Cell, Vol. 28, 315–324, February 1982, Copyright © 1982 by MIT

Incompletely Base-Paired Flip-Flop Terminal Loops Link the Two DNA Strands of the Vaccinia Virus Genome into One Uninterrupted Polynucleotide Chain

Bahige M. Baroudy, Sundararajan Venkatesan and Bernard Moss
Laboratory of Biology of Viruses
National Institute of Allergy and Infectious Diseases
Bethesda, Maryland 20205

Summary

The nature of the ends of the vaccinia virus genome was determined by nucleotide sequencing. Our finding of terminal hairpins indicated that the linear double-stranded DNA molecule consists of a single continuous polynucleotide chain. The 104 nucleotide apex of the hairpin contains predominantly A and T residues and is incompletely base-paired. These loops exist in two forms, which when inverted with respect to each other are complementary in sequence. Both forms of the 104 nucleotide loop are present in nearly equimolar amounts at each end of the genome. A set of 13 tandem 70 bp repeats begins 87 bp from the proximal segment of the terminal loop, followed by a unique sequence of 325 bp, and then by a second set of 18 tandem 70 bp repeats. The sequence of the 70 bp repeats reveals a 13 bp internal redundancy. Self-priming and de novo start replication models, which involve a site-specific nick in one DNA strand proximal to the 104 nucleotide loop, account for the observed sequence inversions and incomplete base-pairing. Similar mechanisms may be involved in replication of the ends of the eucaryotic chromosome.

Introduction

Because all known DNA polymerases require a primer with a free 3′ OH end, special mechanisms exist for replication of the termini of linear genomes. Some phages solve the problem by passing through a circular or concatameric stage (Watson, 1972). As an alternative, Cavalier-Smith (1974) proposed that the ends of eucaryotic and some viral chromosomes are palindromes. By folding into base-paired hairpin structures, the ends of the genome could be replicated by the combined actions of DNA polymerase, DNA ligase and a specific endonuclease. Self-priming variations of this mechanism have been proposed for the replication of single-stranded parvovirus DNA (Straus et al., 1976; Tattersall and Ward, 1976; Hauswirth and Berns, 1977; 1979), which contains palindromic ends (Astell et al., 1979; Berns et al., 1979; Salzman and Fabisch, 1979). As another variation of this model, Bateman (1975) postulated that the ends of eucaryotic chromosomes exist as self-complementary hairpin structures rather than as extended palindromes. Evidence consistent with terminal hairpins or cross-links near the ends of the vaccinia virus genome was obtained by Geshelin and Berns (1974). They found that the two DNA strands could not be separated and were able to visualize large single-stranded DNA circles by electron microscopy of denatured samples. Furthermore, Forte and Fangman (1976, 1979) found that yeast chromosomal DNA rapidly renatured and had properties consistent with hairpin ends. The similar properties of vaccinia viral and yeast DNA suggested that the former might be a useful model for the eucaryotic chromosome.

Vaccinia virus, the prototypal member of the poxvirus family, contains a linear double-stranded DNA genome of approximately 180,000 bp (Geshelin and Berns, 1974). In addition to the putative terminal hairpin or near-terminal cross-link, there is a 10,000 bp inverted terminal repetition (Garon et al., 1978; Wittek et al., 1978) containing two sets of many tandem repeats (Wittek and Moss, 1980). We undertook the present study to determine the physical nature of the cross-links holding the two strands of DNA together as well as the nucleotide sequences of the tandem repeats and flanking regions. We found that at each end of the genome, the two DNA strands formed a continuous polynucleotide loop. However, the terminal 104 nucleotides cannot completely base-pair and exist in two alternative forms that are inverted and complementary to each other. Several DNA replication models that produce flip-flop sequence inversions and maintain incompletely base-paired loops were considered.

Results

Nucleotide Sequence of the 70-Base-Pair Repeat

The 10,000 bp inverted terminal repetition contains two sets of repetitive Hinf I, Taq I and Mbo II sites, suggesting many tandem 70 bp repeating elements (Wittek and Moss, 1980). To sequence the latter, a phage λ recombinant constructed by digesting the ends of the vaccinia genome with a single-strand-specific nuclease and adding synthetic Eco RI linkers was used (Wittek et al., 1980). This recombinant contains the 9000 bp terminal Eco RI fragment except for the final 50 bp (Wittek and Moss, 1980). After digestion of the recombinant DNA with Hinf I or Taq I, the multimolar 70 bp fragment was purified by polyacrylamide gel electrophoresis, then labeled with polynucleotide kinase. Strand separation was achieved by hybridization to one strand of the 9000 bp DNA segment that had been cloned in the single-stranded phage f1 (Venkatesan and Moss, 1981). The sequence, determined by the technique of Maxam and Gilbert (1980), confirmed the 70 bp length of the repeat and the three expected restriction endonuclease sites, and also revealed an internal redundancy (Figure 1). Starting at the fourth nucleotide from the Hinf I cleavage site, there is a 13 base sequence

CCATCAGAAAGAG closely followed by a run of As and Ts. An identical 13 base sequence, interrupted by a 6 base insertion GAAGAG and also followed by As and Ts, begins 36 nucleotides from the Hinf I site.

