Title: Identifying Critical Behavior in Virus Capsid Assembly

Author: Archit K. Vasan

Date of submission: May 8, 2017

Course: Physics 563

Abstract: This essay reviews key aspects of physical virology as they apply to phase transitions in assembly mechanisms of icosahedral viral capsids. Viral architecture is first briefly reviewed with a description of the CK selection rules and some exceptions to them. The assembly mechanisms are then analyzed using empty capsids, loaded capsids and under varying ion and pH conditions.

1. Introduction

Viruses are typically viewed through the lens of a public health official: from the recent Zika virus crisis in South America to Ebola in West Africa. Or on a much more positive note, they are viewed as tools used in batteries, memory devices, nanoscaffolds and drug delivery vectors [1]. However, on a more fundamental level, what is most striking about these biological machines is perhaps not their usefulness as instruments or even their deadliness in disease, but their extreme efficiency and ability to command order in the face of chaos. Within the turmoil of the cell, viruses hijack the basic machinery to drive production of their own copies. Further, new viruses can assemble large conglomerations of essentially identical macromolecules into well-defined structures with very high fidelity.

It is this self-assembly, and in particular, the assembly of viral capsids that is of great interest to physicists. Capsid proteins typically require little to no local energy contribution such as from ATP to self- assemble; in addition, many tend to do so without the presence of internal nucleic acids (i.e. without a polyelectrolyte attractive force) [1]. Therefore, this process could be understood via the theory of equilibrium statistical mechanics. In 1955, Fraenkel-Conrat and Williams confirmed this idea through their work with tobacco-mosaic virus [2]. The researchers showed that when observed in vitro, these mosaic viruses self-assembled into functional viruses without external energy contribution. The idea was further expanded by Klug, who determined a phase diagram for TMVs dependent on ion concentration and acidity [1]. This paper determined that acidity and salinity act as parameters that modulate the hydrophobic and electrostatic interactions between capsid protein amino acids to either drive assembly or disassembly of the capsid.

Zlotnick's group's work on the cowpea chlorotic mottle virus and hepatitis B provided another major achievement in the application of thermodynamics to capsid assembly [1]. Upon measuring concentrations of subunits, they made a vital observation: the population of proteins showed a double peak, one associated with small clusters of 2-5 proteins and the other with fully formed capsids. Furthermore, upon calculating the ratio of the concentration of the small free subunits to that of the fully formed capsids they found that this approximately observed the law of mass action. This was a very important finding as it implied that the assembly of proteins into the assembly followed the behavior of phase transition theory. At a critical concentration of free capsid proteins, the fraction of proteins recruited to the capsid would rise sharply, analogous to critical behavior in an Ising model upon reaching the critical temperature. This critical concentration (C_{crit}) was then found to be related to the Gibbs free energy of assembly (ΔG_o) as described below [1].

$$C^* \propto \exp\left(\frac{\beta \Delta G_o}{N}\right) \quad (1)$$

This work by Zlotnick was a major motivating factor for many physicists to continue work in understanding the critical behavior of capsid assembly. In particular, this review will focus on both how different factors such as ion concentration and nucleic acid content drive assembly and on how different theoretical approaches such as the Landau theory of crystallization produce varying interpretations of this phenomenon.

The general organization of the following sections is as follows: section 2 focuses on a general description of viral capsid architecture, section 3 provides a brief introduction to thermodynamic investigations of empty capsids, and section 4 does the same for loaded capsids (containing nucleic acids).

2. Viral Architecture

2.1 General structure

Viruses have a very simple and elegant design: a fully developed, infectious virion's basic components are simply the outer capsid composed of oligomeric capsid proteins (typically dimers or pentamers) and some type of nucleic acid genome including double stranded DNA (dsDNA), double stranded RNA (dsRNA) or single stranded RNA (ssRNA) [3].

Capsid proteins usually have a stereotypical motif: the outer head is negatively charged while the inner tail is a positive [3]. This positive charge allows the tail to interact with nucleic acids within the virus. A basic model is presented in Figure 1. The internal nucleic acid normally acts as a type of "electrostatic glue"; the typically positively charged internal components of capsid subunits are attracted to these negative polyelectrolytes. Furthermore, the life cycle of a virus is significantly altered by the type of nucleic acid included; this idea presents a central theme in this paper.

