MIT–LCS Technical Memo #TM-527

Local Rule Switching Mechanism for Viral Shell Geometry

Bonnie Berger* Peter W. Shor†

June 16, 1995

Abstract

In a previous paper [Berger et al., PNAS **91** 7732, 1994], a theory of virus shell formation was proposed in which shell assembly is directed by local interactions of the coat and scaffolding subunits. This theory requires that the same chemical subunits assume different, stable conformations depending on their position in the shell. During assembly, the conformation of a protein subunit dictates the conformations of its neighboring subunits. It was shown that these local interactions could be designed so as to generate shells that have the same geometric structure as virus capsids. Different sets of local interactions, or local rules, were designed to produce different final shell geometries. In this paper, local rules are given that assemble a T = 7 shell such that a small change in these rules produces a T = 4 shell. This is intriguing since evidence has been accumulating that some T = 7 shells are closely related to T = 4 shells. These local rules also predict that hexamers in the assembled procapsid would have approximate two-fold rotational symmetry. This symmetry is exemplified by the elongation of hexamers observed in many T = 7 viruses. These rules also provide a possible explanation for spiraling and tubular malformations.

*Mathematics Department and Laboratory for Computer Science, MIT.
†AT&T Bell Laboratories, Murray Hill, NJ.

1 Introduction

Icosahedral virus shells are constructed of repeated protein subunits, or coat proteins, which surround their condensed DNA or RNA genomes. Many of these shells are believed to assemble with only limited aid from cellular machinery; they "self-assemble," or spontaneously polymerize and take shape, in the host cell environment. Sometimes assembly is assisted by scaffolding proteins, which assemble with the coat proteins to form a precursor shell, but are removed before the shell matures. At first glance, the assembly of the shells seems easy to understand, because the structure is so regular. In fact, it has been difficult to determine the actual pathway through which the subunits interact to form a closed shell composed of hundreds of subunits [26]. In icosahedral viruses this has been particularly difficult to explain because very often the same coat protein occurs in non-symmetric positions [26].

In Berger *et al.* [5], a hypothesis for icosahedral viral shell assembly mechanisms was introduced. This hypothesis, the local rule theory of virus shell assembly, proposes that the assembly of virus shells is directed by the local interactions of its coat and scaffolding protein subunits. The hypothesis requires that these subunits can assume different conformations. During assembly, the conformation of a protein subunit in the shell dictates the conformation of its neighboring subunits when they join the shell. It was shown that these local interactions could be designed so as to generate shells having the same geometric structure as observed in actual virus capsids. Different sets of local interactions, or local rules, can be designed which will produce different final shells, as well as malformed shells. Varying the angles and interaction lengths associated with a set of local rules by small amounts can change the shape of the final shell while keeping the same basic geometric structure. The nature of the rules is such that by changing which conformations bind to each other, one can change the geometric structure of the shell, and thus produce shells with various T-numbers.

For some time, evidence has been accumulating that $T = 7$ and $T = 4$ shells are in some cases closely related. The most recent indications that this is true arise in experiments with bacteriophage P2 and its satellite phage P4 [21]. The capsid of bacteriophage P2 is a $T = 7$ icosahedral structure formed from a phage protein gpN. When the bacteria is coinfected with P4, P4 capsids are also formed which have $T = 4$ icosahedral structure but also use the P2 coat protein gpN. The P4 protein that influences size was determined to be gpSid [28, 1]. Since gpSid is present in immature P4 particles, but not in the mature virus, it is believed to act as a scaffolding protein [2, 12]. Similarly, a P2 protein gpO is believed to act as a scaffolding protein [18, 8, 6]. Marvik *et al.* [21] showed that the capsid size is indeed determined by these auxiliary proteins gpO and gpSid. When P2 gene N was cloned and inserted into bacteria without these auxiliary proteins, gpN produced predominantly irregularly shaped malformations, but also a small number of spherical shells of both $T = 4$ and $T = 7$ structure. When the P2 protein gpO was also introduced, gpN formed predominately $T = 7$ structures. When a P4 bacteriophage infection was present, but gpO was absent, the protein gpN formed predominately $T = 4$ shells. Finally, with both gpO and P4 infection present, both size shells were formed, but with many fewer malformations than with gpN alone.

Another indication that $T = 4$ and $T = 7$ shells are related is that in bacteriophage λ, which also has a $T = 7$ capsid, mutations of the coat protein exist that form functional $T=4$ shells [16]. A third indication of this relationship is that one of the common mistakes observed in the assembly of the phage P22 in the absence of scaffolding proteins is the formation of a $T=4$ shell instead of a $T=7$ structure [10]. In fact, the structures resulting from assembly of pure P22 coat protein are quite similar to those resulting from pure P2 coat protein [10, 22, 21], possibly indicating that these two viruses have similar assembly mechanisms, even though the sequences of these two proteins do not appear very similar [19].

If these icosahedral capsids indeed form by following a set of local rules, then this set of local rules should have the property that a relatively small perturbation of the rules directing a $T = 7$ shell should produce a $T = 4$ shell. In this paper, we give such a set of rules. It would be nice if this set of rules could completely predict the structure of the virus. Unfortunately these rules are at too high a level of abstraction to predict exactly where the scaffolding molecules should be in the structure. However, these rules do imply that the coat and scaffolding should have certain properties which have been observed in several $T = 7$ viruses as well as the P2 and P4 viruses. Similarly, these rules give constraints on the placement of the scaffolding molecules in these viruses. The local rules given here also account for the elongation of hexamers observed in phage P22 procapsids [22].

