Proc. Natl. Acad. Sci. USA
Vol. 75, No. 10, pp. 4779–4783, October, 1978
Biochemistry

Symmetry mismatch and DNA packaging in large bacteriophages

(virus assembly/biological rotation)

ROGER W. HENDRIX

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Communicated by J. D. Watson, July 19, 1978

ABSTRACT **A model is presented for the mechanism of packaging double-stranded DNA into phage heads. The model is based on, and rationalizes, the mismatch in symmetry between the heads and tails of large bacteriophages. DNA movement is postulated to be mediated by a rotating protein structure at the tail-proximal vertex of the head.**

A long-standing problem of virus structure concerns the fact that the symmetries of the heads and tails of large bacteriophages are different. All bacteriophage tails for which adequate data are available are now known to have 6-fold rotational symmetry, and they attach to the icosahedral head at one of the axes of 5-fold rotational symmetry (1–5). I propose here a general model for the structure of the large double-stranded DNA-containing phages that accounts for the symmetry mismatch and argue that the symmetry mismatch makes available an interesting new class of models for how double-stranded DNA is packaged into phage heads. The phages to which this discussion applies include λ, T4, P22, P2/P4, T7, T3, T5, φ29, and others. Casjens and King (6) have recently reviewed the general features of the structure and assembly of these viruses.

Phage structure

A structural feature that is probably common to all the phages under consideration is a small protein knob that lies just inside the head membrane at the unique corner of the icosahedral head where the tail attaches. This structure, which I will call the connector, is assembled as part of the head and, in at least several phages, has been implicated in the earliest steps of head assembly. The phage tail attaches directly to the connector after DNA packaging. Connectors appear to have a roughly cylindrical shape, and in some cases (λ, T7, T4, P22) an axial hole has been visualized oriented along the head–tail axis (4, 7, 8; W. C. Earnshaw, personal communication). During injection, the DNA passes out of the head and into the tail through the corner where the connector lies, and it is generally assumed that the DNA passes out through the connector. It is also likely that the DNA passes in through the connector when it is being packaged into the head.

Although it is clear that tails have a different symmetry from the icosahedral head membrane, it has never been established whether the interface between the 5- and 6-fold symmetries is at the junction between the tail and the connector or at the junction between the connector and the head membrane. [A third possibility (1)—that the connector is an adaptor with both 5- and 6-fold symmetry (i.e., with 30-fold symmetry)—seems ruled out for at least many of the phages because no proteins are present in the required 30 copies.] The symmetry of the connector has not been determined directly for any of the phages. Table 1 lists the available measurements of the numbers of copies of the proteins in connectors for several phages. Most of the numbers can plausibly be interpreted as multiples of either five or six copies, and none of the data rule out either hypothesis conclusively.

I propose that the connector has 6-fold symmetry and therefore that the symmetries of connectors and tails match but the symmetries of connectors and head membranes do not. Although a conclusive argument for this point of view cannot yet be made, it seems more attractive than the alternative for the following reason. The tail attaches to the connector after head assembly is complete; it must be held in place primarily by the noncovalent bonds it forms with the connector. This attachment would probably be more secure if the symmetries of the tail and connector matched than if they did not. On the other hand, the connector is assembled into the head very early in the assembly pathway and is probably too large to pass out through the hole in the vertex of the head (unpublished results). It might be held in the head primarily by steric constraints, like a button in a buttonhole. If this were so, then the relative weakness of the bonding between connector and head membrane that might be expected from mismatched symmetries would not compromise the integrity of the virion.

Two lines of evidence seem to have some relevance to the question of the symmetry of the connector. First, several different chemical treatments of virions are known to cause separation of heads and tails (4, 8, 9–12), and when this happens, the connector usually remains attached to the tail. This is at least compatible with the idea that the connector is bonded more firmly to the tail than to the head membrane, as would be expected if it has 6-fold symmetry. Second, Paulson and Laemmli (13) have argued that the morphogenetic core of T4 heads, a transient structure that is present in the earliest stages of head assembly and is then lost, probably has 6-fold rotational symmetry. The 6-fold axis of the core and the 5-fold axis of the head membrane both coincide with the axis of the connector; thus, at early stages of head assembly, the connector is probably sandwiched between a structure with 5-fold symmetry and one with 6-fold symmetry.