Sequence of Individual Repeats and Flanking Regions

To facilitate further analysis, the λ recombinant was cleaved with Eco RI and Sal I, and the terminal 3700 bp of the 9000 bp vaccinia segment was subcloned in pBR322. Restriction endonuclease analysis revealed a unique Bst E II site located between the two sets of tandem repeats (Figure 2). This site was labeled with polynucleotide kinase, and the linearized plasmid was then digested with Eco RI and Sal I. After purification of the two labeled DNA fragments, Hinf I, Taq I and Mbo II sites were mapped by the partial digestion procedure of Smith and Birnstiel (1976). Analysis of the left fragment, corresponding to the end of the vaccinia genome, is shown in Figure 2 (lanes 0–10). The finding of 13 apparently equally spaced Hinf I sites and 14 Taq I sites suggested the presence of 13

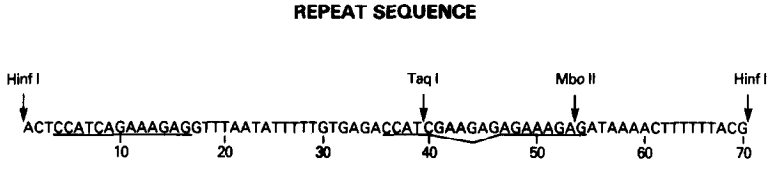

Figure 1. Nucleotide Sequence of 70 Base Pair Repeat

Hinf I and Taq I fragments were end-labeled, strand-separated and sequenced by the Maxam–Gilbert procedure (1980). Internal redundant sequences are underlined.

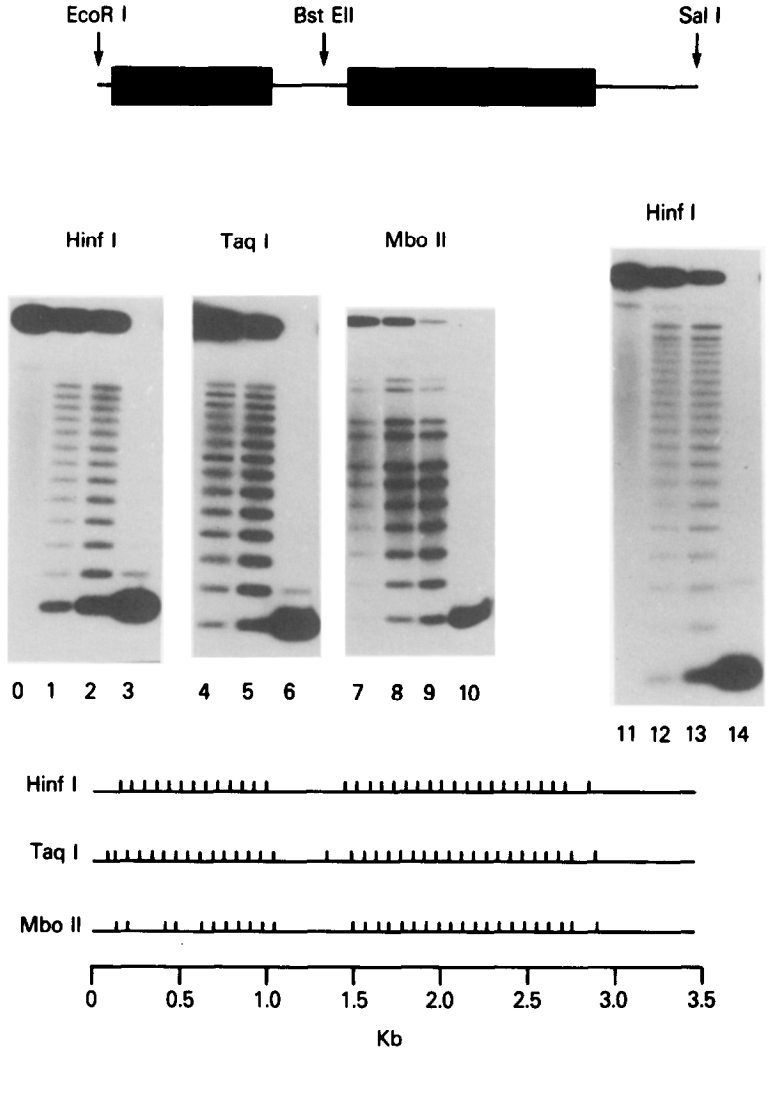

Figure 2. Maps of Repeating Hinf I, Taq I and Mbo II Sites

(Top) Representation of 3700 bp terminal segment of vaccinia genome cloned in pBR322. Synthetic Eco RI site is shown on left, and natural Sal I site on right. The Bst E II site occurs between the two sets of tandem repeats indicated by filled boxes.