2 Theory

2.1 Basic model and assumptions

The *theory of quasi-equivalence* [9] observed that the same protein subunits are not quite equivalent due to differences in the global symmetry, but still it was assumed they really have almost equivalent conformations and nearly identical chemical properties. Since then, evidence has been accumulating that the subunit conformations can be much more flexible [13]. A local rule theory postulates that the coat and scaffolding proteins can assume a number of distinct conformations, and thus it is not bound by the notion that these proteins have identical properties.

A local rule theory is a set of local interactions telling which conformations are allowed to bind to each other, as well as the admissible interaction lengths and relative binding and torsional angles between these interactions. For every T number, it is possible to find a set of local rules producing that T number and using a number of protein conformations that is equal to the T number [5]. It is sometimes also possible to find a different set of local rules resulting in the same shell geometry with fewer conformations by using rules that assign the same conformation to non-equivalent positions.

The conformations used for the assembly rules need not be the same as the conformations in the final shell. The set of rules given below for a $T = 7$ shell assumes only four distinct conformations of the subunits. The local rules require that these four subunit conformations behave differently during the assembly process; however, it does not follow that there are four conformations in the final shell. There could be more, since there are seven non-equivalent positions in the final shell, and the resulting non-equivalent neighborhoods could induce two subunits that behaved similarly during assembly to assume different final conformations. There could also be fewer, since the distinctness of conformations during assembly need not be maintained in the final shell. In phage P22, for example, the conformations of coat proteins in the final capsid appear more similar then the conformations in the procapsid [22].

One of these four conformations will be considered first. Figure 1(a) gives the rule for how one of the four conformations, the type 1 conformation, chemically binds in three dimensions. A type 1 conformation has a binding site for a type 2 conformation, and two binding sites for type 1 conformations. Given the binding interaction to the type 2 neighbor, then, at a position clockwise from this at an angle of about 135°, only a type 1 conformation can attach. Similarly, only a type 1 conformation can attach at an angle of about 108° from this latter binding interaction. We call this representation the type 1 local rule. Note that although the type 1 and 2 conformations are represented by circles (or parts of circles), their labels (i.e., 1 and 2) are intended to imply assymetric conformations. Note also that the angles for the rule do not have to add up to 360° because the rule is three-dimensional.

Similar local rules can be constructed for all the four conformations in a $T = 7$ shell (Figure 1). More than one rule is given for certain conformations because for this set of local rules one conformation needs to be able to bind to two alternate conformations in certain directions. The binding interactions in the local rules need to be present in the shell; however, additional interactions may also be present which would have only a secondary effect on the assembly process.

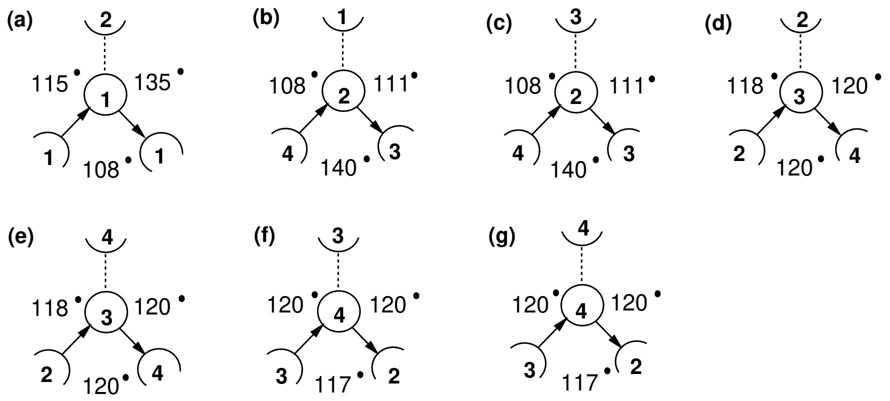

Figure 1: Possible local rules for a left-handed $T = 7$ virus. Each protein subunit is represented as a circle or part of a circle labeled with its conformation. Binding interactions are represented possibly with an associated direction, indicating an interaction that is asymmetric. Angles between binding interactions are the approximate number of degrees between the centers of the protein subunits. Angles are not based on any particular virus, but are derived from a computer simulation.

There are various additional constraints imposed upon this set of rules. For these rules, it is assumed that the shell is initiated at a pentamer. It is also assumed that a protein does not assume its final configuration until there is at least one other protein in the same capsomere. That is, when a protein attaches itself to the growing shell at a point when it does not have an adjacent subunit in the same capsomere, it remains in a state of conformational flexibility and does not yet adopt a final conformation. When another subunit attaches next to it, thus creating two points of attachment for the new capsomere, the two proteins will fall into a low energy configuration and stop fluctuating.

One additional constraint needs to be imposed in addition to the rules already given. During the application of these rules, there is one point during the growth of the shell where there are two possible configurations which are both allowed by the rules given. In order to consistently build $T = 7$ shells, it is necessary to disallow one of them and thus consistently choose the other configuration. In this section, we will not go into the possible biochemical mechanisms for choosing one of these two configurations. However, we will show that if the configuration in Figure 3(a) is consistently chosen, $T = 7$ shells will result, and if the one in 3(b) is consistently chosen, $T = 4$ shells will result. Consequently, the Figure 3(b) configuration is called the *disallowed configuration* for $T = 7$ shells.