Rotation of connectors

What are the implications of a symmetry mismatch between the connector and the remainder of the head? Consider two coaxial rings, A and B, of identical protein subunits turning relative to each other. As an individual subunit of ring A moves around ring B, the energy with which it interacts with ring B will change. This interaction can be expressed as a potential energy that varies as a function of the angle θ by which ring A has been turned relative to ring B. If ring B has six subunits, then the function will repeat with a period of $2\pi/6$. Curve a of Fig. 1 is an arbitrarily chosen function of this sort. The total interaction between rings A and B will be the sum of such functions for all subunits of ring A. If ring A has the same number of subunits as ring B (six), then the functions for the individual subunits will add in phase, and the resulting curve (Fig. 1, curve b) will have the same form as the curve for an individual subunit

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Table 1. Available measurements of numbers of copies of proteins in connectors

Phage	Protein	Copies/ virion	Ref.
λ	gpB*	10–14	14, 15†
	gpB	2–4	
T4	N2	5–6	8
	N6	10–11	
φ29	P10	5.6–5.7	16
T7	P13	38	4
	P14	19.7	
	P15	11.5	
	P16	3.9	
P22	gp1	10 ± 2	17

Proteins that are thought to add to the connector after DNA packaging are not included.
† Also, unpublished data.

and 6 times the amplitude. If, however, ring A has only five subunits, then the total interaction between the rings will be the sum of interactions for the individual subunits added *out of phase* (Fig. 1, curve c). Such a function will, in general, have a small amplitude and will repeat with a period of $2\pi/30$.

Comparison of curves b and c of Fig. 1 suggests that noncovalent bonding between rings of equal numbers of subunits can

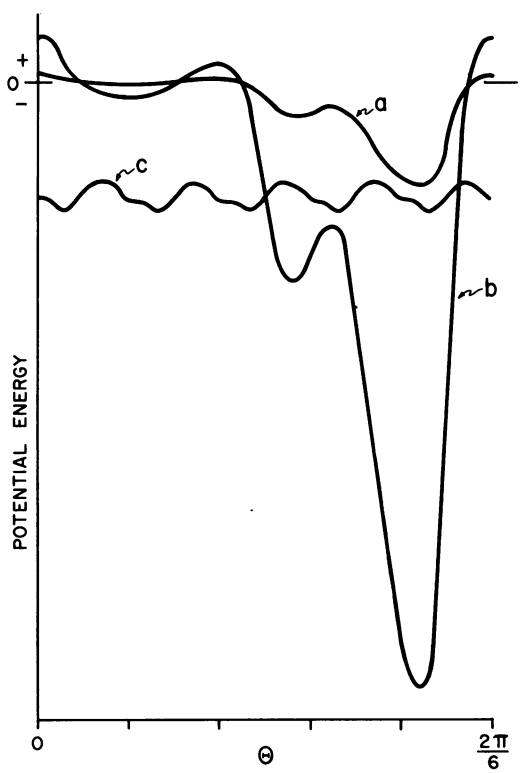

FIG. 1. Relationship between energy of interaction and angle of rotation for the protein rings described in the text. Zero potential energy represents no net interaction, and negative values of potential energy represent net attractive interaction. Curve a: Hypothetical curve for interaction between one subunit of ring A and ring B. Curve b: Curve for the total interaction between rings A and B, calculated from curve a, for the case of equal numbers of subunits. Curve c: Curve for the total interaction between rings A and B, calculated from curve a, for the case of unequal numbers of subunits.

be relatively strong and that the rings will have a strongly preferred orientation relative to each other. On the other hand, for the case of a ring of five subunits next to a ring of six subunits, the bonding will be relatively weak and, most interestingly, there will be no strong energy barriers to rotation of the rings relative to each other. Thus, one would predict that phage connectors, if they do in fact have 6-fold symmetry, are relatively free to rotate with respect to the head membrane.

