Journal of Supramolecular Structure 2:225–238 (1974)

P22 MORPHOGENESIS II: MECHANISM OF DNA ENCAPSULATION

Bik-Kwoon Tye and David Botstein

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts

Phage P22 is known to have a linear duplex chromosome which is circularly permuted and terminally repeated. The propagation of these features of the mature phage DNA is accounted for by the fact that phage DNA lengths (headfuls) are cut from an intracellular intermediate form of phage DNA several phage genomes in length (concatemer) as first suggested by Streisinger. Studies with mutant phages show that cutting of concatemer DNA is intimately connected to the morphogenesis of the phage head.

We have also found, by constructing a partial denaturation map of mature P22 DNA, that circular permutation in P22 DNA is restricted: all of the ends of the mature DNA fall within 20% of each other on the physical map. The limited distribution of ends can be explained by Streisinger's "headful" packaging model with the additional specifications that: a. the intracellular precursor DNA is no longer than ten times the length of mature phage DNA; b. encapsulation of DNA starts at a unique site; c. encapsulation proceeds sequentially therefrom.

This model is supported by the distribution of molecular ends in denaturation maps of two deletion phage DNAs. We found, as expected from our model, that the extent of permutation is a direct function of the length of terminal repetition.

INTRODUCTION

The preceding paper summarizes the current understanding of the morphogenesis of the bacteriophage P22 capsid from the point of view of the proteins which either form part of the final virion or which are active in the correct assembly of the phage head. In this paper, we summarize the known facts about the morphogenesis of the P22 phage head from the point of view of the DNA which it contains.

Just as there are unique protein structures (proheads) which are precursors to the mature virion (5, 6, 15) so are there unique forms of phage DNA which are precursors of the mature phage DNA. Furthermore, just as head assembly is not simply a self-assembly

 © 1974 Alan R. Liss, Inc., 150 Fifth Avenue, New York, N.Y. 10011

of its component proteins, so the DNA which enters the head is quite unlike the DNA which is found in the mature phage.

The mature phage DNA has several unique properties. It is a single linear double-stranded molecule with molecular weight of about 27 million (16). It encodes all of the phage genes, of course; but in addition, each DNA molecule is repetitious at its ends, to the extent of about 2%. Furthermore, the molecules in each phage head are not identical; each is a circular permutation of the other (16, 20). These relations are summarized in Fig. 1, which shows schematically the DNA life-cycle of phage P22.

Streisinger et al. (18) proposed a model that accounts for the generation of these circularly permuted and terminally repeated DNA molecules. They envisioned that the intracellular precursor DNA molecules are repeating polymers (concatemers) of the phage genome. Streisinger suggested that the length of the DNA packaged inside of a phage head is determined by the amount of DNA that can fit into the head, i.e., DNA is packaged by the headful. If the genome length is smaller than the headful, a terminally repreated molecule will be produced when a headful is cut from the concatemer. A headful in the case of P22 is equivalent to the complete wild-type phage genome plus 2%. Streisinger et al. (18) showed by genetic means that phage with deletions in their genome have a longer terminal repetition. Using more direct physical methods, we have recently confirmed this result in the case of P22, showing quantitatively that the genome size directly determines the size of the terminal repetition (20).

RESULTS

The actual existence of concatemers during lytic growth of P22 was shown directly some years ago by Botstein (1). Radioactively labeled replicating phage DNA was analyzed in neutral and alkaline sucrose gradients; essentially all the replicating DNA was found in a complex form called "intermediate I." Most of the DNA in intermediate I is concatemeric, and most of it can be shown to be a precursor of mature phage DNA.

More recently, a comprehensive study of the effect of mutations in the known phage genes on the maturation of concatemeric DNA was undertaken. The results of this work are shown in Fig. 2 and Table I. In Fig. 2, the ability of mutants in the various genes to remove DNA from intermediate I (the concatemeric form; Fig. 2a) and to cut this DNA to the correct mature size (Fig. 2b) is shown. The results (summarized in Table I) indicate that only those mutants which can make a full head mature the DNA; those that make an unstable full head also mature the DNA, which is degraded when the head falls apart. The remaining mutants (those that make proheads only or no head structure) fail to mature the DNA at all; it stays concatemeric. In this latter class are the mutants in gene 5 (the major capsid subunit) and gene 8 (the scaffolding protein). This means that an intact prohead structure is required in the cutting of concatemer DNA to the mature size, and strongly suggests that the DNA cutting, DNA encapsulation, and the exit of the scaffolding protein P8 are all connected, at least temporally.

In this way it has become established that the generation of mature phage particles from the concatemer, as originally envisioned by Streisinger, is responsible for the