We now give a step-by-step explanation of the construction of these shells. Suppose the rules in Figure 1 have been applied to generate the portion of the shell in Figure 4(a). Then by applying the rules in Figures 1(d) and 1(g) to the shell, the portion of the shell in Figure 4(b) results. Notice that

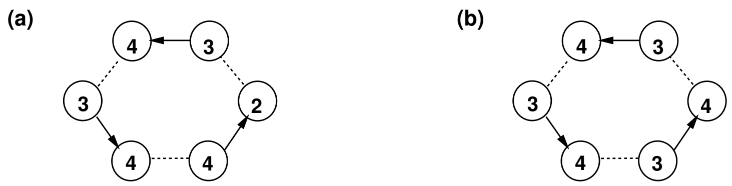

Figure 3: (a) This configuration is allowed. (b) This configuration is not allowed.

the shell now includes the configuration given in Figure 3(a). Alternatively, by applying the rules in Figures 1(e) and 1(f) to the shell in Figure 4(a), the portion of the shell in Figure 4(c) results. However, here the shell includes the disallowed configuration for $T = 7$. This occurs at what we call a "quasi three-fold" axis of symmetry. By consistently choosing between the two configurations in Figure 3 in this manner, it is possible to generate either a $T = 7$ shell (Figure 5(a)) or a $T = 4$ shell (Figure 5(b)).

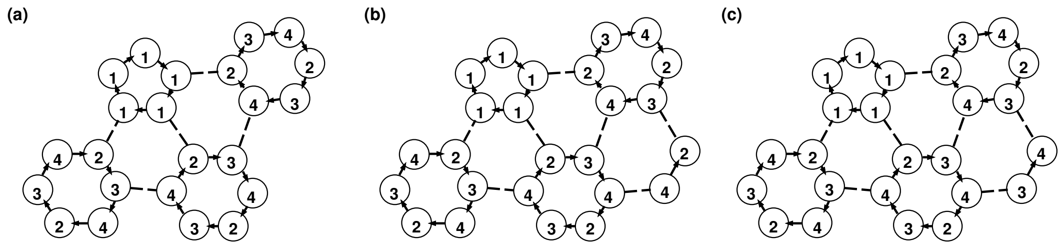

Figure 4: (a) Partial growth of a shell according to the rules in Figure 1. (b) Application of the rules in Figure 1 to the partial shell in (a). (c) Another application of the rules in Figure 1 to the partial shell in (a) yields the disallowed configuration for $T = 7$ shells.

Some care must be taken in the order of adding coat proteins to the virus shell. If, for example, the proteins were added in an order to form a long chain with no interconnections, while this scenario is allowed by the rules, it would not produce the desired icosahedral shell. For the set of $T = 7$ rules, the following guidelines are adequate to consistently produce the correct final structure. The guidelines are:

1. Whenever a new capsomere (i.e., hexamer or pentamer) is started, it should be finished before new capsomeres adjacent to it are started. Note that the conformations of all the proteins in a capsomere are determined by the conformations of the first proteins added to this capsomere.

2. If possible, when starting a new capsomere, add it to a point on the shell where it can attach to two finished capsomeres. This clearly is not possible for the first or second capsomeres added to the shell, but is possible thereafter.

By adjusting the kinetics, it should be possible to make the coat and scaffolding proteins for the most part obey these guidelines. That is, reactions which add proteins to an existing capsomere should be favored over ones starting a new capsomere, and new capsomeres should preferentially be started at spots adjoining two existing capsomeres. The requirement that proteins do not assume