(Middle) Autoradiograms of partial Hinf I, Taq I and Mbo II digests. Recombinant plasmid was cleaved at unique Bst E II site, end-labeled with α-^{32}P-GTP with DNA polymerase (Klenow fragment), then digested with Eco RI and Sal I. The two labeled fragments were purified by agarose gel electrophoresis. The Eco RI–Bst E II fragment incubated with 0 (lane 0), 0.22 (lane 1), 0.45 (lane 2), 4.5 (lane 3) units of Hinf I; 0.1 (lane 4), 0.2 (lane 5), 2 (lane 6) units of Taq I; 0.02 (lane 7), 0.04 (lane 8), 0.06 (lane 9), 3.3 (lane 10) units of Mbo II was subjected to agarose gel electrophoresis and autoradiographed. Similarly, the Bst E II–Sal I fragment was digested with 0 (lane 11), 0.22 (lane 12), 0.45 (lane 13), 4.5 (lane 14) units of Hinf I and analyzed as described above.

(Bottom) Restriction maps derived by partial digestion procedure described above.

complete repeats, in agreement with previous results (Wittek and Moss, 1980). The third, fourth and seventh repeats, however, are missing Mbo II sites and therefore have some sequence divergence. A similar analysis of the adjacent fragment indicated at least 18 repeats (Figure 3, lanes 11–14) with sites for all three restriction enzymes (not shown).

The strategy used for sequencing the terminal region of the cloned segment and the region between the two sets of repeats is shown in Figure 3. Contiguous sequence information was obtained for the terminal portion including some plasmid DNA and the added Eco RI linker, through four complete repeats. The sequence starting immediately after the synthetic linker is shown at the top of Figure 4. Perhaps most notable are the runs of Ts and As in the first 26 nucleotides and an 8 base palindrome, GATCGATC, at nucleotide 67 that is composed of two identical 4 base palindromes. The first repeat starts at nucleotide 87 with the sequence CCATCAG . . . (Figure 4), which also is the start of the 13 bp internal redundant sequence. The sequences of the first and second repeats are identical to the consensus 70 bp sequence of Figure 1, whereas those of the third and fourth repeats differ by the absence of single A residues, causing loss of Mbo II sites. The last repeat of the first set ends with the sequence GAGA, which also precedes the first repeat as well as the internal redundant sequence.

The region between the two sets of repeats appeared unremarkable except for frequent runs of As and Ts. The second set of repeats started with CATCAG . . ., 325 bp after the first set (nucleotide number 381, bottom of Figure 4). However, after the expected 12 nucleotides, the sequence AATAAAA occurs in place of GTTT. In addition, an A–T substitution occurs several nucleotides downstream. After this point, the sequence does not deviate from the consensus sequence for at least the next two repeats. Thus identical tandem 70 bp sequences were found in both sets of repeats.

Isolation of the Ends of Virion DNA

All of the experiments described above were carried out with cloned DNA; however, approximately 50 bp at the end of the vaccinia genome were removed by single-strand nuclease digestion before the addition of linkers necessary for cloning at the Eco RI site of λ. Thus the sequence shown in Figure 4 is missing the most intriguing part, the putative cross-link near or at the end of the genome that holds the two strands together. Only by analyzing virion DNA could this additional information be obtained.

Because of the large size of the vaccinia genome, we used a two-step procedure to isolate the ends. First, the DNA was cleaved with Sal I restriction endonuclease, which cuts within the inverted terminal repetition, and the 3700 bp fragment from both ends of the genome was isolated by preparative agarose gel electrophoresis. Next, the isolated Sal I fragment was cleaved with Pvu I, and terminal transferase was used to add a ^{32}P-cordycepin monophosphate residue to accessible 3′ ends. Pvu I was chosen because it cleaves the Sal I end fragment at a unique site before the start of the tandem repeats (Figure 4). We therefore expected to resolve two fragments by polyacrylamide gel electrophoresis, a large fragment containing all the tandem repeats and a small one representing the terminus. Although this was the basic result, the small fragment was resolved into two bands. By optimizing the electrophoretic conditions, a clear separation of the slow (S) and fast (F) Pvu I fragments was obtained (Figure 5, part I, lane T).

Because the Sal I fragment came from both ends of the genome, we considered the possibility that there were distinctive right and left termini. To resolve this question, we isolated the two ends of the genome. Hind III, which cuts outside of the 10,000 bp inverted terminal repetition, was used to digest virion DNA, and the right 27,000 bp B fragment was separated from the left 20,000 bp C fragment by preparative agarose gel electrophoresis. The two Hind III fragments were cleaved with Pvu I, labeled with ^{32}P-cordycepin tri-