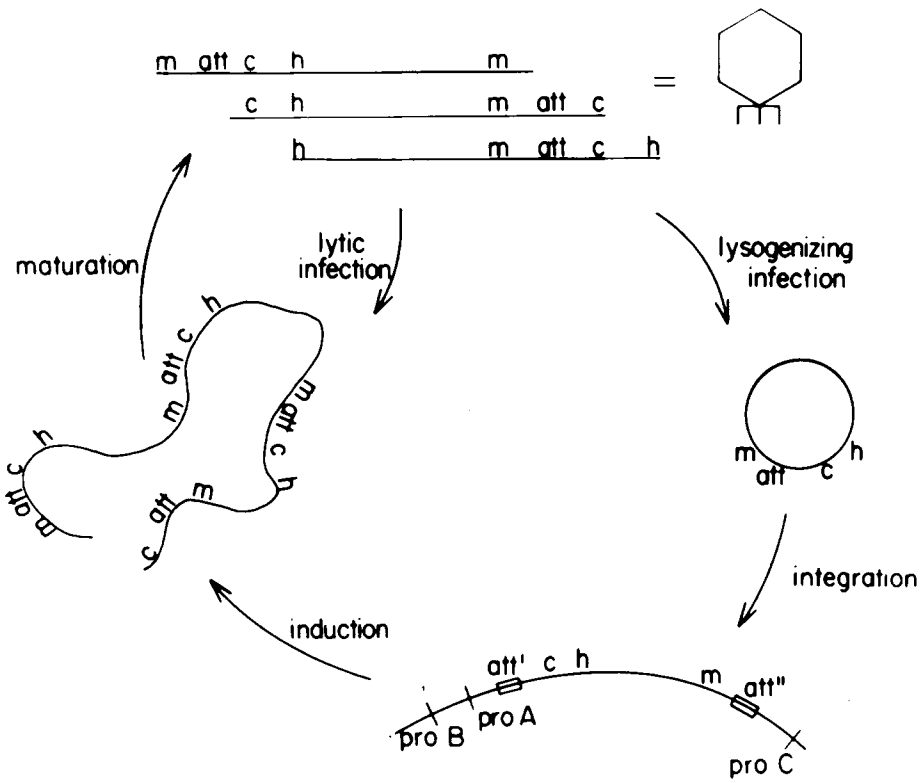

Fig. 1. Rearrangements in the topological form of P22 phage DNA during the life cycle. m, c, and h are genetic markers; att is the site of attachment of the phage to the host chromosome in lysogeny. proA and proC are genetic markers of the host. The DNA in phage heads is circularly permuted and terminally repetitious, as shown. During lytic infections a concatemer is formed; during lysogenizing infections circular forms are observed.

propagation of terminal repetition and circular permutation in mature phage P22 DNA. It should be noted, however, that phages with other mature DNA structures also seem to replicate through concatemeric intermediates: the list of such phages includes at present phage λ, which has single-stranded cohesive ends and is a nonpermuted linear structure (12), and phage T7, which has a terminal repetition but no circular permutation (19).

The generation of circular permutation in the DNA of the progeny of a single phage particle is nicely explained by the Streisinger model as well. If headfuls are cut from the concatemer at different genetic points, circular permutation (as well as terminal repetition) results. However, the Streisinger model does not specify whether the headful encapsulation mechanism cuts out headfuls randomly from the concatemeric precursor or whether headfuls are cut sequentially from an end once packaging begins on a concatemer. If the concatemeric DNA were infinitely long or if the sequential cutting could initiate from any random point, one would expect the resulting population of DNA molecules to

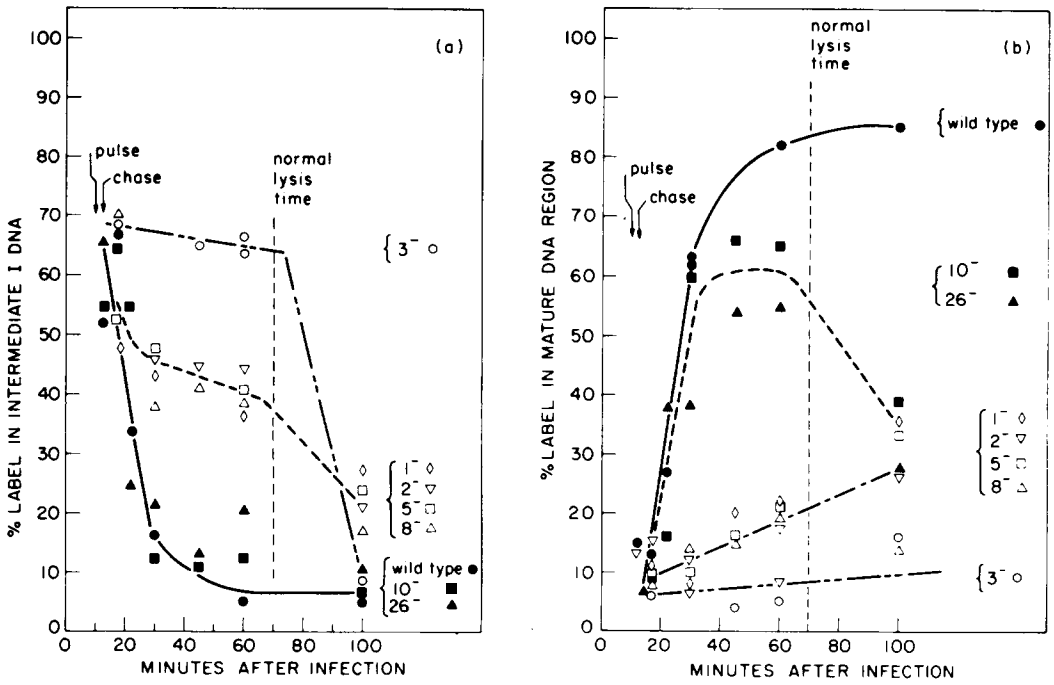

Fig. 2. Analysis of DNA maturation in cells infected with phage mutants. Pulse-chase experiments were performed in which ^{3}H-thymidine was added at the indicated times and followed two minutes later by excess unlabeled thymidine. The radioactivity in DNA was followed in neutral sucrose velocity sedimentation gradients. Each point on these summary figures is the result of such a gradient. In (a) the proportion of the radioactivity in the replication complex (intermediate I) is shown as a function of time; in (b) the proportion of radioactivity cosedimenting with added ^{32}P-labeled mature phage DNA is shown. Panel (a) thus shows appearance of DNA in the mature phage form. All of these experiments were done in conditions which prevent lysis of the infected cells: thus, breakdown of the DNA seen in these experiments represents in vivo degradation occurring after the normal lysis time. Details and further discussion of these experiments are published elsewhere (5). (◇) gene 1^{-}; (▽) gene 2^{-}; (○) gene 3^{-}; (□) gene 5^{-}; (△) gene 8^{-}; (■) gene 10^{-}; (▲) gene 26^{-}; (●) wild-type P22.