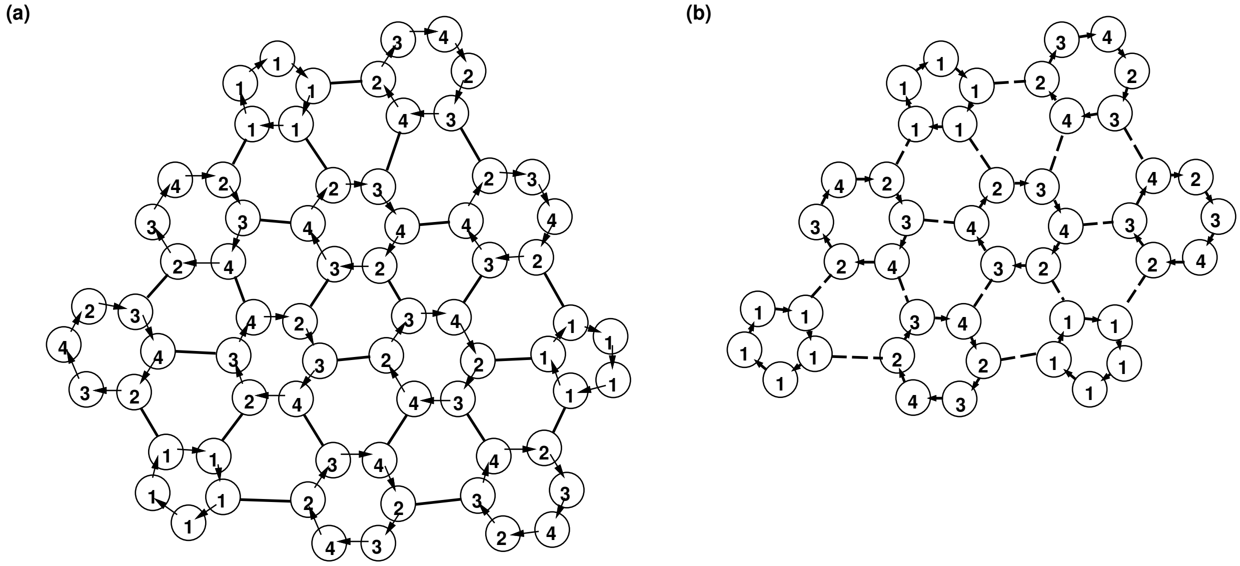

Figure 5: (a) Portion of a $T = 7$ shell produced by the rules in Figure 1. (b) Portion of a $T = 4$ shell produced by the rules in Figure 1 when the disallowed configuration is consistently allowed.

their final conformation until there are two proteins present in the same capsomere matches well with these requirements for kinetics.

For the set of rules in Figure 1, orders of adding proteins consistent with a weaker set of guidelines will also produce the final structure. In particular, as long as proteins are only added onto a protein that is not in conformational flux, then $T = 7$ shells are formed.

2.2 Implications for virus structure

The set of local rules for $T=7$ in Figure 1 is nearly the same as a set of local rules for $T=4$: By consistently choosing the disallowed configuration in Figure 3 (which is what prevents this set of rules from forming $T=4$ shells), the set of rules is restricted to those for a $T=4$ shell with four conformations [4] (*i.e.,* Figures 1(a), 1(b), 1(e), and 1(f)).

In the set of $T=7$ rules, the hexamers are symmetric under rotations of 180°. This symmetry is used in reducing the number of conformations in the rules from seven to four. By using symmetric hexamers, we reduce the number of conformations in a hexamer from six to three. This does make the interaction used to determine the $T = 7$ shell more complicated; however, it reduces the complexity of the number of conformations a protein subunit need assume.

2.3 Possible roles for the scaffolding proteins

One of the key properties of the rules given in Figure 1 is that the conformations of the proteins in the hexamers have 180° symmetry. This reduces the number of coat protein conformations necessary to implement local rules for a $T = 7$ shell from seven to four by putting proteins on the opposite sides of a hexamer into the same conformation. Because of the symmetry, the conformation of any one protein in the hexamer prescribes the conformations for the remaining proteins in the hexamer; this enables the shell to assemble without introducing too much ambiguity in the process. One possible role for the scaffolding protein is to provide this symmetry; that is, to break a natural six-fold symmetry associated with coat protein hexamers. Although it is not mandated biochemically, it seems likely that a hexamer of coat protein would have six-fold symmetry. If either two or four scaffolding molecules naturally bind with this hexamer of coat protein molecules, they could push the hexamer into a configuration having two-fold but not six-fold rotational symmetry. This symmetry is exactly what is required for the set of rules in Figure 1.

This two-fold symmetry on hexamers allows predictions of the number of scaffolding proteins in a virus coat. We believe that this number should be close to one of the three values 120, 240, or 360. Because the rules are symmetric about hexamers, the positions of the scaffolding proteins should also be symmetric about the hexamers. Since the hexamers have two-fold rotational symmetry, the number of scaffolding proteins associated with a hexamer must be even. The most likely values are two, producing a total of 120 scaffolding molecules associated with the hexamers, or four, producing a total of 240. In this latter case, there might also be five scaffolding molecules associated with pentamers, producing 300 scaffolding molecules total. We believe that the number of scaffolding molecules associated with a hexamer is unlikely to be six because the scaffolding molecules may be needed to break six-fold symmetry, as explained earlier in this section; numbers of eight or higher seem too large to be likely.

In the above discussion, we assumed that the scaffolding molecules associated with pentamers had five-fold symmetry, as do the pentamers themselves in the local rules. Except in cases where scaffolding molecules are involved in initiating the shell growth [29], there does not seem to be any necessary role in shell assembly for the scaffolding molecules associated with a pentamer. Even though scaffolding molecules need not associate with a pentamer in order to direct virus shell formation, they could still occupy possible binding sites, which might occur in the coat protein at both hexamers and pentamers. If there are two scaffolding proteins per hexamer, it seems that it would be unlikely that five scaffolding proteins could fit near a pentamer, but if scaffolding molecules are not required to have five-fold symmetry at a pentamer, there might be two scaffolding molecules per hexamer and one or two per pentamer, resulting in between 130 and 144 scaffolding molecules altogether. Similarly, if there are four scaffolding molecules per hexamer, there could be three to five per pentamer, resulting in slightly fewer than 300 scaffolding molecules total.

Another possible role for the scaffolding molecules is to enforce the choice of configuration in Figure 3. The experimental evidence on conformational switching in the P2/P4 virus system indicates that the scaffolding proteins in these viruses have this duty. In the case of P2, the scaffolding must reach across the quasi three-fold axes to interact in some way with scaffolding or coat proteins around them and force the configuration in Figure 3a. In the case of P4, this would correspond to the scaffolding reaching across the three-fold axes of symmetry.