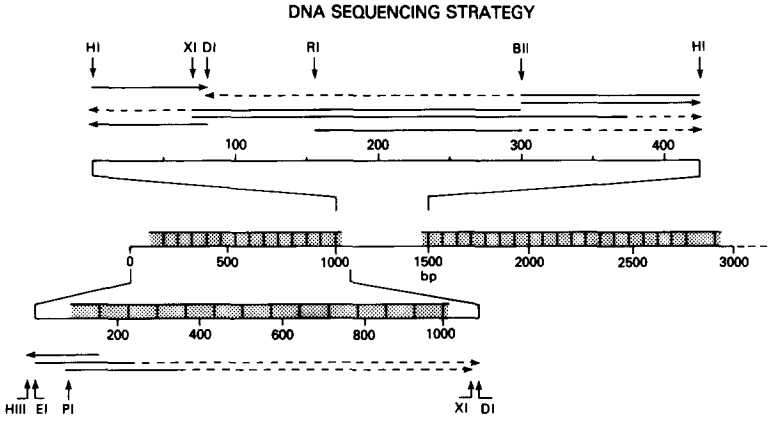

Figure 3. DNA Sequencing Strategy

The left-terminal 3000 bp segment of the cloned Eco RI–Sal I fragment is shown. The region containing tandem repeats is stippled, and the repetitive Hinf I sites are indicated. The strategies used to sequence the terminal segment are indicated at the bottom, and the region between the two sets of repeats at the top. The site of labeling is indicated by the start of the arrow, and the secondary cleavage site by the arrowhead. Useful sequence information was obtained in the region indicated by the solid uninterrupted portion of the arrow. HI, Hinf I; XI, Xba I; RI, Rsa I; BII, Bst II; HIII, Hae III; EI, Eco RI; PI, Pvu I; DI, Dde I. The Hae III site was in plasmid DNA.