TABLE I. Gene Products and Mutant Phenotypes

	Phenotype			Protein Products	
Gene	DNA Cutting	Heads (EM)	In Mature Phage	(MW)	Number per Phage
1	No	Prohead	Yes	94,000	20
2	No	Prohead	No	63,000	–
3	No	Prohead		9,500	–
5	No	None	Yes	55,000	400
8	No	Few aberrant	No	42,000	–
10	Yes	Full (unstable)		56,000	–
26	Yes	Full (unstable)	Yes	23,000	30
9	Yes	Full (tailless)	Yes	76,000	20
16	Yes	Full	Yes	67,000	10
20	Yes	Full	Yes	50,000	20 – 30

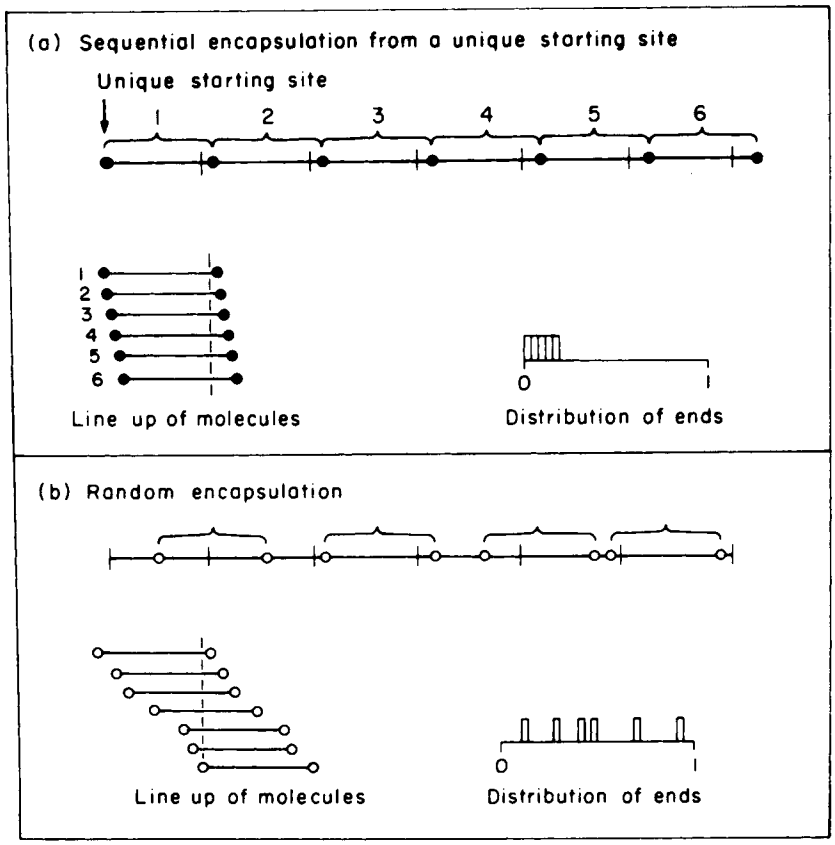

Fig. 3. (a) Diagram showing that if the intracellular multiple-length precursor is only long enough for a small number of headfuls to be made, sequential encapsulation from a unique starting site results in a restricted distribution of ends. (b) Random encapsulation results in a random distribution of ends.

have randomly permuted ends. If, however, the cutting were always initiated at a specific site and the precursor DNA were of finite length, then one could distinguish the random cutting mechanism from the sequential cutting mechanism because each sequential headful would have ends differing from the preceding headful by the length of the terminal repetition. Thus, if the concatemer were 10 headfuls long and the terminal repetition were 2%, then the total distance between the ends of the first and last headful would only span 20% of the genome. A comparison between random and unique site/sequential encapsulation is shown in Fig. 3.

We have shown that the sequential encapsulation hypothesis is correct for P22; what follows is a summary of recently published (21) evidence for this view. The experimental basis for the measurement of the degree of permutation is alignment of partial denaturation maps of P22 DNA molecules.

The partial denaturation technique was invented by Inman (13). It is based on the

observation that one can find partially denaturing conditions for DNA (i.e., high pH or high concentrations of formamide) in which only particular limited regions of the DNA are denatured. In suitably prepared samples, these regions can be visualized in the electron microscope. An example of a partially denatured P22 DNA molecule is shown in Fig. 4. Most of the molecule is double-stranded, but occasionally "bubbles" in which the denatured strands have separated can be seen. Below the electron micrograph is a linear representation of the size and spacing of the denatured regions. The denaturation pattern is characteristic of the DNA's genetic nature, and reflects substantial local inhomogeneities in the percentage of AT base pairs. The partial denaturation pattern can be used as a physical map of the genetic information in the DNA molecule.

