inside the capsid was found to alter its elastic response under nanoindentation, increasing the effective spring constant. From this change, it was possible to infer the associated pressure, which was unequivocally linked to the electrostatic self-repulsion of the confined DNA, and agrees with a simple theoretical estimate using the so-called inverse spool model.

A significant step forward in the comparison with experiments is the development of what we called “Virtual AFM”. This is a BD simulation aimed to mimic the standard setup and conditions of typical AFM nanoindentation experiments. With this virtual AFM we analyzed how the mechanical response of a virus depends on the different relevant parameters such as temperature, the axial symmetry of the capsid, the spontaneous curvature, the bending stiffness, or the binding energy and relative size of the capsomers. We also analyzed the different mechanisms of failure and breaking of viral shells under mechanical stress. Finally, using this Virtual AFM we were able to interpret and justify the astonishing anisotropic response of bacteriophage T7 found experimentally.

In this thesis we have used general theoretical ideas and physical principles that are present in a huge number of viruses. Accordingly, the conclusions and results found are not limited to a particular viral example, but rather are expected to apply to a wide class of viruses. In addition, the main ideas and methodologies developed in the thesis can be used and extended in many different directions in the study of viruses.

In particular, one of the focus of the thesis has been the study of the kinetic mechanisms in the viral self-assembly of all-pentamer viruses. In this context, we plan to extend our kinetic assembly simulations to other triangulation numbers, and to particularize them to analyze *in vitro* experiments of specific viruses such as polyoma [1], papilloma [2] and SV40 [3]. In a next step, we are currently extending these studies to the analysis of the assembly of viruses that use pentamers and hexamers preformed in solution as basic assembly units. This scenario is more complex and the kinetics process of assembly probably depends not only on the temperature, the spontaneous curvature and the bending stiffness but also on the relative interaction between capsomers, and their relative concentrations in solution.

Another important aspect of the assembly that we plan to explore is the additional factors required to assemble elongated structures efficiently. Although we have found that these structures are energetically optimal for some range of parameters, their formation from solution seem to require additional ingredients. In this context, we will explore the role of the genetic material and the torsion term on favoring the appearance of these structures.

A priority in our future research is to incorporate the genetic material in our models and simulations. In fact, we have implemented a preliminary simulation with a genome chain, where each base-pair is modeled as a bead with proper spring, electrostatic and bending interactions to model the stiffness and flexibility of DNA and RNA chains [4, 5]. Using these simulations we plan to study the role of the genetic material in the co-assembly of different viruses, as well as its influence on the mechanical response of the virus in our "Virtual AFM".

We are also interested in analyzing more complex assembly scenarios. One example is the formation of capsids in the presence of scaffolding proteins. In this case the new ingredients will be, essentially, the size and interaction between coat and scaffolding proteins, their stoichiometry in the capsid, and the concentration of scaffold proteins in solution. This mechanism will be compared with the experimental information on the assembly of viruses like bacteriophage P22 [6].

Another long term aspiration in our future research is to develop progressively more realistic physical models for the subunits that form the capsid and their intermolecular interactions. There are many examples of viruses that use dimers, trimers, or even individual proteins as assembly units [7, 8]. We intend to develop increasingly realistic model units and interactions to describe these situations, comparing with the available experimental data.

Finally, the development of the "Virtual AFM" is the first step in the characterization of the mechanical properties of different viruses. In this thesis, we focused on the properties of a $T = 7$ shell, but a similar analysis can be done for other T-number shells. In this sense, we plan to perform a more detailed study on the influence of the T-number, the P-class, and the different relevant parameters of the interaction in the mechanical response of viruses under nanoindentation. A more careful analysis of the buckling and breaking of viruses is also planned. We will also use our virtual AFM to help the interpretation of AFM experiments.

In summary, despite the many new discoveries in virology during the last years, and the emerging interest in the interdisciplinary field of physical virology, there are many open questions about the properties and life-cycle of viruses that remain to be solved. A better understanding of the physical mechanisms involved in the life cycle of viruses might help to control their impact in human life, and also to exploit their amazing properties in pharmaceutical and nanotechnological applications.

References

[1] D. M. Salunke, D. L. D. Caspar, and R. L. Garcea, Polymorphism in the assembly of polyomavirus capsid protein VP1, *Biophys. J.*, **56**, 887-900, 1989.

[2] G. L. Casini, D. Graham, D. Heine, R. L. Garcea, and D. T. Wu, In vitro papillomavirus capsid assembly analyzed by light scattering, *Virology*, **325**, 320-7, 2004.

[3] S. Kanesashi, K. Ishizu, M. Kawano, S. Han, S. Tomita, H. Watanabe, K. Kataoka, and H. Handa, Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro, *J. Gen. Virol.*, **84**, 1899-1905, 2003.

[4] OM. Elrad and MF. Hagan, Encapsulation of a polymer by an icosahedral virus, *Phys. Biol.*, **7**, 045003, 2010.

[5] J. Rudnick and R. Bruinsma, Icosahedral packing of RNA viral genomes, *Phys. Rev. Lett.*, **94**, 1-4, 2005.

[6] P. A. Thuman-Commike, B. Greene, J. A. 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**, 559-568, 1998.

[7] J. M. Johnson, J. Tang, Y. Nyame, D. Willits, M. J. Young, and A. Zlotnick, Regulating self-assembly of spherical oligomers, *Nano Lett.*, **5**, 765-70, 2005.

[8] KN. Parent, RS. Sinkovits, MM. Suhanovsky, CM. Teschke, EH. Egelman, and TS. Baker, Cryo-reconstructions of P22 polyheads suggest that phage assembly is nucleated by trimeric interactions among coat proteins, *Phys. Biol.*, **7**, 045004, 2010.

List of publications

* M. Aznar, A. Luque and D. Reguera, Relevance of capsid structure in the buckling and maturation of spherical viruses, *Physical Biology*, **9**, 3, 2012.

* M. Hernando-Perez, R. Miranda, M. Aznar, JL. Carrascosa, IAT. Schaap, D. Reguera and PJ. de Pablo, Direct Measurement of Phage phi29 Stiffness Provides Evidence of Internal Pressure, *Small*, **8**, 15, 2366-2370, 2012.

* M. Hernando-Pérez, M. Aznar, E. Pascual-Vega, A. Ionel , JR. Caston, A. Luque, JL. Carrascosa, D. Reguera and P J. de Pablo, The interplay between chemical stability and mechanical fragility of viral protein cages, *Preprint to be submitted to PNAS.*

* M. Aznar and D. Reguera, Physical ingredients controlling the polymorphism and stability of viral capsids, *Preprint.*

* M. Aznar and D. Reguera, Elastic modeling in viral capsids, *Preprint.*

* M. Aznar and D. Reguera, Simulations of the assembly kinetics of empty viral capsids, *Preprint.*