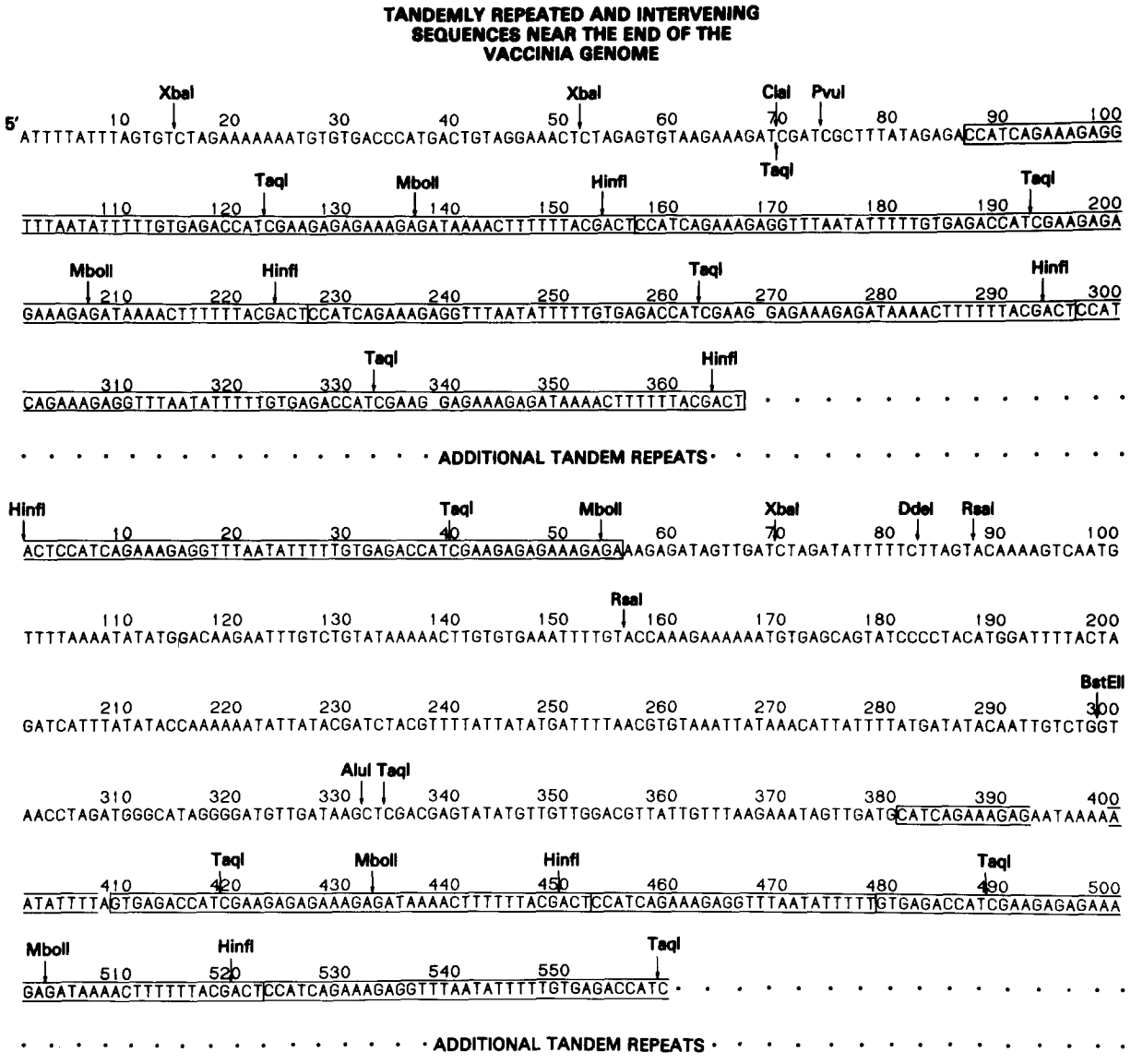

Figure 4. Nucleotide Sequence of Cloned Terminal Region of the Vaccinia Genome

(Top) Terminal sequence and four complete tandem repeats of first set. The first nucleotide shown was immediately adjacent to the synthetic Eco RI linker. Tandem 70 bp repeats are boxed. The consensus A residues at positions 268 and 338 of the second and third repeats are absent. (Bottom) Sequence of the region between the sets of repeats flanked by the last repeat of the first set and the first few repeats of the second set. The first nucleotide is the Hinf I cleavage site in the last repeat of the first set.

phosphate and subjected to polyacrylamide gel electrophoresis. Significantly, both F and S bands were obtained from each end of the genome (Figure 5, part I, lanes B and C). The amount of radioactive material was determined by cutting out the bands and counting them. In each case, the ratio of F to S was approximately 0.8. We concluded that F and S bands were derived, in nearly equimolar amounts, from both ends of the genome.

To examine the possibility of technical artifacts associated with electrophoresis, the F and S bands were eluted from the gel and rerun. As shown in Figure 5, part II, each band ran true. When the F and S band were subjected to electrophoresis under denaturing conditions, however, they were found to have identical lengths of about 250 nucleotides (Figure 5, part III). Thus the F and S fragments appeared to differ in sequence only.

Sequence of the Terminal Loop

Our primary goal in end-labeling the terminal Pvu I fragments was to sequence them. We considered that if the terminal hairpin model was correct, only a single 3′ end would be available for terminal transferase labeling, and therefore the F and S Pvu I fragments should be suitable for DNA sequencing without need for further manipulation. Furthermore, if the sequence could be read around the loop, direct proof of the

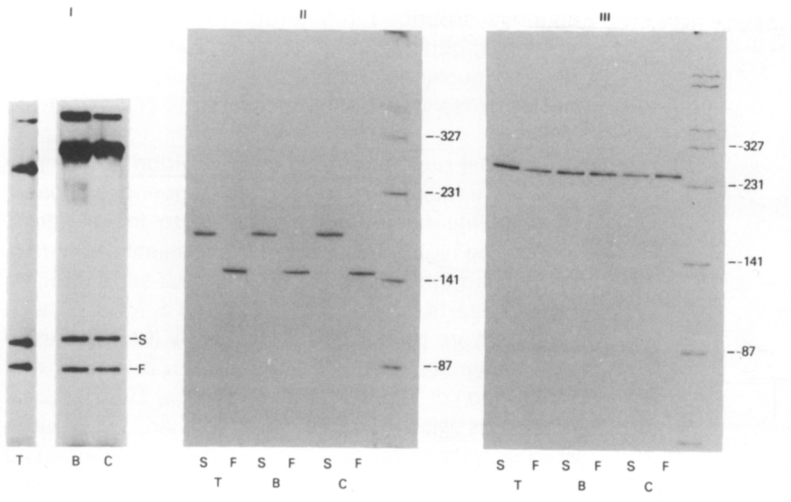

Figure 5. Terminal Pvu I Fragments from the Left and Right Sides of the Vaccinia Genome

(I) DNA isolated from vaccinia virus particles was cleaved with Sal I or Hind III. The single terminal Sal I fragment (lane T) or the right (lane B) and left (lane C) terminal Hind III fragments were purified by preparative agarose gel electrophoresis and cleaved with Pvu I. Accessible 3′ ends were labeled with terminal transferase and ^{32}P-cordycepin triphosphate. Autoradiograms of 6% polyacrylamide gels are shown. Lane S, slow migrating band; lane F, fast migrating band. (II) The S and F bands from I were eluted and subjected to a second round of polyacrylamide gel electrophoresis. Extreme right lane, Taq I fragments of ϕX174 replicative form DNA; lengths indicated in bases. (III) The S and F fragments from I were denatured and subjected to electrophoresis using conditions for analyzing nucleotide sequences. Extreme right lane, Taq I fragments of ϕX174; lengths indicated in bases.

model would be obtained.