By preparing partial denaturation maps of P22 DNA molecules, we were able to measure the extent of permutation by aligning the maps with respect to the partial denaturation pattern and looking at the distribution of molecular ends. Such an alignment of 63 P22 wild-type DNA molecules is shown in Fig. 5. A weight-average histogram of the aligned maps is shown in Fig. 6b; Fig. 6a shows the distribution of molecular ends. It is clear that the extent of circular permutation of the P22 DNA molecules is limited: none of the ends is displaced from any other by more than about 20%, of the map.

The Sequential Encapsulation Hypothesis

One way to explain the generation of molecules with permuted but clustered ends is to hypothesize that during DNA maturation, encapsulation of DNA by phage heads takes place at a unique site and proceeds sequentially, as shown in Fig. 3a. If the terminal repetition is small, and the concatemer is only a few headfuls long, then sequential cutting from a unique site will result in limited circular permutation. The extent of permutation will depend on the size of the terminal repetition and the length of the concatemer.

An alternate explanation for the permuted but clustered distribution of ends is to hypothesize that the encapsulation mechanism has a specificity for certain G–C rich sequences, which are clustered within a G–C rich region comprising about 20% of the P22 genome.

These explanations can be distinguished by the observation that the unique site/sequential encapsulation model predicts that the extent of permutation in the mature phage DNA should depend on the size of the terminal repetition. Alternative site-specific models should not share this property.

Tye et al. (20) have shown that a deletion in the phage genome is compensated for by a lengthening of the region of terminal repetition by an amount equal to the length of the deletion; and, similarly, an insertion in the phage genome is compensated for by a shortening of the region of terminal repetition. Thus, the materials for testing the unique-site/sequential encapsulation hypothesis are readily available in the form of deletion and insertion mutants of P22 which have differing genome lengths and, therefore, different amounts of terminal repetition.

Figure 7 illustrates the detailed quantitative expectations of the sequential encapsulation model. One can see that the displacement of the beginning of the nth headful

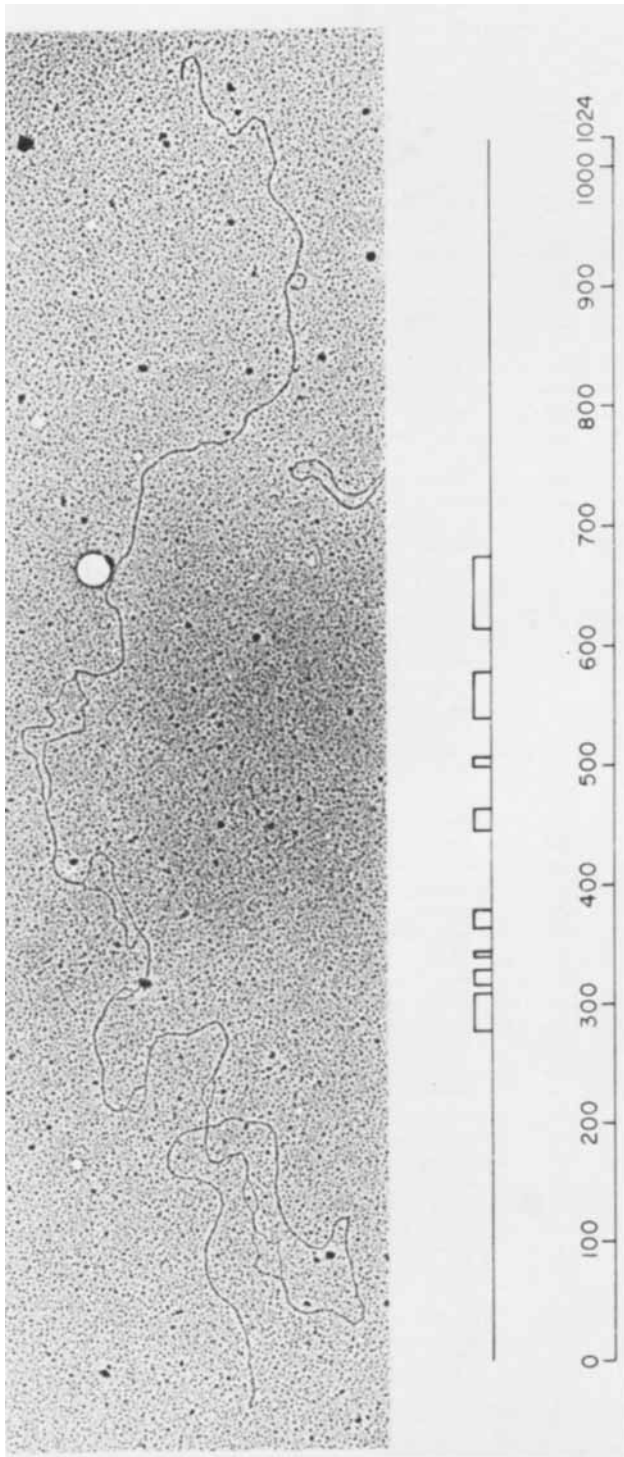

Fig. 4. (Top) Electron micrograph of a P22 DNA molecule partially denatured in 83% formamide. The details of sample preparations are published (21). (Bottom) Linear representation of the partial denaturation map.

