Developments of a Simple Model to Elucidate the Shape of Enveloped Viruses: Motivated by Monkeypox and SARS-CoV-2

Student: Hua Deng

Mentor: Dr. Chwen-Yang Shew

February 21, 2025

Objective

To develop new models and simulations to study how the viral genome architecture(its shape, length and flexibility) influences the membrane. By combining polymer physics and liquid-state physics, such as crowding effects, we hope to gain a deeper understanding of viral assembly. By simulating how viral genomes behave in confined environments, we can uncover key principles behind virus formation and stability. It could contribute to more effective strategies for preventing and treating viral infections, especially for viruses like Monkeypox and SARS-CoV-2.

Introduction

Viral infections have a critical impact on global health, emphasizing their role in pandemics and outbreaks. These infections pose significant risks to public health and encompass a wide range of diseases, from seasonal influenza to more severe infectious diseases such as Monkeypox and COVID-19. Both Monkeypox and COVID-19 viruses are classified as enveloped viruses, which are distinguished by their lipid membranes that surround and protect their genetic material (genomes). These membranes serve several crucial functions: they provide protection by shielding the viral genome from environmental factors and potential damage, and they facilitate transmission by enabling the virus to enter host cells. The membrane proteins and polysaccharides play a key role in this process by interacting with specific receptors on host cells, thereby promoting viral infection and the delivery of the genome into the host. We need to further understand these mechanisms to develop effective prevention and treatment strategies.

Virions are acellular, meaning they are not made up of cells and therefore lack cellular components like organelles, ribosomes, or a plasma membrane. Instead, a virion consists primarily of a nucleic acid genome encased within a protein coat called a capsid.

SARS-CoV-2, the virus responsible for COVID-19, enters a host cell by first using its spike (S) protein to bind to a specific receptor on the surface of human cells called ACE2 (angiotensin-converting enzyme 2). This spike protein acts like a key, allowing the virus to attach tightly to the host cell. After attachment, the virus enters the cell either through direct fusion with the cell membrane or via endocytosis. Once inside, the viral envelope is removed (a process called uncoating), releasing its positive-sense single-stranded RNA genome into the host cell's cytoplasm. This viral RNA acts directly as messenger RNA (mRNA) and is translated by the host's ribosomes to produce viral proteins, including more spike proteins, structural proteins, and enzymes necessary for viral replication. The viral genome is also copied to make more RNA strands. These components are then assembled into new viral particles. Finally, the new viruses are released from the cell by budding, often taking a piece of the host's membrane with embedded spike proteins, allowing them to infect new cells. The spike protein is essential for both the entry of the virus and for determining which cells it can infect.

Viruses often have highly symmetrical structures, such as icosahedrons (20-sided shapes). This symmetry helps distribute internal pressure evenly across the viral capsid, preventing weak spots and making the virus stable and protective of its genetic material.

When a virus infects a host, it deliberately breaks this symmetry. Doing so creates weaker areas on the surface of the capsid, which then open up to allow the virus to inject its genetic material (RNA or DNA) into the host cell. In this way, breaking symmetry is an essential part of the infection mechanism.

Viruses pack their genetic material DNA or RNA into tiny protein shells called capsids under extremely high pressure, often reaching tens of atmospheres, far greater than the pressure inside a champagne bottle. During the assembly of a virus, a specialized motor protein powered by ATP works to push the viral genome into the small, enclosed space of the capsid. This step requires a significant amount of energy because the DNA or RNA has to be crammed into a space much smaller than its relaxed form. As a result, the genetic material becomes tightly packed,

creating considerable internal pressure. This pressure mainly arises from two factors: the repulsion between the negatively charged parts of the nucleic acid strands and the physical strain of bending a rigid, long molecule into such a confined area.[1]

This pressurization plays a crucial role in the infection process. When a virus encounters a host cell, the built-up pressure inside the capsid acts like a compressed spring, forcefully ejecting the genome into the host cell. This mechanism ensures that the viral genetic material enters the cell rapidly and efficiently, allowing the virus to take over the host's cellular machinery almost immediately. This strategy is not limited to a single type of virus; it is conserved across many viral families, including those that infect humans, bacteria, and archaea. For instance, studies have shown that herpes simplex virus type 1 (HSV-1) can exert internal pressures of up to 20 atmospheres, which is essential for successfully delivering its DNA into the host.

Understanding this mechanism at a deeper level has important implications for medicine. By targeting the molecular motor or the structural features that allow such high pressure to build up and be maintained, scientists are exploring ways to develop antiviral drugs that could block the genome ejection step. If successful, these treatments could prevent a virus from completing its infection cycle, offering a new class of antiviral therapies that stop viruses before they gain a foothold in the body.