Sequences of the F and S Pvu I fragments, determined by the technique of Maxam and Gilbert (1980), are shown in Figure 6. Reading upward in the 3′ to 5′ direction, one can see that, until the position indicated by dots, the sequences of the F and S fragments are identical to each other and to the cloned DNA sequence in Figure 4 (nucleotides 32 to 1). At that point, the cloned DNA sequence ends, because the terminal segment of the genome was removed by single-strand nuclease digestion before addition of synthetic linkers. Moreover, at that point, the sequence of the F and S fragments diverged. After 104 nucleotides of divergent sequence (marked by a second dot), the sequences of F and S again became identical to each other and complementary to the initial common region as well as to the strand of cloned DNA indicated in Figure 4 (nucleotides 1 to 20). Using additional applications of material and different electrophoretic conditions, we obtained a readable sequence of 216 nucleotides beginning 30 nucleotides after the labeled Pvu I site, extending around the 104 nucleotide loop, which must form the apex of the hairpin, and continuing to within seven nucleotides of the Pvu I site on the opposite strand. Except for the 104 nucleotide loop, there was base-for-base agreement with both strands of the cloned DNA segment. The sequence was confirmed by isolating the terminal Cla I fragments from both ends of the genome and sequencing them after labeling with DNA polymerase.

The relationship of the 104 nucleotide divergent F and S sequences is shown in Figure 7. When compared in the same orientation, they appear quite different. When examined in opposite polarities, however, it is evident that they are inverted and complementary (flip-flopped).

Another significant feature was noted when attempts were made to write the terminal loop structure in a base-paired fashion (Figure 8). For the proximal region—that is, that portion in which the F and S fragments are identical and which is present in the cloned DNA segment—perfect base-pairing was obtained. For the distal 104 nucleotides that form the terminal loop, however, there was extensive mismatching.

Discussion

The nature of the ends of the vaccinia virus genome was determined by nucleotide sequencing. Our finding of terminal hairpins indicated that the linear double-stranded DNA molecule consists of a single continuous polynucleotide chain. Interestingly, two alternative hairpin loops were found at each end of the genome. Restriction endonuclease fragments containing a fast (F) loop migrated ahead of those containing a slow (S) loop in nondenaturing polyacrylamide gels, whereas their mobilities were identical under denaturing conditions. Further analysis indicated that both F and S hairpin loops are 104 nucleotides long, but are inverted and complementary (flip-flopped) in sequence relative to each other. Other novel features of the loops are their high, 92%, AT content and incomplete base-pairing. Maximum base-paired F and S loop structures, shown in Figure 8, each contain two G–C pairs, 44 A–T pairs and 16 unpaired bases, including four in the apex. All but the most proximal unpaired bases are on the same side of the two-dimensional structure, so that it resembles a bobby pin rather than a straight hairpin. An almost identical folded structure was obtained by a computer program, originally designed for RNA by Zuker and Stiegler (1981) and modified by R. Feldman, which uses stacking and destabilizing energies to determine confor-

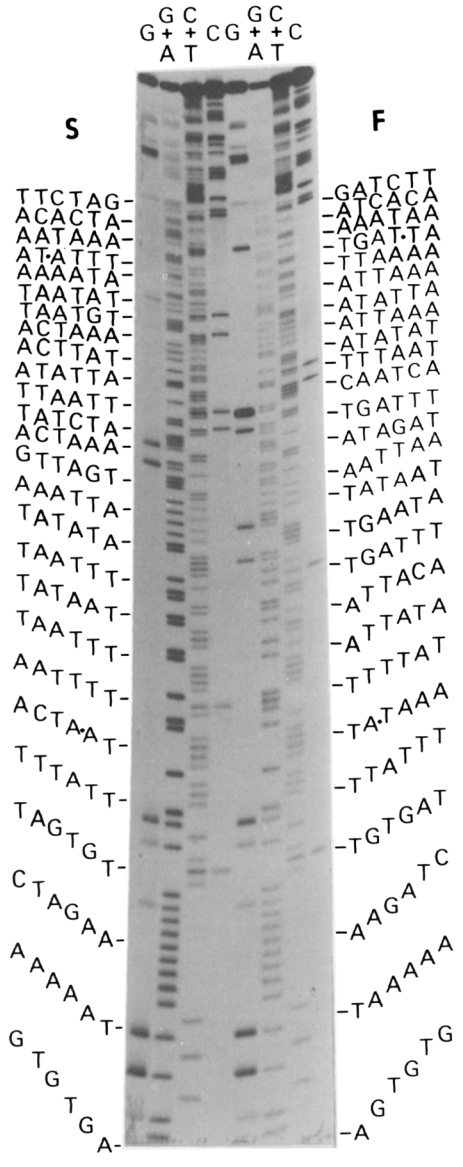

Figure 6. Nucleotide Sequence of the F and S fragments

Pvu I fragments F and S were 3'-end-labeled by use of terminal transferase and ^{32}P-cordycepin triphosphate, purified by polyacrylamide gel electrophoresis and sequenced by the Maxam-Gilbert technique (1980). An autoradiogram of the sequence ladder is shown. The 104 nucleotide divergent portion of the written sequence has been set off by dots.

mation. The only difference was that in our structure (Figure 8), the first of the 104 nucleotides in the loop was unpaired, whereas in the computer model (not shown), the adjacent proximal nucleotide was unpaired instead. In view of their similar base-pairing potentials, the difference in electrophoretic migration rates of DNA fragments containing F and S loops must depend on subtle differences in their three-dimensional structures. It may be useful to probe these structures by examining the products of single-strand nuclease digestion. In this regard, the cloned DNA segment described in this report evidently had been formed by cutting the genome precisely at the base of the 104 nucleotide loop with a single-strand-specific nuclease isolated from vaccinia virus (Wittek et al., 1980).