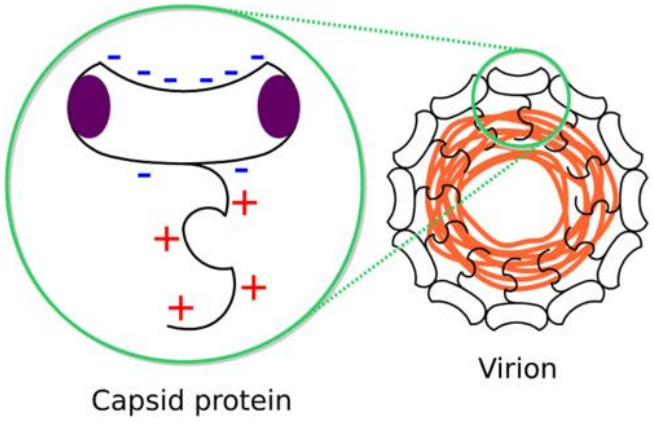

Figure 1: Reprinted from [3]. A basic model of an infectious virion: white elements on the outside are capsid proteins and orange internal coiling is nucleic acid. In addition, a detailed view of the

capsid protein is presented: it contains a negatively charged head and coil like positively charged tail.

2.2 Capsid structure

Capsids also come in a large variety of shapes: some are spherical, cylindrical, or have the characteristic syringe structure of the bacteriophage. However, most studies on viral assembly focus on viruses with approximately icosahedral symmetry such as the hepatitis B virus because of the relative prevalence of this structure: about half of all viral families share this feature [1].

In Caspar and Klug's watershed paper, it is reasoned that an icosahedral shell is preferred by so many spherical viral structures as it provided a minimum for the elastic strain felt by identical proteins on a shell. In addition to this, the paper provides four basic tenets on icosahedral viral structure that form part of the backbone of physical virology [4].

1. The symmetry of the capsid must be lower than a normal icosahedron as the capsid proteins themselves lack symmetry.
2. The total number of viral proteins in a capsid are $60T$, where T is some positive integer to be described further in principle 4.
3. Self-assembly can be modeled analogously to crystallization.
4. T is governed by a selection relation: $T = h^2 + k^2 + hk$, where h and k are described below, in Figure 2.

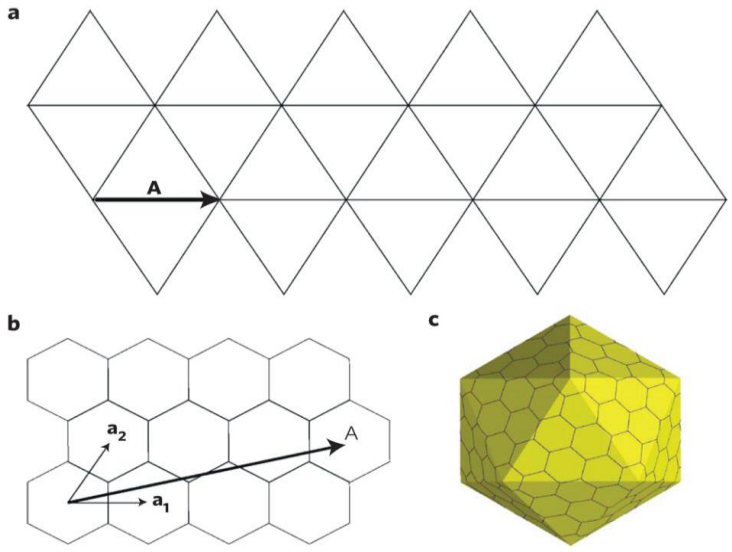

Figure 2: Reprinted from [1]. (a) This describes a template to describe the facets of a full icosahedral capsid with a length vector A. (b) The vector A is equivalent to ha_1+ka_2, where h and k are the variables in the T number CK selection rule. In this case, $h=3$, while $k=1$, so $T=13$. (c) The $T=13$ fully folded icosahedral virus.

A very large number of capsids agree with all four of the above points, however, many viruses such as is investigated in H. Naitow et al.'s work show that the *T* number selection rule is not strictly followed. In a series of papers, Lorman and Rochal design a density wave approximation of protein distribution [4] to generalize this T number selection criteria in terms of the wave number, *l*, of the density wave (Eq. 2).

$$(2) \quad l = 15 + 6i + 10j$$

This generalized rule accounts for a much larger percentage of experimentally observed icosahedral structures that violate the CK T number selection rule. Their work will be described in a bit more detail in the following section.