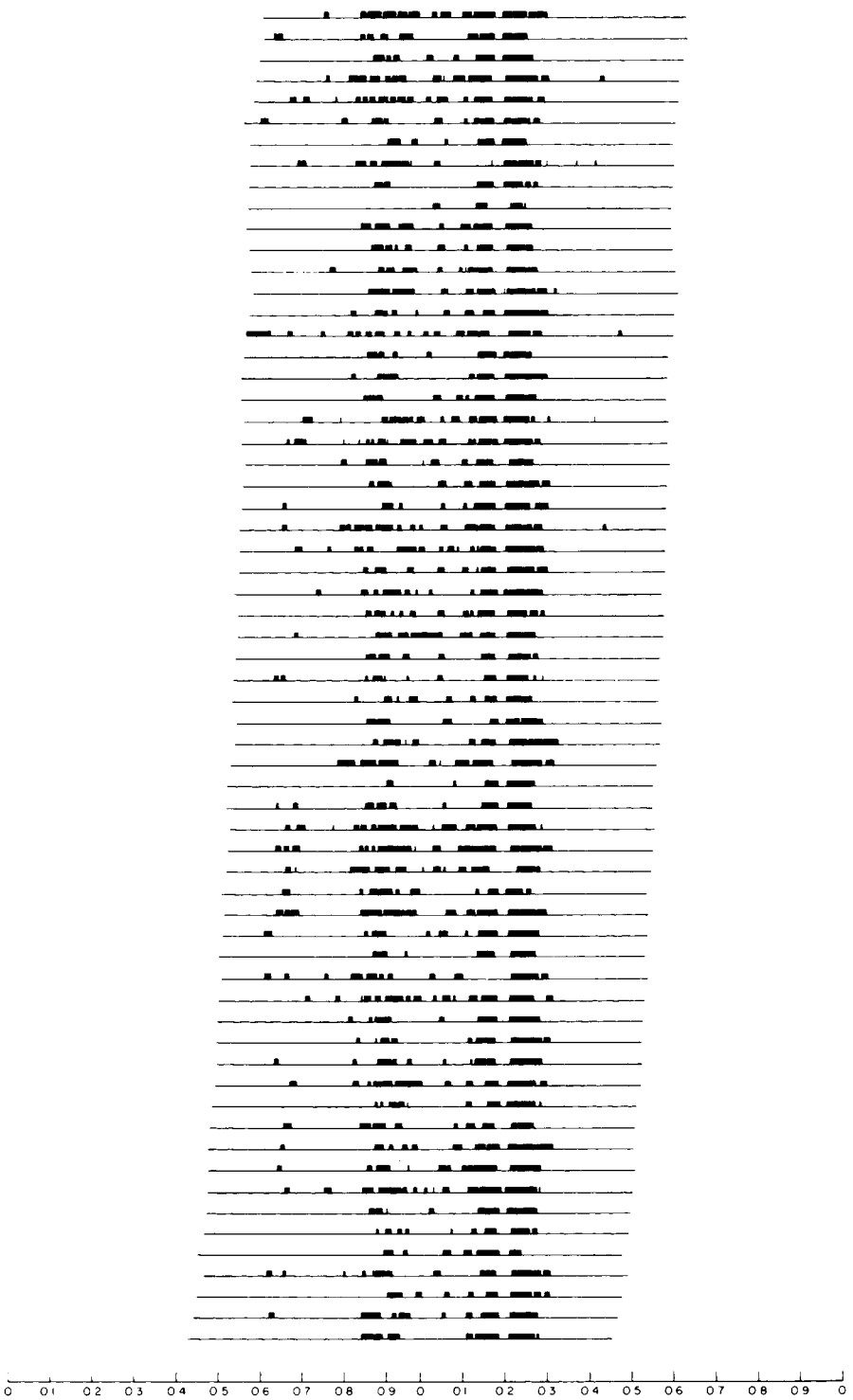

Fig. 5. Partial denaturation maps of P22 wild-type DNA denatured in 83% formamide as in Fig. 4. All molecules are normalized to 100% P22 mature DNA length and the length distribution of the 63 molecules shown is (100 ± 6.7%). The arrow indicates the reference (a characteristic gap between two denaturation sites) of alignment, which in this case was done by eye. The maps are represented in such a way that molecules with ends closest to one another are adjacent.