Terminal restriction fragments of virion and cloned DNA were found to have identical sequences capable of complete base-pairing proximal to the F and S loops. The high AT content of the terminal segment is maintained for the first 26 bp. The first set of tandem repeats starts 87 bp from the F and S loops. These repeats were recognized originally by the generation of a multimolar 70 bp fragment upon digestion with Hinf I, Taq I or Mbo II (Wittek and Moss, 1980). Partial digestions with these enzymes indicated that there are two sets of repeats separated by an intervening region. The first set contains 13 repeats, of which three are missing Mbo II sites. The second set contains 18 repeats, with all three restriction sites. Analysis of virion DNA revealed that neither the arrangement nor the approximate number of repeats was altered during cloning in λ (J. A. Cooper, personal communication). In addition, the Mbo II sites were missing from the original recombinant prior to subcloning in pBR322 (R. Wittek, personal communication). The present finding of one additional repeat in the second set of the pBR322 recombinant is probably due to the improved method of analysis.

A consensus repeat sequence was obtained by analysis of the multimolar 70 bp fragment obtained by Hinf I and by Taq I digestions. The 70 bp length of the repeat as well as the restriction sites were confirmed, and an internal periodicity was recognized. The latter consisted of a 13 bp sequence and a run of A and T residues followed by an identical 13 bp sequence, albeit with a 6 bp insertion, and a similar number of A and T residues. The sequences of the first two repeats are identical to that of the consensus sequence, but both the third and fourth repeats are missing a single A residue. This deletion results in loss of the Mbo II site as predicted from the partial mapping procedure.

An unremarkable sequence of 325 bp separates the two sets of repeats. Previously, this intervening region was considered to be longer because it was defined as the segment between two Hinf I sites (Wittek and Moss, 1980). The second set of repeats begins almost precisely in phase with the first set; however, the first repeat of the second set contains small deletions and substitutions. Nevertheless, the next two repeats are identical to the consensus sequence, indicating that the standard repeating element is the same in both sets. Presumably, unequal cross-overs are responsible for correcting and multiplying the tandem sequences. In this regard, virus isolates with different numbers of tandem repeats and unstable isolates in which an entire set of repeats and intervening region have been amplified many times have been found (Moss et al., 1981). Because of the

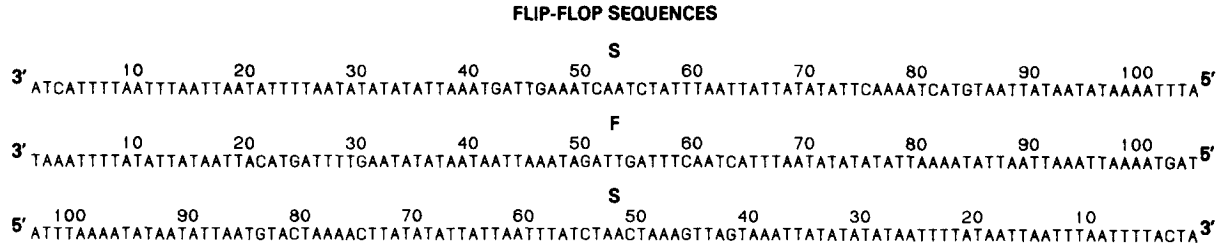

Figure 7. Inverted Complementary (Flip-Flop) Sequences

The 104 nucleotide divergent segments of the F and S forms of the terminal loop are shown in 3′ to 5′ or 5′ to 3′ direction as indicated.

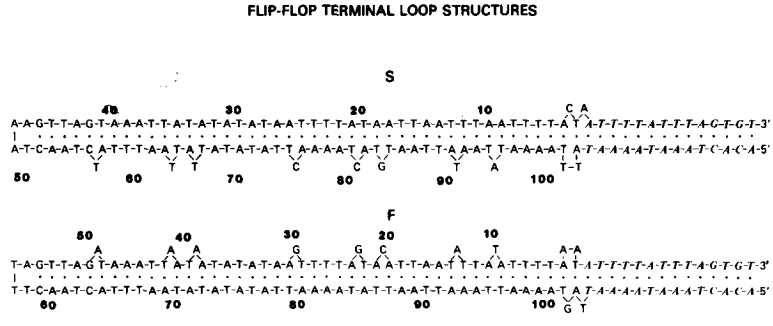

Figure 8. Incomplete Base-Pairing of the Terminal Loops

The F and S terminal loop forms are shown with maximum base-pairing. The nucleotides of the divergent 104 nucleotides are numbered. A portion of the contiguous identical sequence is shown in italics on the right.

paucity of restriction endonuclease sites in the region just beyond the second set of repeats, there is still more than 600 bp of unsequenced DNA before the transcriptionally active region of the vaccinia genome (Venkatesan et al., 1981).