It is difficult to predict exactly where the scaffolding protein might actually be from this set of local rules. For $T = 7$ capsids, when there are four scaffolding proteins per hexamer, if the hexamers are stretched, an attractive hypothesis is that the scaffolding proteins form a roughly symmetric tetramer binding to the four coat proteins that come closest to forming a square. In phage P22, the hexamers are stretched along the 3-3 conformation axis of Figure 5a, so this would indicate the scaffolding proteins were bonded to the 2 and 4 conformations. An opposing argument would be that the scaffolding proteins need to interact across the quasi three-fold axis that might produce the disallowed configuration, which would indicate that they are bonded to the 3 and 4 conformations of the hexamers.

Scaffolding protein is not actually necessary to implement this set of rules. The chemical properties of the coat protein itself might force it into hexamers with two-fold but not six-fold symmetry. Since the three type 4 conformations in the disallowed conformation could easily be spatially adjacent, they may form a trimer that has higher energy than the 4-4-2 conformation trimer and thus is energetically unfavorable (see Figure 3). Alternatively, the disallowed configuration might be impossible because the sum of the bonding angles around it is too large or too small to allow it to close, although this scenario seems less likely.

2.4 Spiraling malformations

The coat protein for P2, gpN, forms spiral structures, and "telephone"-shaped structures, where two spirals, often having different diameters, are connected, especially in the absence of scaffolding protein. The local rules provide a possible explanation for the standard spiral structure [5]; in particular, if a hexamer occurs at a 5-fold axis of symmetry, spiraling occurs in computer simulations (Figure 6). One may also suggest that once this initial flat region is formed, it is more likely that this flat region gives rise to other such "mistakes", producing an even larger flat region. When pentamers are finally added to the flat region, they introduce curvature. Then if the local rules are correctly followed thereafter, the "telephone"-shaped structure could result. If pentamers are never added to the large flat region, then tubular malformations would be likely to result.

Another possible explanation for both types of spirals arises when the constraint that the application of the rules in Figure 1 must start with a pentamer is violated. Suppose that the shell started building with a hexamer. Then it would be ambiguous how to build out from this initial hexamer, whereas it is not ambiguous how to build out from a pentamer. Furthermore, it would be ambiguous when to add a pentamer to the growing structure. Then large flat regions could result with the same consequences as those discussed above.

It still remains to address why different diameter spirals (one corresponding to a T=4 shell and the other to a $T = 7$ shell spiral) may result in a telephone-shaped spiral. When present, the scaffolding proteins gpO and gpSid may attach to a pentamer and then attach only to themselves (see discussion below). In the absence of scaffolding protein, the switching mechanism may be responsible for initiating one type of spiral or the other.

2.5 Switching mechanisms for other T numbers

Similar switching mechanisms seem to exist for other T numbers. There is a set of local rules with seven conformations where $T = 13$ is related to $T = 7$ shells in much the same way that the $T = 7$ set of rules presented in this paper is related to $T = 4$ shells (Figure 7). These rules have the same additional constraints as those imposed on the rules for the $T = 7$ shell (Figure 1), except the disallowed configuration which occurs at a quasi three-fold axis of symmetry is different. Consistently choosing the disallowed configuration will restrict the rules chosen to precisely those rules for a $T = 7$ capsid with seven conformations (*i.e.*, Figures 7(a), 7(b), 7(d), 7(f), 7(h), 7(i), and 7(k)) [5].

It would be interesting to know whether there is any biological evidence for the relation between $T = 13$ and $T = 7$ shells. Similar relationships between other pairs of T-numbers would also be interesting.

3 Discussion

Experimentally [21], the P2 coat protein assembles into fewer malformations in the presence of both scaffolding proteins than it does in the presence of either alone. This seems counter-intuitive, because it might be the case that these scaffolding molecules interfere with each other by both bonding to the same shell. Since one type makes $T = 4$ shells and the other $T = 7$ shells, a shell containing both types of scaffolding particles would seem very likely to be malformed. One possible reason this does not happen would be if, in P2 and P4, all the scaffolding proteins are internally linked. In this case, the first scaffolding protein to bind with a forming shell would force the other scaffolding proteins to be of the same type, because the other type could not bind with a scaffolding protein of the first type. This would explain why the presence of both kinds of scaffolding molecules does not lead to more malformations than in the presence of either alone; it may result in fewer malformations because there are more likely to be scaffolding molecules present at initiation, and the absense of scaffolding molecules during the initiation step could lead to malformations. We thus believe that it is likely that the gpO and gpSid particles in P2 and P4 form a connected inner scaffold.

After the above paragraph was written, we learned that Lindqvist *et al.* had found that the gpSid molecules in P4 form a connected *outer* scaffold [20]. This gives a better explanation of why there are fewer malformations in the presense of both scaffolding proteins than in the presense of either alone: in P4, both scaffolding proteins are simultaneously present during assembly, so the assembly is being stabilized by both scaffolds cooperatively (one internal and one external), which apparently works better together than either does alone. The gpSid molecules do indeed appear to be connected to the shell at the three-fold axes of symmetry (as well as other points), which is where our theory predicts they should operate to direct the assembly of a $T = 4$ rather than a $T = 7$ shell.

Recall that we needed the initiation complex to be a pentamer. This assumption seems realistic since for several bacteriophages, an initiation complex located at one of the pentamers is required to guarantee formation of the desired structure in vivo [29]. Further evidence in support of this assumption appears in [25], which gives evidence that *in vitro* initiation of P22 capsids occurs at a pentamer.

Conformational flexibility is required for the sets of rules given here, since proteins do not assume their final conformation until after having first bound and also waited for a neighboring protein to attach. However, this is quite reasonable for real viral coat proteins (at least for P22) since by Galisteo et al.'s results [13], these proteins are quite flexible in solution.