In our previous study, we developed a simple yet insightful model to understand how monomers can self-organize into trimers and tetramers on spherical surfaces. Our findings show that trimers are easier to form and more stable than dimers, and in many conditions, they are as stable—or even more stable—than tetramers. Tetramers require stronger attractive forces to form and are more prone to aggregation when these forces become too strong. These results highlight the importance of both energetic and entropic contributions to self-organization, particularly how interaction geometry and asymmetry influence the outcome.

Beyond virus-related modeling, we also explored how rod-like molecules behave inside confined spaces, such as droplets or cavities, under molecular crowding. These simulations revealed that shorter rods tend to accumulate near the boundary, while longer rods align in the interior—behaviors that closely match experimental observations. Altogether, our work provides a foundation for understanding how structure and organization emerge in complex systems.
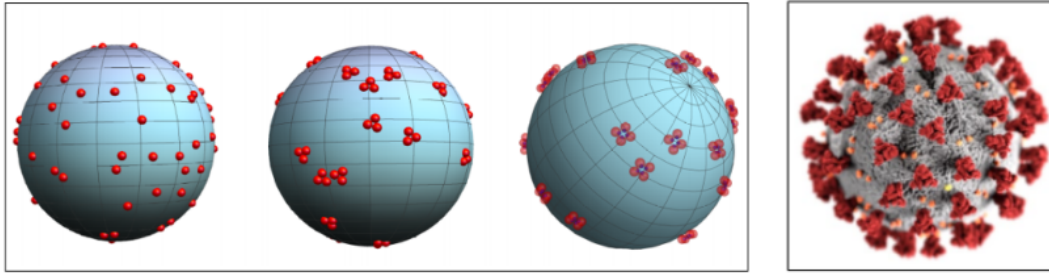
Figure 1: Left image: Simulation of the SARS-CoV-2 spike protein on a spherical surface, illustrating the progression from randomly distributed monomers (left) to trimer (center) and tetramer (right) formations. Right image: SARS-CoV-2 for Covid-19 schematic image (CDC Public Health Image Library) [2]
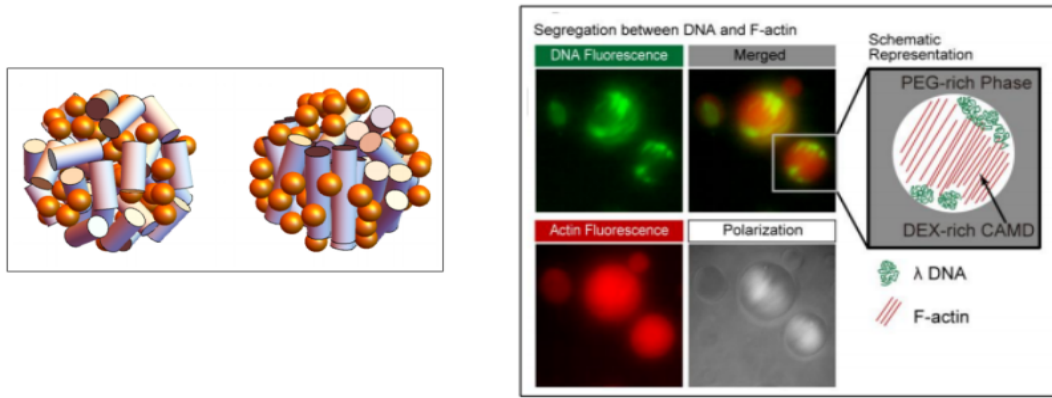

Figure 2: **Left image:** Short DNA rods distribute near wall under molecular crowding. Long F-actin rods distribute in cavity interior and show high orientation ordering. Long F-actin rods distribute in cavity interior and show high orientation ordering. Specific localization of long DNA (λ DNA) and F-actin in DEX-rich CAMDs. **Right image:** Fluorescence microscopic images of DNA (GelGreen), actin (Alexa Fluor 546), and merged views are shown, along with polarization microscopy observations (four panels on the left). The images were captured under the conditions of 120 μM λ DNA, 10 μM actin, and 4.0 mM KCl. F-actins were observed to be in a nematic liquid-crystal state at the center of DEX-rich CAMDs, while DNA molecules appeared segregated from the center. In other words, long DNA strands are compressed by the aligned F-actin region, as schematically illustrated in the right panel. Scale bar: 100 μm.[3]

Background

The genomic material of SARS-CoV-2 (single-stranded RNA) and Monkeypox (double-stranded DNA) behaves like a polymer.