3. Assembly of empty capsids

Many viruses, especially those containing dsDNA, will first assemble an empty "procapsid" before injecting nucleic acid material and forming a final, infectious virion. A primary reason for this is that dsDNA has a large degree of rigidity, so it cannot act as an "electrostatic glue" like ssRNA [1]. Thus, many studies have been performed on empty structures. In addition, empty structures simplify calculations and can produce results that can be generalized to include polyelectrolyte interactions. This section will provide a review of experimental approaches as well as theoretical approaches using basic thermodynamics or the Landau theory of crystallization.

3.1 Basic thermodynamic calculations

Zlotnick et al provide a simple model to describe self-assembly in which the free energy of interaction is described as proportional to the number of subunit-subunit interactions (Eq. 3). Here, C_n is the number of subunit-subunit interactions, g_b is the subunit-subunit binding free energy and S_n represents a symmetry factor [5].

$$(3) \quad G_n^{cap} = g_b C_n - TS_n$$

As was described in the introduction, the population of capsid proteins was described via a two peaked distribution; therefore Zlotnick only considered free subunits and complete capsids in his model ; that is, $C_{tot} = C_{free} + NC_N$ [5] , where *N* represents the number of subunits within a capsid.

Then in the limit of large *N*, the fraction of fully formed capsids displays the following behavior for concentrations below and above the critical concentration: C^*. (Eq. 3 and 4)

$$(4) \quad f_c \approx \left(\frac{C_T}{C^*}\right)^N \; when \; C_T \ll C^*$$

$$(5) \quad f_c \approx 1 - \frac{C^*}{C_T} \; when \; C_T \gg C^*$$

This critical behavior was shown experimentally in HBV with the subunit-subunit energy used as a fitting parameter; their data suggested a small g_b of approximately 4kcal/mol.

Zlotnick's results identify a transition at the critical subunit concentration at which assembly begins; however, they are not explicit as to whether this transition is first order or continuous. As was alluded to earlier, Lorman and Rochal [4, 6] take a density wave theory approach to investigate transitions in capsid assembly: this approach is capable of identifying the type of phase transition. In addition, this would allow them to account for the asymmetry of capsid subunits.

3.2 Density wave theory

This methodology is as follows: in the vicinity of a crystallization point, the probability density distribution of capsid proteins is given by Eq. 4, where ρ_o is the isotropic protein density and $\Delta\rho$ is the density associated with assembly ordering.

$$(6) \quad \rho = \rho_o + \Delta\rho$$

In accordance with the Landau theory of crystallization, the order parameter is a system of density waves (CSDW) made up of spherical harmonics: Y_{lm} on the icosahedral surface [6]. This system of density waves is designed such that the free energy of transition to an ordered state is a function of CSDW amplitudes. In addition, the positions of capsid proteins are approximately defined as the maxima of the CSDW when the system is within the vicinity of a crystallization point.

The critical part of the density is $\Delta\rho$, the density associated with ordering; this value takes on a defined value for a particular wave number, l, given in Eq. 6. Here, Y_{lm} is a spherical harmonic, defined by its angular parameters and A_{lm} are the associated amplitudes [4].

$$(7) \quad \Delta\rho_l = \sum_{m=-l}^{m=l} A_{lm} Y_{lm}(\theta, \varphi)$$

Lorman and Rochal make the argument that the CK rules (asymmetry of capsid subunits) prevent the symmetry group involved in this density wave structure from having any mirror planes [4]. Thus, they claim all spherical harmonics in $\Delta\rho_l$ must have even l values going to zero.

The most important part of their work in the scope of this review paper is in describing the free energy of crystallization; however, the terms in this free energy cannot be included in this paper due to space limitations. The main point, though, is that due to requirements on the wave

number, *l*, the free energy is devoid of cubic terms [4]; this implies that according to the theory of crystallization [6], assembly resembles a continuous phase transition.

This is of further interest, as Lorman and Rochal point out, a phase transition of this type would imply that the process takes place without a nucleation step. In fact, in Zlotnik's work [7] on small viruses, this lack of a nucleation process is seen.

3.3 Nucleation and growth kinetics in bacteriophage

However, this seems to contradict work done by Prevelige, which showed that empty capsids assemble via an initial nucleation step [8]. Electron microscopy and scattering experiments on a bacteriophage procapsid displayed a "nucleation- and growth" type mechanism of capsid assembly, as detailed in Figure 3. In this model, the first intermediate products have very few subunit-subunit interactions, implying little stability until a critical nucleus is formed. The critical nucleus was defined to be the intermediate with the smallest size such that the probability of assembly was over fifty percent. With a strong analogy to the correlation length, the critical nucleus would continue to grow as the free energy of subunit-subunit binding decreases. This would drive the assembly towards a half capsid.