The local rule theory is described in terms of conformationally flexible monomers binding to a growing shell. Whereas certain bacteriophages such as P22 and λ are known to form from monomers, the theory need not be so restrictive. Similar local rules can also be given that build capsids from the association of dimer and capsomere building blocks. These rules also make use of 2-fold symmetry in the hexamers to direct assembly.

The requirement in the rules for two-fold rotational symmetry of hexamers is intriguing because the micrographs of P22 show near-symmetry of the hexamers under $180°$ rotations [22]. This is seen in the elongation of the hexamers of P22 precursor virus along the axis connecting the two type 3 conformations.

Our estimates for numbers of scaffolding proteins are consistent with experimental data. The $T = 7$ bacteriophage T7 procapsid is believed to contain only approximately 140 scaffolding proteins [27], which matches well for the case when two scaffolding proteins bind to each hexamer. The bacteriophage P22 procapsid has approximately 300 scaffolding proteins [17, 7, 11], which would indicate that four scaffolding proteins are associated with each hexamer and five with each pentamer. That scaffolding proteins are indeed associated with the pentamers in P22 is indicated by evidence that scaffolding is involved in initiation of shells *in vitro* [25] and *in vivo* [3]. Further, Greene and King [14] have shown that when scaffolding protein is added to empty P22 precursor capsids, the scaffolding rebinds in two stages, a fast stage comprising approximately half of the 300 scaffolding proteins per capsid, and a slow stage where the remaining scaffolding proteins bind. It would be consistent with our hypothesis that there are 120 scaffolding proteins tightly bound to the coat (two per hexamer), whereas in the remaining binding sites the scaffolding subunits are bound much more loosely. These 120 sites with tightly bound scaffolding protein would likely be the important sites for control of virus assembly. Further evidence for 120 important sites for virus assembly comes from work of Prevelige *et al.* [23], which shows that the minimum number of scaffolding molecules required for assembly approaches the value of 130 or so, in good agreement with 120 tightly-bound sites which are necessary for control of assembly.

The arrangement of four scaffolding proteins associated with each hexamer may explain recent results of Prevelige [24] that P22 scaffolding protein forms tetramers in solution. These results show that at low concentrations, scaffolding proteins are monomers in solution, and with increasing concentration they first form dimers and then tetramers.

For P22, the differences between the protein conformations in non-equivalent positions are noticeably less in the mature form than in the precursor form [22]. Perhaps functionally different protein conformations are required for assembly, while in the mature form all the proteins assume the same functionality and need only be different enough to hold the shell together stably. This may also be the reason for the substantial changes that other bacteriophages undergo between their precursor and mature forms [15].

Whereas in P4 and P2, scaffolding proteins appear responsible for the switch between $T = 4$ or $T = 7$ shells, in λ, several mutations of the coat protein produce $T = 4$ shells [16]. This may indicate that the role of scaffolding in the local rules differs in these viruses. Whereas in P22 and P2 the choice of hexamer as in Figure 3 is enforced by the scaffolding, in λ it may be enforced through interactions between coat proteins. Alternatively, the mutation of λ producing $T = 4$ shells may be in the section of the coat protein responsible for interacting with the scaffolding, in which case the scaffolding molecule could still be involved in directing the choice of $T = 7$ rather than $T = 4$ shells.

Whereas some bacteriophages, such as T4, have a separate coat protein for the pentameric vertices, most $T = 7$ bacteriophages only have one protein. This set of local rules would give an explanation for this phenomenon. If there were a separate coat protein for the pentameric vertices, this protein would have to be able to distinguish a potentially penetameric site from a hexameric site when joining to the forming shell. For the local rule theory given in this paper, this distinction sometimes does not occur until another coat protein joins adjacent to the first. This would indicate that for this set of local rules, the same protein needs to be able to take on both hexameric and pentameric conformations.

The relationship between the two sets of rules for $T = 7$ and $T = 4$ may indicate a pathway for the evolution of the assembly mechanism of a $T = 7$ virus. Since these two sets of rules are relatively similar, a $T = 4$ virus capsid protein with rules involving four conformations may have mutated to produce a larger $T = 7$ capsid. The relationship between $T = 7$ and $T = 13$ shells may similarly have led to the evolution of $T = 13$ bacteriophages.

Acknowledgements

We are grateful to Jonathan King and Peter Prevelige for their assistance with this work. Russell Schwartz's programming assistance has been a great help. B.B. was supported in part by an Arthur D. Little Career Development Award and ARPA grant #N00014-92-J-1799.

References

[1] M. Agarwal, M. Arthur, R. D. Arbeit, and R. Goldstein. Regulation of icosahedral virion capsid size by the *in vivo* activity of a cloned gene product. *Proc. Natl. Acad. Sci. U.S.A.*, 87:2428–2432, 1990.

[2] K. J. Barrett, M. L. Marsh, and R. Calendar. Interactions between a satellite bacteriophage and its helper. *J. Mol. Biol.*, 106:683–707, 1976.

[3] C. Bazinet and J. King. Initiation of P22 procapsid assembly in vivo. *J. Mol. Biol.*, 202:77–86, 1988.

[4] B. Berger and P. W. Shor. On the mathematics of virus shell assembly. Technical Report 519, MIT Lab. for Computer Science, April 1995.