I. Polymer

The polymer is a large molecule composed of identical monomers or different types of monomers linked by covalent bonds. Both RNA and DNA are composed of long sequences of nucleotides linked together by phosphodiester bonds.
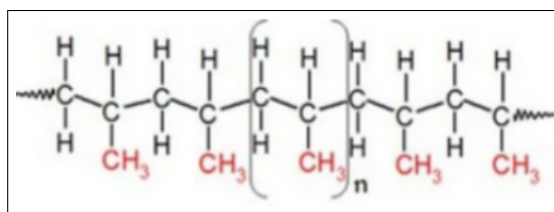
Figure 3: Polymer

II. Membrane/Phospholipids

Phospholipids are the main ingredients of cell and viral membranes, and their structure is actually a lot like a simple polymer. Each one has a head that loves water and two tails that hate
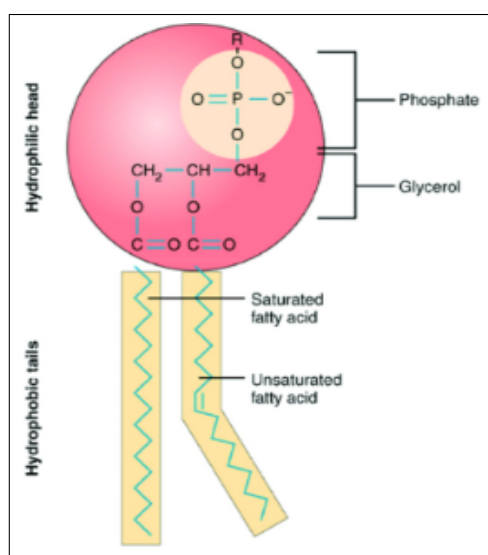

Figure 4: membrane subunit

it. The head is connected to the tails through a small backbone, so the whole molecule looks kind of like a matchstick or a two-tailed kite. Because one part of the molecule wants to be near water and the other part wants to avoid it, phospholipids naturally arrange themselves into layers when they're in water. These layers form flexible membranes that act a lot like soft polymer sheets—they can bend, move, and even stretch a little. In fact, scientists often use ideas from polymer physics to study how these membranes behave, especially when trying to understand things like how viruses wrap themselves up or how they enter cells.

III. Genome/DNA/RNA

DNA and RNA are long chains made up of smaller units called nucleotides. Each nucleotide is like a puzzle piece, made of a sugar, a phosphate group, and a nitrogen base. In DNA, the sugar is called deoxyribose, and in RNA, it's ribose. The bases are the parts that carry genetic
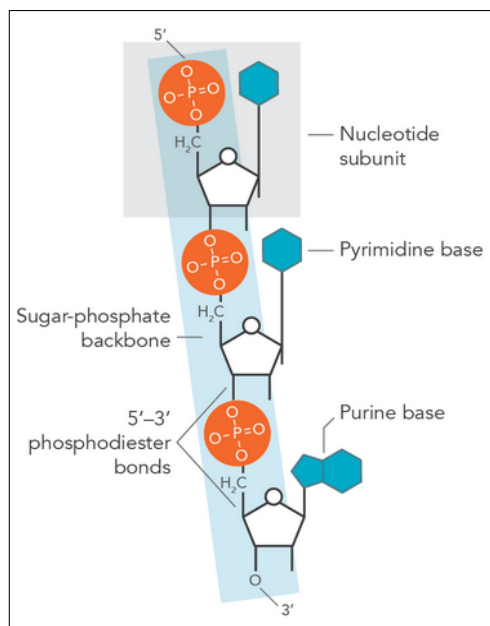

Figure 5: Nucletide subunits

information—A, T, G, and C in DNA, and A, U, G, and C in RNA (with U replacing T in RNA). These nucleotides link together through bonds between the phosphate of one and the sugar of the next, forming a backbone, kind of like beads on a string.

Nucleotides, the building blocks of RNA and DNA, act as repeating units that form larger macromolecules, similar to how polymers are made from repeating monomers. The polymeric nature allows genomic material to adopt different shapes and configurations. Just as polymers can interact with other compounds, RNA and DNA can form secondary structures and interact with proteins, which is essential for processes like gene expression and viral assembly. Polymers can have various structures—linear, branched, or crosslinked—allowing them to be flexible, adaptable, and capable of forming diverse shapes and configurations. DNA and RNA are biopolymers, and it means they have a polymeric nature that enables them to fold, coil, and interact with other molecules based on factors like sequence, chemical interactions, and environmental conditions.