DNA packaging

For the phages under consideration, it is clear that DNA is packaged into a preformed protein shell. Once packaged, the DNA is in a highly condensed state, with a packing density approximately equal to that of crystalline DNA. It is highly ordered and is in the crystallographic B form (18, 19). The packaged DNA is apparently in a metastable state, because it can leave the head spontaneously.

The mechanism by which DNA packaging is accomplished has been a subject of active investigation for several years, but none of the mechanisms proposed thus far seems entirely satisfactory (see discussion below). Most existing models see the connector as a passive hole through which the DNA is either pulled or pushed. I wish to propose that the connector is not passive but rather is a machine that actively moves the DNA into the head. This type of model has not been considered seriously in the past for the following reason. If the connector interacts directly with the DNA to move it into the head, then it seems likely that the geometrical relationships between the proteins of the connector and the two strands of the DNA molecule would need to remain the same as the connector passes along the DNA. Because the DNA is helical, the connector and the DNA must turn relative to each other. If the connector is fixed to the head, then the DNA and the head must turn relative to each other as the DNA screws into the head along its helical axis. Assuming that, once a substantial amount of DNA has entered the head, it is not free to rotate within the head, then this model would require that the two strands of the DNA become almost completely unwound during the process of packaging. This contradicts the data cited above on the state of DNA within the head.

However, this type of model becomes much more attractive if the connector is allowed to rotate relative to the head membrane. In this case, the geometrical relationship between the connector and the DNA can be maintained while the DNA passes into the head, but the DNA need not turn relative to the head membrane. Fig. 2 illustrates the proposed process. The connector rotates relative to both the head membrane and the DNA molecule and progresses along the DNA molecule much like a nut moving along a threaded rod. The DNA molecule and the head membrane turn relative to each other only enough to accommodate the superhelical twists being introduced into the DNA as it is wound around the inside of the head—perhaps one rotation for each 500 base pairs packaged. Initial orientation of the DNA molecule inside the head might be accomplished by binding of the DNA by the proteins of the head membrane. There is evidence from both λ and T7 for binding between head proteins and DNA (4, 20).

How is the packaging driven? I assume that the energy for packaging the DNA is provided by hydrolysis of ATP. With the possible exception of T7 (21), *in vitro* packaging of phage DNA is strongly dependent on added ATP (22–26). The mechanism by which DNA movement is accomplished is less clear. Two types of mechanism seem possible.

According to mechanism I, the active process would be the movement of the DNA relative to the connector. Hydrolysis of ATP would be coupled to movement of the DNA relative to

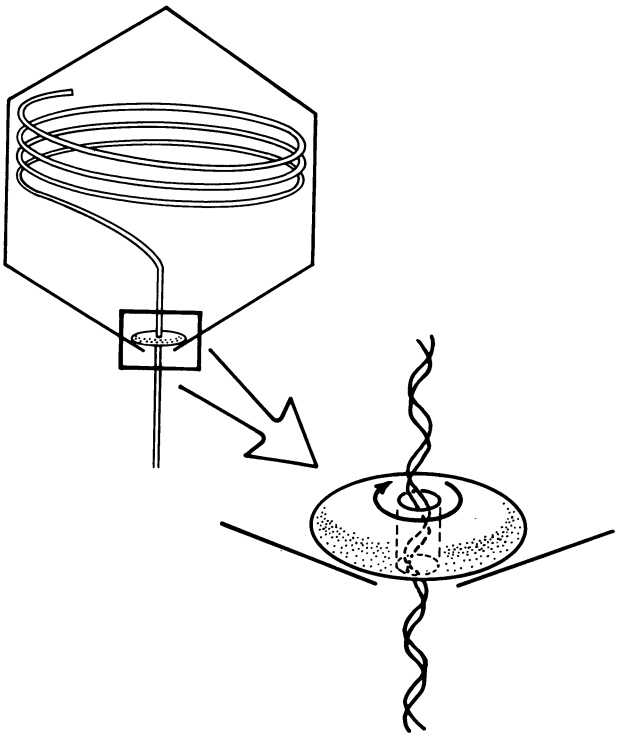

FIG. 2. Proposed model for DNA packaging. The connector rotates counterclockwise as viewed from outside the head.

the connector by one base pair along and 36° around its axis. Rotation of the connector relative to the head membrane would follow passively in an amount sufficient to relieve any twisting strain in the DNA. The molecular details of how such a movement of the connector along the DNA might work are obscure, but they might be expected to share features in common with the movement of polymerases, nucleases, or ribosomes along nucleic acid molecules.