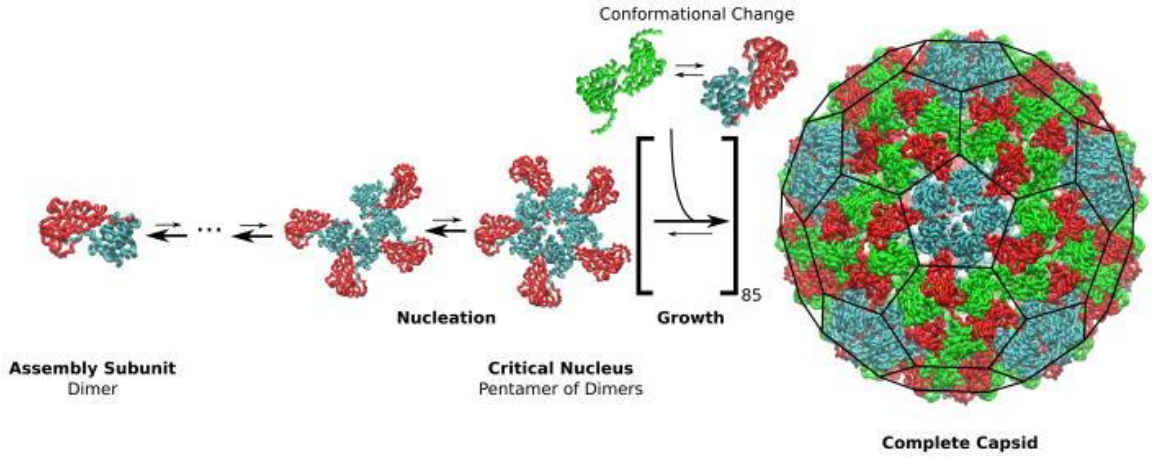

Figure 3: Reprinted from [8]. This describes the nucleation and growth kinetics observed in Prevelige's studies on bacteriophage.

In Lorman and Rochal's work, they suggest that their density wave model can only apply for models which can be explained fully in terms of physical models [9]. The fact that the density wave model does not predict this initial nucleation phase may simply be because the density waves did not account for biochemical behavior that occur in an actual assembly process. This biochemical behavior includes any large scale conformational change as time evolves, the presence of various amino acids on capsid proteins that might present additional forces, or any new bond formation that might occur.

4. Assembly of Loaded Capsids

In general, three competing interactions direct a virus towards a fully formed virion and maintain it in its metastable state. These interactions are as follows: a weak attractive interaction between capsid proteins (consists of hydrophobic and van der Waals forces), an electrostatic repulsion between capsid proteins, and a long range electrostatic attraction between positive tails of capsid proteins and nucleic acids [3]. Researchers described in the previous section simplified their study of assembly by reducing assembly to zero: they were able to do so with success for their systems.

Studying transitions in viral structures with non-negligible polyelectrolyte concentrations can present some interesting results, yet these studies can get much more complicated. Not only is there another force present, it seems that transitions in the viral capsid structure are also dependent on those in the nucleic acid. This will be explained in more detail in the following model.

4.1 Self-assembly of RNA viruses

As was described earlier, most studies on self-assembly of loaded capsids will make use of ssRNA as the internal nucleic acid. The logic behind this is simply that ds DNA's persistence length is too long for it to properly undergo a transition from coil to globule, a key step in providing a sufficient capsid-polyelectrolyte interaction.

Bruinsma et al. developed a comprehensive model of self-assembly [10] motivated by experiments performed on CCMV. The important argument they tackled in their article was whether the neutralization of capsid proteins by negatively charged vRNA (viral RNA) is the driving force to decrease electrostatic repulsion and onset assembly. This argument was found to be false on CCMV by experiments that showed removing the tails (effectively "neutralizing" them) of capsid proteins was not sufficient criteria for the onset of viral assembly [10].

The logic of this argument was proven invalid as well: since most RNA viral capsid proteins have a dipolar charge distribution, neutralization of positive internal charge will actually serve to increase the net negative charge which should increase electrostatic repulsions. Capsid assembly should statistically only occur when net attractive interactions outweigh repulsions; therefore, mere neutralization should not cause the onset of assembly. The researchers proposed that a different interaction between capsid proteins must take place, one that promotes a transition in the state of viral RNA [10]. The interaction they reasoned was one that promotes a "coil-to-globule" transition in the distribution of vRNA, where capsid proteins are the condensing agents [10]. Due to space limitations, this paper does not provide details on this particular polymeric phase transition.