This makes DNA and RNA behave like polymers—long, repeating chains that can twist,

bend, and fold. DNA usually forms a double helix, with two strands held together by base pairing (A with T, and G with C), while RNA is usually single-stranded and folds into all kinds of shapes. In science, especially in physics, we can think of DNA and RNA as flexible polymers, which helps us understand how they fold up inside tiny spaces—like the inside of a virus.

IV. Protein/Amino Acid chain

Proteins are natural polymers made up of amino acids, which are like the individual links in a chain. Each amino acid has the same basic structure: a central carbon atom (called the alpha carbon) bonded to a hydrogen, an amino group ($-NH_2$), a carboxyl group ($-COOH$), and a unique side chain, known as the R group. These amino acids connect through peptide bonds, forming a
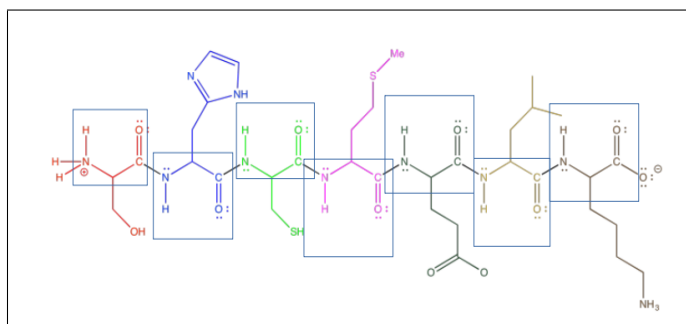

Figure 6: Amino acid chain

backbone that repeats along the entire length of the chain—just like a synthetic polymer has repeating units. But proteins don't stay as simple straight chains. They start folding right away into more complex shapes. First, they form local patterns like alpha helices (coils) and beta sheets (pleated folds), which are stabilized by hydrogen bonds. Then the entire chain folds into a specific 3D structure, known as the tertiary structure, shaped by interactions between the R groups—like hydrogen bonding, hydrophobic packing, or ionic attractions. Sometimes multiple folded chains come together to form a quaternary structure. This whole folding process is what makes a protein functional. So while a protein is a polymer at its core, its complexity and flexibility come from the variety in its amino acids and the way they interact in space.

After a protein folds into its final shape, it still technically is a polymer—it's just a much more complex and organized one. Instead of looking like a simple chain or string, a folded protein looks more like a tangled but purposeful knot or a compact 3D sculpture. The backbone of the

chain is still there, repeating like in any polymer, but now it's been twisted, looped, and packed in a specific way to create a functional structure. Some parts form spirals (alpha helices), others form sheets (beta sheets), and the rest bends and loops to connect them all together. Even though it's not as uniform-looking as a plastic polymer, it still behaves like a polymer chain in the sense that it's made from repeating units (amino acids), it has flexibility, and it responds to its environment. So a folded protein is like a smart, dynamic polymer—one that's shaped by chemistry, and whose final form is directly tied to its job inside the body.
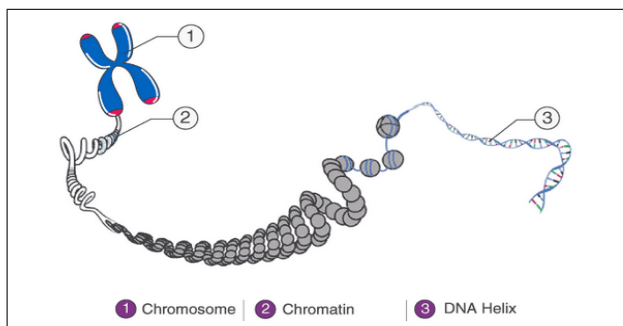

Figure 7: DNA, Chromatin and Chromosome

A coarse-grained model simplifies the structure of chromatin (DNA and proteins) and membrane for computational simulations, focusing on larger-scale organization rather than individual atoms or base pairs.

Method

1 Models

1.1 Coase-grain models

1.1.1 Adveantage of Coase-grain model

Using a coarse-grain model in polymer science is really helpful because it makes things simpler and more efficient, especially when dealing with large and complex systems. One of the biggest reasons to use coarse-graining is that it saves a lot of computational power. Polymers are often made up of thousands or even millions of atoms, so simulating each individual atom takes a

huge amount of time and resources. Instead of modeling each atom, a coarse-grain model groups several atoms or monomers into a "bead," which reduces the number of particles in the simulation. This allows researchers to study bigger systems and longer timescales without overwhelming their computers.