Mechanism II would propose that the active, ATP-coupled process is the rotation of the connector. The DNA would then be screwed through the connector, just as a bolt is moved through a nut when the nut is turned. For mechanism II to work, the connector must in some sense be threaded. One way that this might be accomplished would be if the 6-fold symmetry of the hole in the connector that the DNA passes through is broken by proteins or parts of proteins that interact with a particular part of the DNA helix. Two of the phage connectors listed in Table 1 contain proteins that are present in two to four copies and that might fulfill this role. The uncleaved molecules of λ protein gpB are especially interesting in this regard. An alternative possibility is that the threads could be made of proteins that associate transiently with the connector during DNA packaging. λ proteins gpA and possibly gpNul have this property (27; R. Weisberg and N. Sternberg, personal communication).

The mechanism by which one ring of protein subunits might be driven relative to another is not known. However, if the two rings have unequal numbers of subunits as proposed, then an interesting mechanism suggests itself. As noted above, a ring of six subunits turning against a ring of five has 30 equivalent positions. Fig. 3 shows such a system progressing through 4 of the 30 equivalent positions. It can be imagined that each time a subunit reaches a certain position relative to the subunits of the other ring, indicated in the figure by an asterisk, it is capable of coupling ATP hydrolysis to movement of the ring in a unique direction. If each step illustrated in Fig. 3 required hydrolysis of 1 ATP, then one full rotation would consume 30 ATPs. (This assumes that there is only one functional site per subunit.) In the case of DNA packaging, this would amount to three ATP hydrolyses per base pair packaged. Fig. 3 could also have been drawn so that the rings move through 2 or 3 of the 30 equivalent positions at each ATP hydrolysis; for these cases, DNA packaging would use 1.5 or 1.0 ATPs per base pair packaged. In the evolution of such a mechanism, the energy required to go through one step of movement would presumably be tailored to the energy available from an ATP hydrolysis, possibly by changes in the protein structure which would affect the size of the energy barrier between equivalent positions.

For either mechanism I or II, the energy requirement for packaging is expected to be on the order of one ATP per base pair. For phage λ, this amounts to about 5×10^4 ATPs per virion. Although this may seem extravagant at first sight, it is only a small fraction of the total energy required to make a virion. Biosynthesis of the macromolecules found in a λ viron requires the equivalent of about 6×10^5 ATPs. In addition, virion production requires synthesis of considerable amounts of RNA and protein that are not incorporated into the virion. Thus, an energy expenditure of one ATP per base pair to package the DNA would amount to on the order of 5% of the total energy required for virion production.

Another way to ask if the model is reasonable is to ask whether it allows DNA to be packaged at a reasonable rate. There are no good estimates of how long DNA packaging actually takes. However, suppose that it takes 1 min to package a λ DNA molecule. This is a rate of about 800 base pairs per second, or a step time of 1.25 msec. If mechanism II were correct, then the individual steps would be carried out sequentially by more than one protein and the step time for an individual protein would be increased to 2.5 msec. For comparison, the step time for *Escherichia coli* RNA polymerase has been estimated to be about 12 msec for an ATP substrate at 37° (28), a single DNA replication fork in *E. coli* moves at a rate of about one base pair per 1.6 msec (29), and conjugative transfer of DNA in *E. coli* proceeds at a rate of about one base per 1.5 msec (30).

Injection

This model for DNA packaging suggests a solution to an important problem of DNA injection—namely, how injection is driven. If DNA packaging is driven against a free energy gradient by a process involving ATP hydrolysis, then some of the energy released in this process will be stored in the structure of the virion in the form of tightly constrained DNA, and this energy should be available for moving the DNA out of the head and into the next host cell.