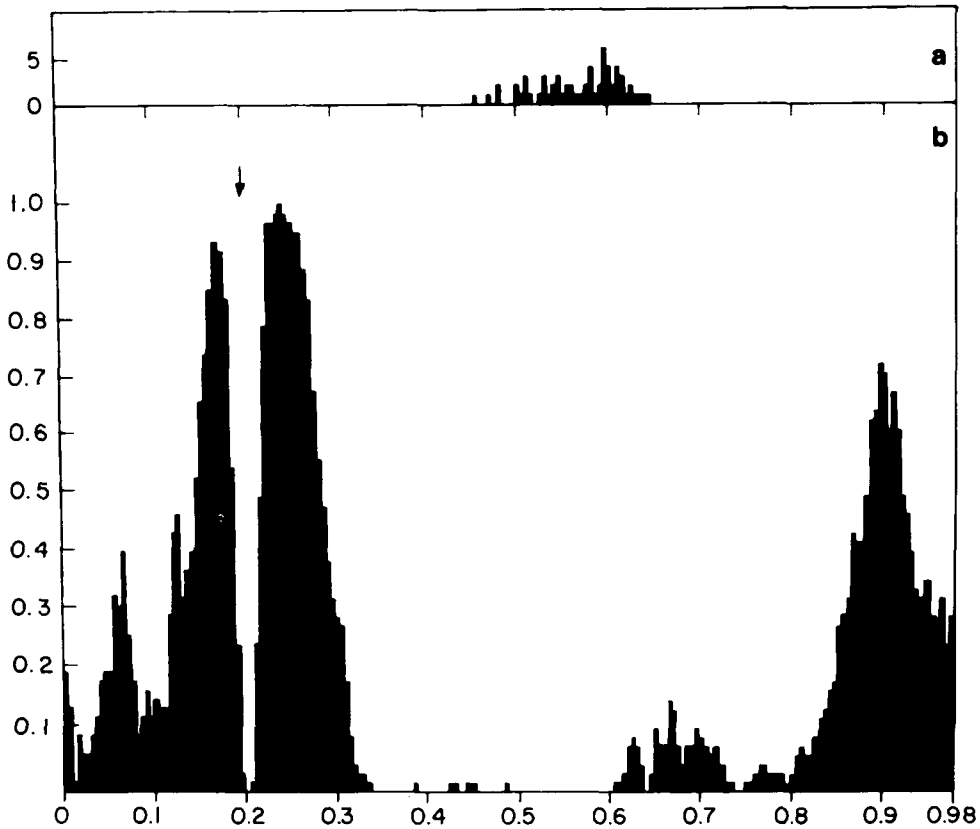

Fig. 6. (a) Histogram of the distribution of the ends of the molecules with respect to the denaturation map. The positions of ends are positions of the right ends in the alignment (see Fig. 2). The distribution of ends is nonrandom, clustering within 20% of the map length. (b) Histogram of the partial denaturation maps of P22 wild-type DNA (in Fig. 5) showing the positions and frequency of the denatured sites. The X-axis is plotted as a function of the genome length such that the terminal repetition is not included in the map range, making the total length 98% P22. This is done by taking out the amount of terminal repetition (2%) from each molecule, all from the same end (left end in Fig. 2). The arrow indicates the reference for the alignment.

from the initiation site is a simple linear function of the terminal repetition characteristic of a particular phage genome. Thus, the terminal repetition in a P22 phage with a 5% deletion is 7% [the sum of the deletion (D = 5%) plus the wild-type terminal repetition (T = 2%)]. Then the tenth headful will begin 63% displaced from the unique site (taken as position 0% on the map). The tenth headful from a concatamer of wild-type genomes (T = 2%) would, by a similar calculation, only be displaced 18%.

Another prediction of the model applies to phages with very large terminal repetitions. In this case, the successive headfuls will be far apart, so far, in fact, that the

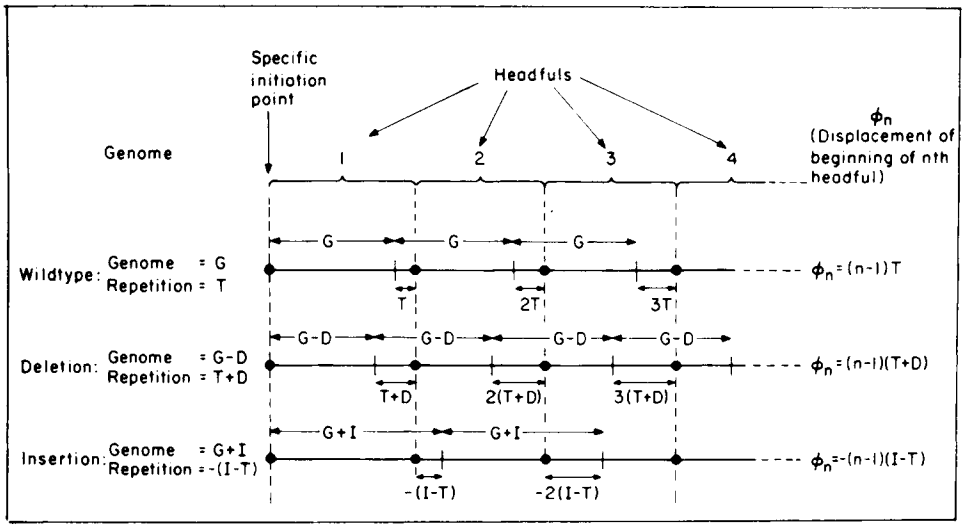

Fig. 7. Diagram illustrating the relationship between genome size and the size of the terminal repetition (TR); a prediction of the Streisinger model. Streisinger's model proposes that the length of the DNA inside a phage head is determined by the volume of the phage head. Thus, a deletion (D) in the phage genome (G) would be compensated for by a lengthening of the terminal repetition by an amount of D to (T + D). Similarly, an insertion (I) in the phage genome would be compensated for by a shortening of the TR by an amount (I) to (T – I). If I is greater than T, then only a fraction of the phage genome can be encapsulated by each phage head. The difference between the length of each headful (G + T) and the insertion phage genome (I + G) is –(I – T); this we term negative terminal repetition.

The diagram also illustrates the relation between the size of the TR and the extent of permutation in the sequential encapsulation of DNA starting from a unique site. The longer the TR, the larger is the displacement of each sequential headful from its previous headful resulting in a dependence of the extent of permutation on the size of the terminal repetition. For example, in the case of the wild-type phage, the fourth headful would be displaced from the first by an amount = 3T; i.e., the extent of permutation in 4 sequential headfuls is 3T. In the case of the deletion phage, the extent of permutation in 4 sequential headfuls is 3(T + D); and in the case of the insertion phage, the extent of permutation in 4 sequential headfuls is –3(I – T).

total errors in measurement might be small enough to allow distinction of each successive class.