Another advantage is that coarse-graining allows scientists to focus on the big picture of polymer behavior. At the atomic level, polymers have tons of detailed interactions, but often what we care about are the overall properties, like how the polymer stretches, bends, or reacts to forces. Coarse-graining simplifies this by letting researchers study these larger-scale behaviors without getting caught up in the tiny atomic details, which are often less important for understanding real-world polymer behavior.

Coarse-grain models also make it easier to model interactions between polymers. Instead of focusing on each individual monomer, coarse-graining represents groups of monomers as beads. This makes the interactions between polymer chains less complex and easier to study. This is really useful for studying complicated systems like polymer networks, gels, or blends, which would be too difficult to model atom by atom.

Another reason coarse-grain models are so valuable is that they help researchers simulate larger systems over longer periods of time. By reducing the number of particles in the model, simulations can run much faster and allow for studies over longer timescales. Instead of simulating each tiny atom's movement, which can take a long time, coarse-graining lets researchers study the overall behavior of the polymer chain over much longer periods. This is really useful for looking at things like polymerization processes or phase transitions that occur over time.

Coarse-grain models also focuses on the important polymer properties that matter most, like flexibility, shape changes, and how polymers self-assemble. These properties are often determined by the overall structure of the polymer, not by the exact positions of every single atom. So by simplifying the model, researchers can study how polymers fold, how they form membranes, or how they interact with other molecules, without having to get into the atomic-level details.

Finally, coarse-grain models are essential for designing new materials. When researchers are developing new types of polymers—whether for drug delivery, biodegradable plastics, or high-performance materials, these models help predict how the polymer will behave in different

environments. This helps researchers optimize the materials for real-world uses without needing to simulate every tiny detail. In the end, coarse-grain models are incredibly useful in polymer science because they simplify complex systems, making it easier to study their behavior. They save time and computational power while still allowing scientists to focus on the big picture—like how polymers stretch, bend, and interact—helping to design better materials and understand how polymers work in the real world.

1.1.2 Assumption of Coarse grain model

Coarse-grained models come with several assumptions that are critical for their application:

1. Homogeneity:

It is often assumed that the system is uniform, meaning properties do not vary significantly across space.

2. Bead Representation:

Each bead represents a group of atoms, assuming they behave as a single entity in terms of interactions, which may not always reflect the nuances of molecular behavior.

3. Simplified Interactions:

The models typically ignore specific interactions at the atomic level, such as hydrogen bonding or van der Waals forces, unless they are effective at the coarse-grained level.

4. Mean-Field Approximations:

Coarse-grained models often use mean-field approximations, averaging the effects of all other particles on any given particle, simplifying the interactions that can be complex in real systems.

5. Fixed Connectivity:

In polymer applications, it usually assumes fixed connectivity between beads, which may not fully account for the flexibility in the actual polymer chains.

1.2 Kremer-Grest Model for DNA chain

The Kremer-Grest (KG) model is a bead-spring type of coarse-grained model.KG model is specifically for polymers,including those relevant to chromosome structure and genome organization, and a popular way to simulate polymers in computer-based molecular dynamics. It is a specific, well-defined bead-spring model that includes Lennard-Jones (LJ) potential for excluded volume (non-bonded interactions) and FENE (Finitely Extensible Nonlinear Elastic) potential for the bonds between adjacent beads.

Instead of tracking every atom, it simplifies the system by representing each polymer chain as a string of connected beads. These beads interact with each other in two main ways: non-bonded beads repel each other using something called the Lennard-Jones (LJ) potential, which helps prevent them from overlapping, and bonded beads (those directly connected in the chain) are held together by a FENE spring, which stretches like a real polymer bond but can't break or stretch too far. To keep the system at a steady temperature—like it would be in a real fluid—a Langevin thermostat is often used, adding both friction and random kicks to the beads. The model is simple, efficient, and captures the large-scale behavior of polymers really well, which is why it's used so widely in polymer research.

Kresmer−grest Model is more physically realistic for polymers than a basic harmonic model and bead−spring model. It prevents bond crossing (thanks to FENE + LJ) and it has been extensively validated and used in thousands of studies.
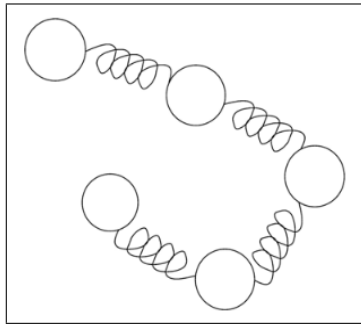

Figure 8: Bead-Spring Representation of a Polymer Chain (Kremer–Grest Model)

- **FENE spring:** $k = 30$, $R_0 = 1.5$

- **Lennard-Jones (LJ) potential:** $\varepsilon = 1$, $\sigma = 1$, cutoff $r_c = 2^{1/6}\sigma$

- **Time integration (Langevin thermostat):** friction $\gamma = 0.5$, temperature $T = 1.0$

1.3 Simple liquid Models for the organelles surrounding DNA

Simple liquid models that help understand the structure and behavior of polymer fluids in crowded environments. Specifically, the simple liquid models referenced include Hard Sphere Model and Lennard-Jones Model.