In order to accomodate the model to all of the requirements of injection, however, it may be necessary to make additional assumptions about the mechanism of injection. According to the arguments presented above, the tail should be fixed firmly to the connector, and the tail presumably also attaches firmly to the cell surface. Thus, if the DNA is to rotate relative to the connector as it passes out of the head and into the cell, it must also rotate relative to the cell. Although I cannot rule out this possibility, it seems more attractive to postulate that injection is not simply the reverse of packaging. If the connector were to undergo a conformational change at some time subsequent to packaging, such that the DNA could pass freely through the connector without rotation, then the DNA could enter the cell without rotation. In the case of λ, there is evidence that the DNA is released to move partway down the tail when the tail attaches to the head (31–34), and this is consistent with the possibility of a conformational change in the connector. For phages with contractile tails, such as T4, it is clear that there are

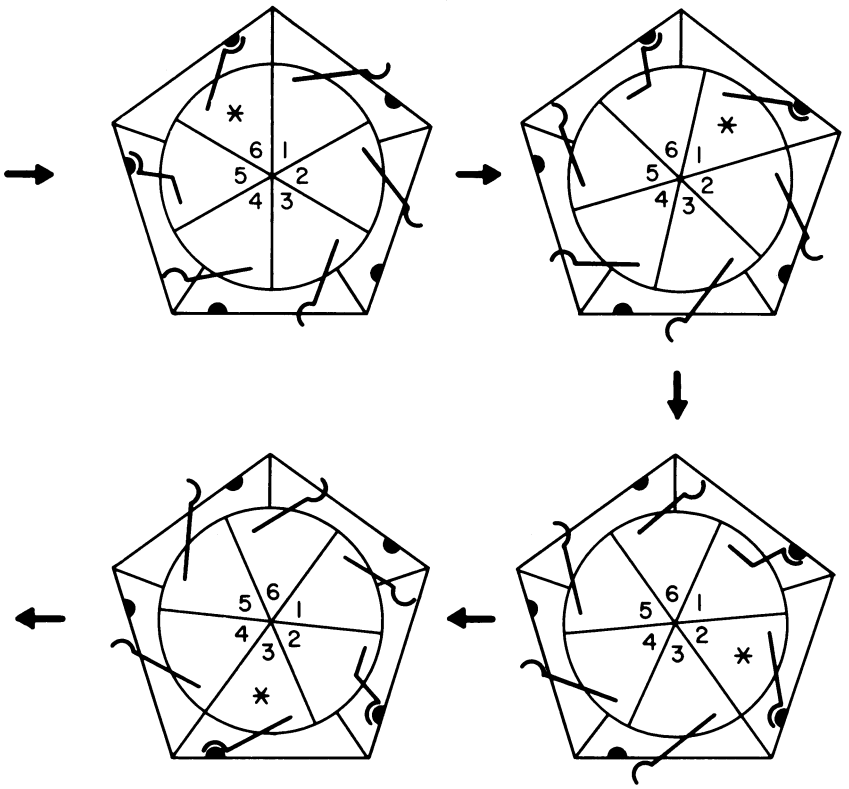

FIG. 3. Possible mechanism for driving rotation of the connector. The circular structure with 6-fold symmetry represents the connector and the surrounding material with 5-fold symmetry represents the head membrane.

extensive conformational changes in the virion when it interacts with the cell surface.

Other DNA packaging models

The assembly pathways of the protein capsids of the various phages under consideration have several common features. Among these are the presence of a core or scaffold early in assembly, subsequent loss of the core, and expansion of the capsid to its final size. There are also several details that differ among the different phages. These include whether or not covalent processing of proteins occurs, whether loss of the core is accompanied by degradation of the core protein, whether proteins are added after capsid expansion, and others. Most of the models previously proposed to explain DNA packaging have suggested that DNA entry is coupled to one of the steps of capsid assembly, including degradation of the core (35), exit of the scaffolding (core) protein (17), or expansion of the capsid (36, 37).