[5] B. Berger, P. W. Shor, L. Tucker-Kellogg, and J. King. Local rule-based theory of virus shell assembly. *Proc. of the Natl. Academy of Sci.*, 91(16), August 1994.

[6] E. Bertani and E. W. Six. The P2-like phages and their parasite, P4. In R. Calendar, editor, *The Bacteriophages*, volume 2, pages 73–143. Plenum Press, New York, 1988.

[7] D. Botstein, C. Waddell, and J. King. Mechanism of head assembly and DNA encapsulation in Salmonella phage P22. I. Genes, proteins, structures and DNA maturation. *J. Mol. Biol.*, 80:669–695, 1973.

[8] S. Casjens and R. W. Hendrix. Control mechanisms in dsDNA bacteriophage assembly. In R. Calendar, editor, *The Bacteriophages*, volume 1, pages 15–91. Plenum Press, New York, 1988.

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

[10] W. Earnshaw and J. King. Structure of phage P22 coat protein aggregates formed in the absence of the scaffolding protein. *J. Mol. Biol.*, 126:721–747, 1978.

[11] K. Eppler, E. Wyckoff, J. Goates, R. Parr, and S. Casjens. Nucleotide sequence of the bacteriophage P22 genes required for DNA packaging. *Virology*, 183:519–538, 1991.

[12] C. G. Fossdal. *GpSid Functions as a transient, P4 capsid size determining scaffold.* PhD thesis, University of Oslow, Norway, 1991.

[13] M. L. Galisteo, C. L. Gordon, and J. King. Stability of wild type and temperature sensitive protein subunits of the bacteriophage P22 capsid. *J. Biological Chemistry*, 1995. In press.

[14] B. Greene and J. King. Binding of scaffolding subunits within the P22 procapsid lattice. *Virology*, 205:188–197, 1994.

[15] R. W. Hendrix. Shape determination in virus assembly: The bacteriophage example. In S. Casjens, editor, *Virus Structure and Assembly*, pages 169–204. Jones and Bartlett, Boston, MA, 1985.

[16] I. Katsura. Structure and inherent properties of the bacteriophage lambda head shell IV: Small-head mutants. *J. Mol. Biol.*, 171:297–317, 1983.

[17] J. King and S. Casjens. Catalytic head assembling protein in virus morphogenesis. *Nature*, 251:112–119, 1974.

[18] J. A. Lengyel, R. N. Goldstein, M. Marsh, M. G. Sunshine, and R. Calendar. Bacteriophage P2 head morphogenesis: Cleavage of the major capsid protein. *Virology*, 53:1–23, 1973.

[19] N. A. Linderoth, R. Ziermann, E. Haggard-Ljunguist, G. E. Christie, and R. Calendar. Nucleotide sequence of the DNA packaging and capsid synthesis genes of bacteriophage P2. *Nucleic Acids Res.*, 19:7207–7214, 1991.

[20] B. H. Lindqvist, T. Dokland, R. N. Hanoa, E. Jacobsen, T. Larsen, and O. J. Marvik. Sid determines the size of phage P4 capsid by imposing the formation of an external procapsid scaffold during assembly. *J. Mol. Biol*, 1995. To appear.

[21] O. J. Marvik, P. Sharma, T. Dokland, and B. H. Lindqvist. Bacteriophage P2 and P4 assembly: alternate scaffolding proteins regulate capsid size. *Virology*, 200:702–714, 1994.

[22] B. V. V. Prasad, P. E. Prevelige, Jr., E. Marietta, R. O. Chen, D. Thomas, J. King, and W. Chiu. Three-dimensional transformation of capsids associated with genome packing in a bacterial virus. *J. Mol. Biol.*, 231, 1993.

[23] P. Prevelige, D. Thomas, and J. King. Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells *in vitro*. *J. Mol. Biol.*, 202:743–757, 1988.

[24] P. Prevelige, Jr. Personal communication.

[25] P. E. Prevelige, Jr., D. Thomas, and J. King. Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. *Biophys. J.*, 64:824–835, March 1993.

[26] M. G. Rossmann. Constraints on the assembly of spherical virus particles. *Virology*, 134:1–13, 1984.

[27] P. Serwer and R. H. Watson. Internal proteins of bacteriophage T7. *J. Mol. Biol.*, 194:271–291, 1982.

[28] D. Shore, G. Deho, J. Tsipis, and R. N. Goldstein. Determination of capsid size by satellite bacteriophage P4. *Proc. Natl. Acad. Sci. U.S.A.*, 75:400–404, 1978.

[29] J. M. Valpuesta and J. L. Carrascosa. Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Quart. Rev. Biophys.*, 27:107–155, 1994.