Hard sphere model is used to understand the behavior of particles in crowded environments, where the excluded volume effect becomes important.

The Lennard-Jones model accounts for both attractive and repulsive forces between particles, which is important in understanding the interactions within fluids and condensed matter. It helps model the behavior of molecules, especially in non-ideal conditions like those found in crowded environments.
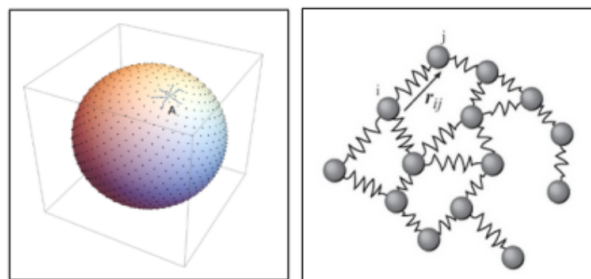
1.4 Helfrich Model with Fibonacci for the membrane

Figure 9: Left image: Fiboncci model Right image:

Continuum models are a powerful framework used to describe and simulate the large-scale, smooth behavior of biological membranes (and other materials) without tracking individual atoms or molecules. Instead of focusing on discrete particles like in molecular dynamics, continuum models treat the membrane as a continuous surface or material with defined physical properties.

To create beads on a spherical surface using the Gaussian Network Model (GNM), the first step is to place the beads evenly across the sphere.The way we use is Fibonacci lattice, which

helps distribute the points uniformly without crowding near the poles. Each bead is assigned a 3D position on the surface of the unit sphere. This means each bead's coordinates satisfy:

$$x^2 + y^2 + z^2 = r^2$$

where x, y, and z are the coordinates of a point on the sphere, and r is the radius of the sphere.

After that, we define a cutoff distance—a threshold that determines whether two beads are close enough to be connected. For each pair of beads, we calculate the distance between them, and if it's less than or equal to the cutoff, we connect them with a virtual spring. These springs represent interactions or flexibility between the nodes.

Once all the connections are made, we construct what's called the Kirchhoff matrix (or the GNM matrix). In this matrix, a value of -1 is placed at the off-diagonal positions where two beads are connected, and 0 where they're not. The diagonal entries are set so that each row adds up to zero, which simply reflects how many connections each bead has. This matrix is the core of the GNM approach—it captures the structure of the network and can be used to study how the system might flex, shift, or vibrate as a whole. By following these steps, you turn a static arrangement of beads on a sphere into a dynamic network model that can simulate real-world molecular or physical behaviors.

Bead 1 is connected to beads 2 and 3

Bead 2 is connected to beads 1 and 3

Bead 3 is connected to beads 1, 2, and 4

Bead 4 is connected to bead 3

Bead 5 is not connected to any bead

The Kirchhoff matrix Γ for a simple Gaussian Network Model with 5 beads is:

$$\Gamma = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1.6 Summary of the models

Coarse-grain model and simple liquid models are valuable tools in understanding how particles interact in condensed matter systems, such as polymer fluids, from both an energetic (energy of interactions) and entropic (disorder and randomness) perspective.

II.Simulation

1.1 Monto Carlo simulation

Monte Carlo (MC) simulation is a technique used to explore different configurations of a system by randomly sampling possible states, based on probability rules that often relate to energy. Instead of following every tiny movement of particles over time—like molecular dynamics does—Monte Carlo takes a different approach. It "jumps" from one possible state to another and decides whether to accept or reject each move depending on how favorable it is, typically using criteria like energy minimization. This method becomes especially powerful when combined with coarse-grain (CG) models. Coarse-graining simplifies complex systems by grouping atoms or monomers into larger units, reducing the number of particles and interactions. This makes the system easier and faster to simulate, which is perfect for Monte Carlo techniques. With fewer variables to track, MC simulations can explore a much wider range of configurations in less time. Using MC with CG models is particularly useful for studying large-scale behaviors like polymer folding, self-assembly, phase transitions, or how polymers interact with other surfaces or structures. Since you don't need to worry about every atomic detail, you can focus on understanding the broader behavior of the material. For instance, in a coarse-grain polymer simulation, MC methods might rotate parts of the polymer chain, stretch or shift segments, or rearrange them—checking each time how these changes affect the system's overall energy or structure. In short, Monte Carlo simulations and coarse-grain modeling work hand in hand. Together, they make it possible to study complex polymer systems more efficiently and gain insights that would be hard to reach using more detailed, atomistic approaches.