The phages being considered are sufficiently similar in the basic features of their structure and assembly that it would be surprising if they did not use a common mechanism for packaging DNA. If this is assumed to be the case, then most of the previous models seem untenable. For example, in the case of λ, loss of the assembly core clearly occurs prior to DNA packaging (20, 36, 38–41), arguing against models in which cleavage or loss of the core is coupled to DNA entry. In P22, protein cleavage apparently does not occur (17), and this makes models requiring cleavage unattractive. A recent report by Hsiao and Black (42) argues convincingly in the case of T4 that DNA packaging cannot be obligatorily coupled to cleavage or loss of the core, cleavage of the major head protein, expansion of the capsid, or concomitant DNA synthesis. Their data implicate protein P20 in DNA packaging. Although P20 is not a component of the T4 connector (8), it is apparently located at the same corner of the head as the connector (43). It may form the part of the head membrane against which the connector rests.

The DNA packaging model proposed here would seem to have several advantages over previous models. First, it is not in obvious conflict with any of the experimental data now available, and it appears to be applicable to all of the well-studied double-stranded DNA phages. Second, it provides a plausible means for coupling ATP hydrolysis to DNA movement and, consequently, a mechanism for storing energy in the structure of the virion to drive DNA injection. Third, it provides a rational explanation for the symmetry mismatch between phage heads and tails. Furthermore, the model seems amenable to extensions that can explain other features of DNA packaging. For example, in the case of phages that cut the DNA after a "headful" has been packaged (44), the increased difficulty of forcing DNA into the head as a headful is approached might be reflected in a strain-induced conformational change in the connector which could activate an associated nuclease. A more complete assessment of the model and its variations will only be possible when the detailed structures of connectors are better understood.

The principal new idea which makes the proposed DNA packaging model possible is that protein rings with unequal numbers of subunits should be relatively free to rotate against each other. There is currently only one clear example of rotary motion in a biological system. Bacterial flagella are driven in rotary motion by a motor embedded in the cell wall and membrane (45, 46). The way in which this motor works is not yet clear, but its morphology is compatible with the possibility of unequal protein rings driven by the sort of mechanism illustrated in Fig. 3. An interesting feature of this mechanism which might recommend it as a general mechanism arises from the fact that the number of equivalent positions that can be assumed by two rings of n and $n + 1$ subunits as they rotate is $n(n + 1)$. This means that the number of ATP hydrolyses or other energy-yielding events that can be coupled to a rotation varies roughly as the square of the ring size. Consequently, small differences in ring size would result in motors with large differences in gearing ratio.

I thank the many colleagues who provided valuable discussions and criticisms of the model and the manuscript, especially Howard Berg, David Botstein, Allan Campbell, Sherwood Casjens, Ron Davis, Bill Earnshaw, Fred Eiserling, Costa Georgopoulos, Thomas Hohn, Barbara Hohn, Dale Kaiser, Edward Kellenberger, Ned Minkley, Helios Murialdo, Jim Paulson, Dick Russell, and Nat Sternberg. This work was supported by research Grant AI12227 and Research Career Development Award AI00023 from the National Institutes of Health.