The foregoing analysis is the basis for a simple test of the sequential encapsulation idea: namely, if one used P22 phages with different amounts of terminal repetition in their DNA, one should find predictably different distributions of molecular ends among mature phage DNA molecules. Different amounts of terminal repetition are easily produced by shortening the genome by deletions (thus making the difference between the headful and the genome larger). Figure 8 is a summary of the result of partial denaturation analysis of mature DNA from two deletion mutants of phage P22. One of these has a terminal repetition of about 7%; the other has a terminal repetition of about 16%. It should be emphasized that the extent of deletion and terminal repetition were determined independently of partial denaturation maps (20, 21). The main result is very

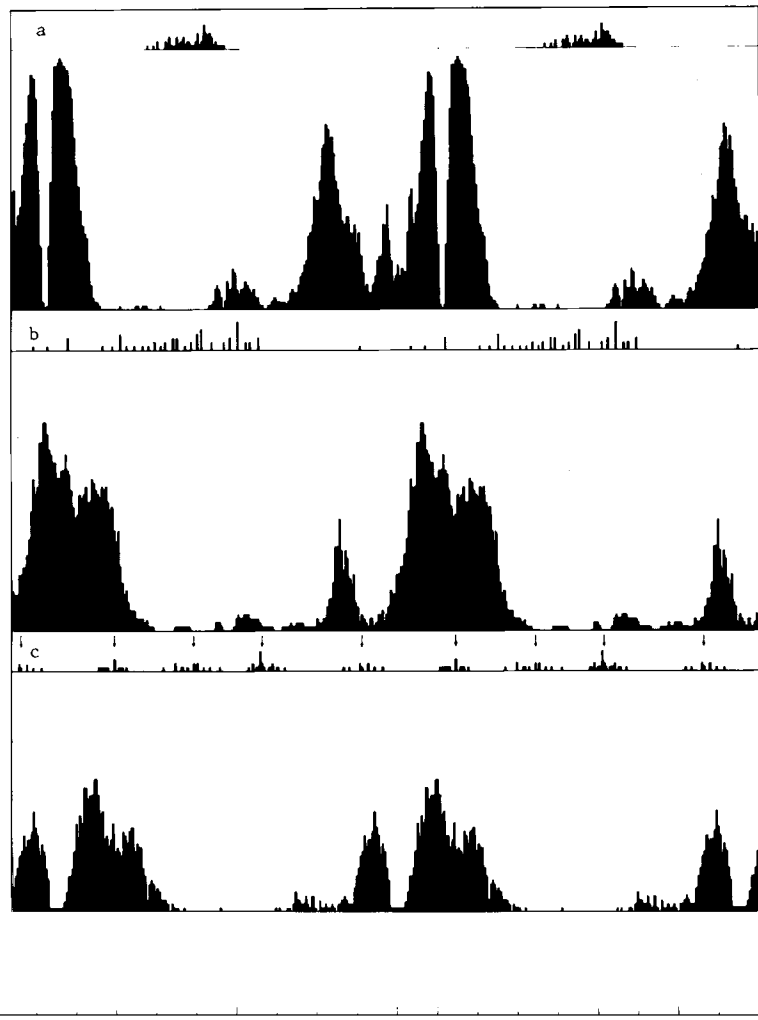

Fig. 8. Histograms (see Fig. 6) of the partial denaturation maps of (a) P22 wild-type DNA (2% terminal repetition), (b) P22bpl DNA (7% terminal repetition), (c) P22bp5 DNA (16% terminal repetition). The histograms are plotted as double map lengths for an easier comparison of the profiles of these histograms. The histograms are lined up with one another by the two major denaturation sites at the center. A comparison of these histograms shows: the relation between the extent of permutation and the size of the TR; that the most frequent headfuls are the first headfuls (this is strengthened by the fact that they all coincide at the same region of the partial denaturation map); and encapsulation seems to proceed sequentially and unidirectionally from right to left since the number of headfuls decreases unidirectionally from right to left in bpl and in bp5.

obvious: in the case of the 7% terminal repetition, the molecular ends are spread over a region encompassing about 70% of the genome (as opposed to only 20% for wild-type P22). In the case of the 16% terminal repetition, discrete classes of ends are found about 16% apart, just as anticipated in the analysis above.

The data in Fig. 8 also suggest that sequential encapsulation from the unique starting site is in one direction only. This suggestion is strongly made by the fact that the distribution of ends in the 7% terminal repetition case begins at a point in the map which coincides with the more restricted distribution of ends found in wild-type P22 DNA molecules. If encapsulation were occurring in two directions, then the wild-type and deletion distribution would be centered around the same point, rather than having one end in common.

DISCUSSION

Our picture of the events in DNA metabolism during the encapsulation of phage DNA is now quite detailed. We know that at some point in the lytic growth of the phage, concatemers are produced, probably as a direct product of DNA replication (4). These concatemers are reduced in size to the mature phage length if and only if the appropriate protein precursor structures (proheads) can encapsulate the DNA. If the proheads are defective or absent, then the concatemers are not cut.

If one combines these findings with those of the preceding paper, we see that the critical step in head morphogenesis and in DNA maturation must involve something like the following events: the prohead and the DNA interact; P8 (the scaffolding protein) is ejected and the DNA enters the nascent head; when the head is full, all the P8 is gone and the DNA is cut. Some time in this process P2 and P3 act (14, 15) and the actual size of the head structure increases some 30% in volume (Earnshaw, Casjens, and Harrison, personal communication).