1.2 Assumpations of Monto Carlo simulation

1. Randomness is representative

The simulation assumes that the random inputs it uses (often generated using pseudo-random number generators) adequately represent the variability and uncertainty of real-life scenarios.

2. Independent trials

Each simulation run (or trial) is assumed to be independent of the others. The outcome of one run does not affect the outcome of another.

3. Known probability distributions

The input variables must follow known probability distributions (e.g., normal, uniform, triangular, etc.). The accuracy of the results relies heavily on how well these distributions reflect real-world behavior.

4. Large number of iterations

The law of large numbers is a key assumption — with a large enough number of trials, the average of the simulated outcomes should approximate the expected value. More iterations generally improve accuracy.

5. System/process remains stable

The simulation assumes that the underlying process or system doesn't change over time during the simulation (i.e., it's stationary).

6. Model structure is correct

The simulation is only as good as the model it's based on. It assumes the mathematical or logical structure you build reflects the real-world process correctly.

We simplify the genome−membrane system using a coarse-grained model,and use Monte Carlo simulation to explore its behavior of the system. Monte Carlo can help fit the parameters of a coarse-grained model.Coarse-grained models and Monto carlo simulate large, complex systems (like polymers, membranes) efficiently and with reasonable accuracy.

III. Summary

Viral genomes (like those in Monkeypox and SARS-CoV-2) behave similarly to a polymer chain packed inside a small, crowded space (the viral capsid/membrane). The project aims to extend existing models and Monto Carlo simulation to study how the genome's architecture (shape, length, and flexibility) affects the overall shape of a virus. This will help us understand how viral genomes and membranes interact, which is crucial for predicting viral structures and behavior.

Formular

$$E(r) = -\Gamma \frac{\exp(-\kappa_1 r)}{r} K_B T [4] + V \tag{1}$$

The Yukawa potential V(r) describes how two particles interact when their force has a finite range

E(r) potential energy as a function of distance r

Γ is the interaction strength,

κ_1 is the screening parameter

r is the distance,

Negative sign (-) indicates the interaction is attractive.

The left-hand side $\frac{V(r)}{k_B T}$ indicates how strong the interaction is compared to the typical thermal energy $k_B T$.

The potential is negative $(-\Gamma)$, meaning it is attractive — it tends to pull particles together.

The form $\frac{e^{-\kappa_1 r}}{r}$ means:

- At short distances $(r \to 0)$, the potential behaves like $\sim \frac{1}{r}$ — like a Coulomb (electrostatic) potential.

- At large distances $(r \to \infty)$, the potential decays exponentially because of the $e^{-\kappa_1 r}$ — it becomes very small quickly.

- κ_1 controls how short-ranged the potential is: larger κ_1 means the force dies off quicker.

$$V = \int \left(\frac{\kappa_2}{2}(C_1 + C_2 - C_0)^2 + \sigma \right) dA \qquad (2)$$

The Helfrich–Canham equation describes the total energy of a membrane by combining two main contributions: bending energy and surface tension energy. The bending energy reflects how much the membrane curves away from its preferred, or spontaneous, shape. Meanwhile, the surface tension energy accounts for the cost of stretching the membrane — it increases when the membrane's surface area grows. Together, these terms help explain how membranes bend, stretch, and maintain their shape in different environments.

κ_2 is bending rigidity — it quantifies the stiffness of the membrane.

C_1, C_2 are the principal curvatures of the membraneat surface

C_0 is the spontaneous curvature, which represents the preferred cuvaturea due to asymmetry in the membrane.

σ is the surface tension(accounts for the energy cost of increasing the membrane area)

dA is integration over the entire membrane surface area.

Simulating realistic physical conditions, since many experiments happen at constant pressure (example:1 atm) and temperature (example: 298 K).

NPT Ensemble (Isothermal–Isobaric Ensemble)

NPT stands for:

N = Number of particles (constant)

P = Pressure (constant)

T = Temperature (constant)

Under NPT conditions, the system's pressure is controlled, allowing the volume to change naturally over time. In one approach, the membrane is initialized with a fixed oval or ellipsoidal shape, defined by the input parameters a, b, and c, which represent the lengths along its three principal axes. These parameters remain constant during the simulation, so the membrane maintains its overall shape while still allowing local fluctuations and interactions with the genome inside.