1. Moody, M. F. (1965) *Virology* **26**, 567–576.
2. Caspar, D. L. D. & Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
3. Crowther, R. A. & Klug, A. (1975) *Ann. Rev. Biochem.* **44**, 161–182.
4. Serwer, P. (1976) *J. Mol. Biol.* **107**, 271–291.
5. Anderson, D. L., Hickman, D. D. & Reilly, B. E. (1966) *J. Bacteriol.* **91**, 2081–2089.
6. Casjens, S. R. & King, J. (1975) *Annu. Rev. Biochem.* **44**, 555–611.
7. Mazza, A. & Felluga, B. (1973) *J. Ultrastr. Res.* **45**, 259–278.
8. Coombs, D. H. & Eiserling, F. A. (1977) *J. Mol. Biol.* **116**, 375–405.
9. Cummings, D. J., Chapman, V. A. & DeLong, S. S. (1968) *J. Virol.* **2**, 610–620.
10. Yanagida, M. & Ahmed-Zadeh, C. (1970) *J. Mol. Biol.* **51**, 411–421.
11. Bassel, A., Shaw, M. & Campbell, L. L. (1971) *J. Virol.* **7**, 663–672.
12. Harrison, D. P., Brown, D. J. & Bode, V. C. (1973) *J. Mol. Biol.* **79**, 437–449.
13. Paulson, J. R. & Laemmli, U. K. (1977) *J. Mol. Biol.* **111**, 459–485.
14. Hohn, T., Flick, H. & Hohn, B. (1975) *J. Mol. Biol.* **98**, 107–120.
15. Casjens, S. R. & Hendrix, R. W. (1974) *J. Mol. Biol.* **88**, 535–545.
16. Hagen, E. W., Reilly, B. E., Tosi, M. E. & Anderson, D. W. (1976) *J. Virol.* **19**, 501–517.
17. Casjens, S. R. & King, J. (1974) *J. Supramol. Str.* **2**, 202–224.
18. Earnshaw, W. C., Casjens, S. R. & Harrison, S. C. (1976) *J. Mol. Biol.* **104**, 387–410.
19. Earnshaw, W. C. & Harrison, S. C. (1977) *Nature (London)* **268**, 598–602.
20. Brody, T. (1973) *Virology* **54**, 441–451.
21. Kerr, C. & Sadowski, P. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3545–3549.
22. Becker, A., Murialdo, H. & Gold, M. (1977) *Virology* **78**, 277–290.
23. Hohn, B. & Hohn, T. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2372–2376.
24. Pruss, G., Goldstein, R. & Calendar, R. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2367–2371.
25. Kaiser, D., Syvanen, M. & Masuda, T. (1975) *J. Mol. Biol.* **91**, 175–186.
26. Poteete, A. (1976) Dissertation (Massachusetts Institute of Technology, Cambridge, MA).
27. Becker, A., Marko, M., & Gold, M. (1977) *Virology* **78**, 291–305.
28. Manor, H., Goodman, D. & Stent, G. (1969) *J. Mol. Biol.* **39**, 1–29.
29. Helmstetter, C. E. & Cooper, S. (1968) *J. Mol. Biol.* **31**, 507–518.
30. Bachmann, B. J., Low, K. B. & Taylor, A. L. (1976) *Bacteriol. Rev.* **40**, 116–167.
31. Padmanabhan, R., Wu, R. & Bode, V. C. (1972) *J. Mol. Biol.* **69**, 201–207.
32. Saigo, K. & Uchida, H. (1974) *Virology* **61**, 524–536.
33. Thomas, J. (1974) *J. Mol. Biol.* **87**, 1–9.
34. Chattoraj, D. K. & Inman, R. B. (1974) *J. Mol. Biol.* **87**, 11–22.
35. Laemmli, U., Paulson, J. & Hitchins, V. (1974) *J. Supramol. Str.* **2**, 276–301.
36. Hohn, B., Wurtz, M., Klein, B., Lustig, A. & Hohn, T. (1974) *J. Supramol. Str.* **2**, 302–317.
37. Serwer, P. (1975) *J. Mol. Biol.* **92**, 433–448.
38. Hendrix, R. W. & Casjens, S. R. (1975) *J. Mol. Biol.* **91**, 187–199.
39. Kaiser, D. & Masuda, T. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 260–264.
40. Zachary, A., Simon, L. D. & Litwin, S. (1976) *Virology* **72**, 429–442.
41. Lickfeld, K., Menge, B., Hohn, B. & Hohn, T. (1976) *J. Mol. Biol.* **103**, 299–318.
42. Hsiao, C. L. & Black, L. W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3652–3656.
43. Müller-Salamin L., Onorato, L. & Showe, M. K. (1977) *J. Virol.* **24**, 121–134.
44. Streisinger, G., Emrich, J. & Stahl, M. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 292–295.
45. Berg, H. C. & Anderson, R. A. (1973) *Nature (London)* **245**, 380–382.
46. Silverman, M. R. & Simon, M. I. (1974) *Nature (London)* **249**, 73–74.