Important questions remain about this process. First of all, we cannot decide for certain whether the unique starting site for sequential encapsulation involves recognition of a particular DNA sequence by the prohead or whether this site is a unique structural feature of the concatemer DNA itself. One attractive hypothesis for a unique structural feature in the DNA is that DNA replication proceeds by some asymmetric mechanism such as the rolling circle (11). The rolling circle form has a free end which might be the origin of DNA replication. Then the prohead simply needs to recognize free double-stranded ends – the first such available is the end of the rolling circle; subsequent ends are produced by each encapsulation event. This idea is attractive because each of the encapsulation events is the same mechanistically; all the specificity is in the formation of the DNA substrate. Preliminary data of Chan (7) suggest that the unique starting site might be located at or near the origin of DNA replication. However, many uncertainties remain to be resolved about the origin and structure of the concatemers before any solid conclusion can be drawn on this point.

Another question which remains unresolved is the way in which the DNA actually comes to be inside the prohead. Although we know something about the identity and fates of the proteins involved, we have no real clue about the mechanism. The result

presented above seems to indicate that the DNA enters the prohead from an end – however, we cannot rule out more complicated models involving initial binding of the middle of the DNA. Similarly, the results favor the idea that the prohead has formed before it attaches to the DNA which will wind up inside it – after all, the end of the DNA headful becomes available only after the previous headful has been completed.

It is worth mentioning that the foregoing ideas neatly account for the ability of phage P22 to act as vector for generalized transduction. All we need to assume is that occasionally double-strand scissions in the bacterial DNA occur. Then the sequential encapsulation mechanism will use this DNA, producing phage particles with pre-existing host DNA in them instead of phage DNA, as shown previously to be the case for P22 by Ebel-Tsipis et al. (8). If the occurrence of such double-strand breaks in Salmonella DNA is not random, then the systematic differences in transduction frequency of different markers as well as the influence of neighboring DNA on the transduction frequency can readily be accounted for by our model (Schmeiger, 1972; Chelala and Margolin, in press).

It should also be noted that coliphage λ, which has recently been shown to have an intimate evolutionary relationship to the Salmonella phage P 22 (2, 3, 10, 17), also appears to encapsulate its DNA from concatemers in a sequential manner (9). This is surprising in view of the facts that DNA encapsulation is in almost all cases site-specific and readily tolerates heads filled with much less than a headful of DNA. On the other hand, Weisberg and Sternberg (personal communication) have recently found conditions which allow λ to encapsulate bacterial DNA by the headful and thereby to act as vector in generalized transduction of E. coli genetic markers. Thus, we might speculate that λ and P22 might have some similarity in DNA encapsulation mechanism underlying their greatly different modes of head assembly and DNA maturation.

In conclusion, the P22 head assembly and DNA encapsulation process is now quite well defined in terms of the substrates (DNA concatemers and proheads) and the mechanism (sequential encapsulation from ends).

ACKNOWLEDGMENTS

We thank J. A. King and S. Casjens for valuable discussions. This work was supported by grants from the National Institutes of Health (GM18973-02) and the American Cancer Society (VC18B). D. Botstein holds a Career Development Award from the National Institutes of Health (5K04GM7032502); B. -K. Tye was supported by NIH Grant No. CA12174-03.

REFERENCES

1. Botstein, D., J. Mol. Biol. 34:621–641 (1968).
2. Botstein, D., Chan. R. K., and Waddell, C. H., Virology 49:268–282 (1972).
3. Botstein, D., and Herskowitz, I., Nature (in press).
4. Botstein, D. and Matz, M. J., J. Mol. Biol. 54:417–440 (1970).

5. Botstein, D., Waddell, C. H., and King, J., J. Mol. Biol. 80:669–695 (1973).
6. Casjens, S. and King, J. A., J. Supramol. Struct. 2:150 (1974).
7. Chan, R. K., Ph. D. Thesis, Massachusetts Institute of Technology (1974).
8. Ebel-Tsipis, J., Botstein, D., and Fox, M. S., J. Mol. Biol. 71:433–448 (1972).
9. Emmons, S. W., J. Mol. Biol. 83:511–525 (1974).
10. Gemski, P., Jr., Baron, L. S., and Yamamoto, N., Proc. Nat. Acad. Sci. U.S. 69:3110–3114 (1972).
11. Gilbert, W., and Dressler, D., Cold Spring Harbor Symp. Quant. Biol. 33:473 (1968).
12. Hershey, A. D. and Burgi, E., Proc. Nat. Acad. Sci. U.S. 53:325–328 (1965).
13. Inman, R. B., J. Mol. Biol. 28:103–116 (1967).
14. Jarvik, J. and Botstein, D., Proc. Nat. Acad. Sci. U.S. 70:2046–2050 (1973).
15. King, J., Lenk, E. V., and Botstein, D., J. Mol. Biol. 80:697–731 (1973).
16. Rhoades, M., MacHattie, L. A., and Thomas, C. A., Jr., J. Mol. Biol. 37:21–40 (1968).
17. Skalka, A. and Hanson, P., J. Virol. 9:583–593 (1972).
18. Streisinger, G., Emrich, J., and Stahl, M. M., Proc. Nat. Acad. Sci. U.S. 57:292–295 (1967).
19. Thomas, C. A., Jr., Kelly, T., Jr., and Rhoades, M., Cold Spring Harbor Symp. Quant. Biol. 33:417–424 (1968).
20. Tye, B. K., Chan, R. K., and Botstein, D., J. Mol. Biol. 85:485–500 (1974).
21. Tye, B. K., Huberman, J. A., and Botstein, D., J. Mol. Biol. 85:501–532 (1974).