In another approach, the parameters a, b, and c are allowed to fluctuate during the simulation. This means the membrane can expand, shrink, or reshape dynamically in response to internal forces exerted by the genome. Such flexibility promotes more realistic interactions between the genome and the membrane — for example, the genome may push outward and deform the membrane, while changes in membrane tension or curvature can influence how the genome is organized. Using NPT conditions in this way provides a natural and dynamic view of the interplay between membranes and genomes in biological systems.

NVT Ensemble (Canonical Ensemble)

NVT stands for:

N = Number of particles (constant)

V = Volume (constant)

T = Temperature (constant)

Under NVT conditions, the volume of the system is fixed, meaning the membrane is constrained and cannot expand or contract. In one approach, the membrane is initialized with an oval or ellipsoidal shape, defined by input parameters a, b, and c, representing the lengths along its three principal axes. These parameters remain constant throughout the simulation, so the membrane keeps its shape even as the genome inside moves or grows. Because the membrane cannot adapt naturally to changes, any pressure exerted by the genome builds up inside, sometimes leading to artificial forces that would not normally appear in a flexible system.

In another approach, even though the overall volume remains fixed under NVT, small local fluctuations in the shape of the membrane are allowed while still keeping a, b, and c centered around their initial values. In this case, the membrane can slightly deform in response to internal forces from the genome without changing its total volume. While the confinement effect is still strong, allowing limited local fluctuations provides a somewhat more realistic interaction between the genome and the membrane. Studying the membrane–genome relationship in either method under NVT is useful for early-stage stabilization or when specifically exploring how genomes behave in confined, rigid environments.

The genome of the Monkeypox virus is a double-stranded DNA, approximately 190 kb (kilobases) in length, which equals about 95 kbp (kilobase pairs). This genomic length

corresponds to a substantial contour length of approximately 3000 nm. This indicates that the genetic material is significantly longer than the physical dimensions of the virus itself, which measures around 250 nm long. These viral particles can display spherical and oval shapes, indicating different developmental stages: immature virions tend to be spherical, while mature virions adopt an oval shape. This morphological differentiation can be important for recognizing stages in the virus life cycle and understanding how these changes may affect infectivity. The transition from spherical to oval shapes may be linked to the maturation process, influencing the virus's ability to enter host cells and replicate effectively.

The shape of viral particles, like those seen in the Monkeypox virus, can reveal a lot about what stage they're at in their life cycle. In the early phases, the virus forms immature virions that are usually spherical. This round shape isn't random—it helps the virus begin assembling its genetic material and proteins in a flexible, efficient way. At this point, the virus is just starting to come together, packaging its core components. As it matures, the virus goes through structural
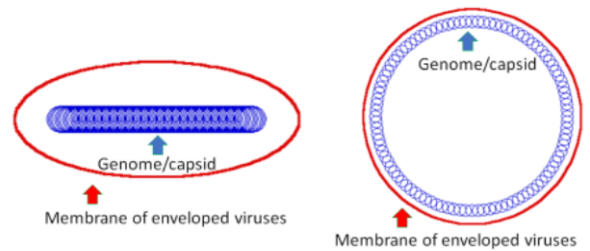

Figure 10: Preliminary 2D model studies to investigate the effect of the geometry of a genome on the shape of a virus. A rod-like genome induces an elliptic shape whereas a circular genome leads to a circular shape.

changes that shift its shape from spherical to oval. This transformation involves the rearrangement of viral proteins, which makes the particle more stable. The oval shape not only helps the mature virion survive better outside the host but also improves its ability to attach to and enter host cells. This increased stability and better interaction with host cell receptors make the mature, oval-shaped virus more infectious and better equipped to replicate once inside a host.

References

[1] A. Brandariz-Nuñez, T. Liu, T. Du, and A. Evilevitch. Pressure-driven release of viral genome into a host nucleus is a mechanism leading to herpes infection. *eLife*, 8:e47212, 2019.

[2] Centers for Disease Control and Prevention (CDC). Sars-cov-2 illustration. `https://www.cdc.gov/museum/timeline/covid19.html`, n.d. Accessed April 18, 2025.

[3] Naoki Nakatani, Hiroshi Sakuta, Makoto Hayashi, Shunsuke Tanaka, Kohei Takiguchi, Kouhei Tsumoto, and Kenichi Yoshikawa. Specific spatial localization of actin and dna in a water/water microdroplet: Self-emergence of a cell-like structure. *Chembiochem*, 19(13):1370–1374, 2018.

[4] Hideki Yukawa. On the interaction of elementary particles i. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, 17:48–57, 1935.