Based on this coil to globule transition, Bruinsma et al proposed the phase diagram shown in Figure 4. If both CP (capsid protein)-vRNA interactions (ε) and CP-CP interactions (μ) are small, then vRNA should be uncondensed or swollen while CP should be predominantly free. At a certain critical ε^*, the solution vRNA would disproportionate into CP rich and CP poor populations which were held in thermodynamic equilibrium. Eventually, angle dependent CP-CP interactions (μ) and surface tension force CPs out of the vicinity of vRNA and towards the outside,

forming a structure with liquid. Characteristics of the provirion 1 structure include net negative internal ionic charge, as well as a few proteins that are adsorbed on the surface of the virion.

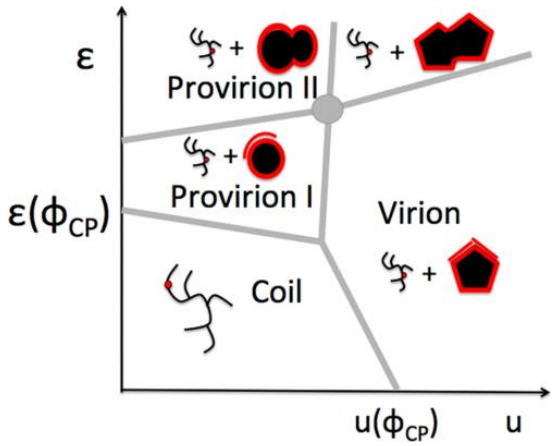

Figure 4: Reprinted from [10]. This is the phase diagram for small viruses proposed by Bruinsma et al. ε represents CP-vRNA interactions, and μ is CP-CP interactions, ε (Φcp) and μ(Φcp) are the critical points for each respectively.

If, however, ε is further increased while holding μ constant, this promotes another transition past the provirion 1 to provirion 2, a state characterized by a much larger surface area but with similar volume. In addition, this provirion 2 state is particularly interesting as it possesses very little to no surface tension and so the surface is highly susceptible to fluctuations. When solidified, the fully formed provirion 2 state would be expected to be misshapen and would not be expected to form a typical virion.

A particularly interesting finding from this phase diagram is how the behavior of assembly changes when ε passes over its critical point, ε*. For $\varepsilon > \varepsilon^*$, the assembly would not be expected to follow the Law of Mass Action. The logic behind this is that if the interactions between CP and vRNA are strong enough, then this would drive a high CP concentration into the CP vRNA aggregate even if there was not a large solution CP concentration. Thus, as μ is increased, the mechanism of the final assembly would not be dependent on the solution concentration. In addition, at a large ε, the assembly into a final virion state would pass by the provirion 1 state.

On the other hand, for $\varepsilon < \varepsilon^*$, the mechanism of assembly follows the typical Law of Mass Action and the kinetics should follow nucleation and growth phases. An empty capsid is the limit that ε goes to zero; thus, this model explains results that Prevelige achieved in his experiments on provirion bacteriophages [8].

(Go back and expand on these points as well as include the 2D phase transitions article+ include conclusion)

References:

1. W. H. Roos, R. Bruinsma, and G. J. L. Wuite, Nature Physics 6, 733 (2010).
2. H. Fraenkel-Conrat and R. C. Williams, Proceedings of the National Academy of Sciences 41, 690 (1955).
3. G. Tresset, J. Chen, M. Chevreuil, N. Nhiri, E. Jacquet, and Y. Lansac, Physical Review Applied 7, (2017).
4. V. L. Lorman and S. B. Rochal, Physical Review Letters 98, (2007).
5. J. D. Perlmutter and M. F. Hagan, Annual Review of Physical Chemistry 66, 217 (2015).
6. L. Landau, Nature 138, 840 (1936).
7. A. Zlotnick, Journal of Molecular Recognition 18, 479 (2005).
8. P. Prevelige, D. Thomas, and J. King, Biophysical Journal 64, 824 (1993).
9. V. L. Lorman and S. B. Rochal, Physical Review B 77, (2008).
10. R. F. Bruinsma, M. Comas-Garcia, R. F. Garmann, and A. Y. Grosberg, Physical Review E 93, (2016).