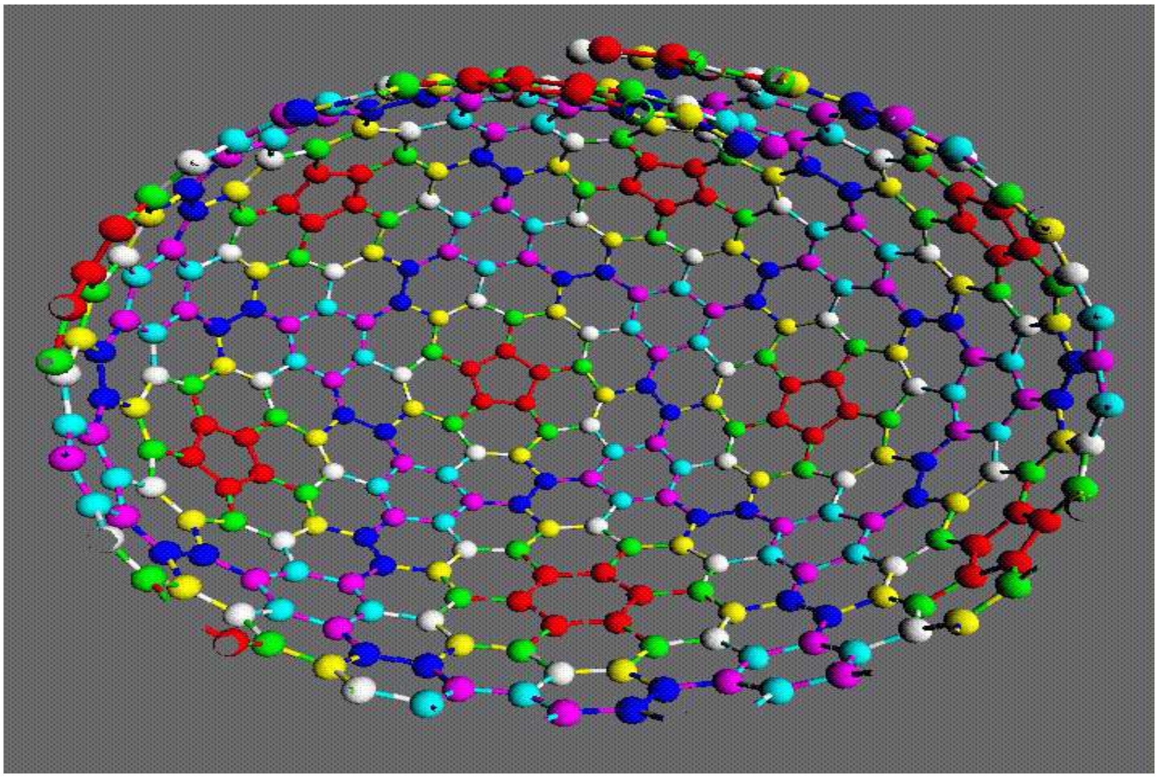

Figure 6: *Silicon Graphics Indigo 2* computer graphics image of slice of a spiraling malformation.

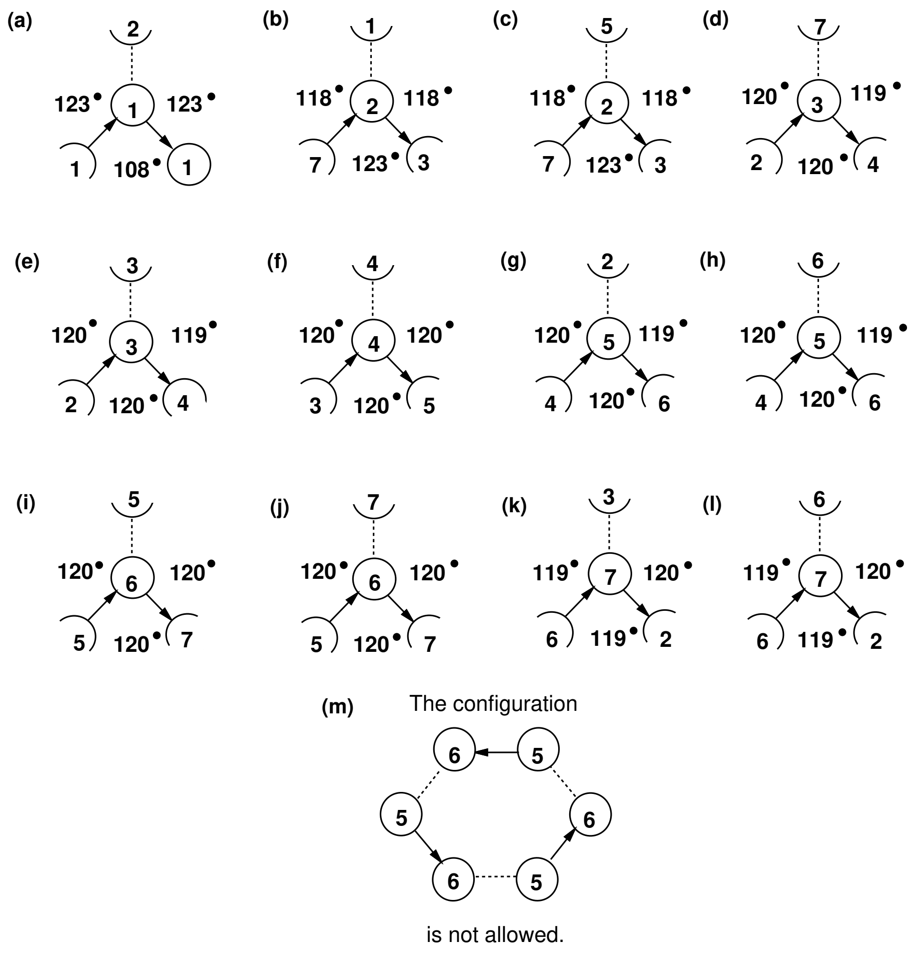

Figure 7: Possible local rules for a left-handed $T=13$ virus. The disallowed configuration is given with the rules.