Although no sequence information is yet available, there are data consistent with hairpin ends in the genomes of other orthopoxviruses (Wittek et al., 1977; Mackett and Archard, 1979) and of other members of the poxvirus family including avian poxviruses (Gafford et al., 1978), molluscum contagiosum (Parr et al., 1977) and stomatitis papulosa virus (Menna et al., 1979). African swine fever virus, a member of the iridovirus family, also has a genome that rapidly renatures, suggesting the presence of cross-linked ends (Ortin et al., 1979).

The small single-stranded DNA genomes of parvoviruses have imperfect terminal palindromes that can base-pair to form partial duplex structures (Astell et al., 1979; Berns et al., 1979; Salzman and Fabisch, 1979). The 3′ end palindromes of defective and nondefective parvoviruses are 115 to 125 nucleotides long and therefore similar in length to the 104 nucleotide, incompletely base-paired hairpin of vaccinia. Moreover, in the case of the adenovirus-associated virus, there is a flip-flop sequence inversion (Spear et al., 1977). The finding of covalently linked double-stranded forms of parvovirus DNA as replicative intermediates (Straus et al., 1976; Bourguignon et al., 1976; Hauswirth and Berns, 1977, 1979) makes the analogy with vaccinia virus compelling. In contrast to vaccinia virus, however, the parvovirus palindromes are GC-rich and form Y-shaped maximum base-paired structures.

Studies with yeast chromosomal DNA indicate that they rapidly renature, suggesting the presence of cross-linked ends (Forte and Fangman, 1976, 1979). There is also evidence that the DNA strands at one end of replicating mitochondrial DNA of Paramecium aurelia are covalently linked (Goddard and Cummings, 1977). The presence of hairpin ends on DNA viruses and eucaryotic chromosomes implies that this feature has an important role in DNA replication. Such structures were predicted on theoretical grounds as a way of replicating the ends of linear chromosomes without need for circularization or concatamerization (Cavalier-Smith, 1974; Bateman, 1975). Moreover, evidence that parvoviruses replicate by a self-priming variation of the Cavalier-Smith model has been obtained (Straus et al., 1976; Bourguignon et al., 1976; Hauswirth and Berns, 1977, 1979).

Although there is a paucity of data regarding the mechanism of vaccinia virus DNA replication, reports suggest that nicks are introduced near the ends of the parental genomes soon after infection, and that replication starts near the ends of the DNA molecule and occurs in a semiconservative fashion involving discontinuous DNA synthesis and possibly RNA primers (Esteban and Holowczak, 1977a, 1977b; Pogo, 1977; Pogo and O'Shea, 1978; Pogo et al., 1981). Eye structures and forked ends have been seen by electron microscopy (Esteban et al., 1977). Introduction of cross-links—that is, ligation of nicks—occurs relatively late in infection (Esteban and Holowczak, 1977a; Pogo, 1977). Although no evidence for concatamers was obtained by alkaline sucrose gradient centrifugation (Esteban and Holowczak, 1977b), DNA restriction fragments apparently derived from head-

to-head and tail-to-tail concatamers have been identified (Moyer and Graves, 1981).

A self-priming DNA replication model for vaccinia virus that accounts for the occurrence of incompletely base-paired hairpin loops with flip-flop sequence heterogeneity is shown on the left side of Figure 9. The first step is postulated to be a site-specific nick resulting in a free 3′ OH end that then serves as a primer for DNA replication. Strand displacement with unfolding of the hairpin loop occurs in the second step. In the third step, the extended palindrome reforms hairpin loops, and in the fourth, DNA replication continues by a strand displacement mechanism. In the simple model shown, replication has been initiated at both ends, and mature DNA molecules are formed by ligation (step 5). Inspection of the mature DNA molecules formed by this scheme reveals that a flip-flop sequence inversion has occurred. Moreover, since the ends are replicated as extended palindromes, the lack of complete base-pairing in the loops is readily explained. Nevertheless, more complex variations of this model can be envisioned. For example, if initiation has occurred only at one end of the DNA, then replication around the opposite unnicked hairpin would lead to formation of a concatamer, as in parvovirus DNA replication (Straus et al., 1976; Tattersall and Ward, 1976; Hauswirth and Berns, 1979). In this situation, nicking would be required to separate the daughter strands. Furthermore, self-priming need be used only for replication of the ends of the genome, and discontinuous synthesis of leading or lagging strands or both also could occur.

A de novo start model, similar to that proposed by Bateman (1975) for the replication of eucaryotic chromosomes, is shown on the right side of Figure 9. DNA replication is assumed to initiate with RNA primers. Nicking and rearrangement is necessary to separate the daughter strands of DNA leading to flip-flop sequence inversions.

A third model in which the first step is a site-specific nick to form sticky ends and the second is circularization or concatamerization is also possible (Figure 10). Following replication by mechanisms similar to those used in procaryotic systems (Kornberg, 1981), site-specific nicks, refolding and ligation would occur. One constraint of the circularization model, however,

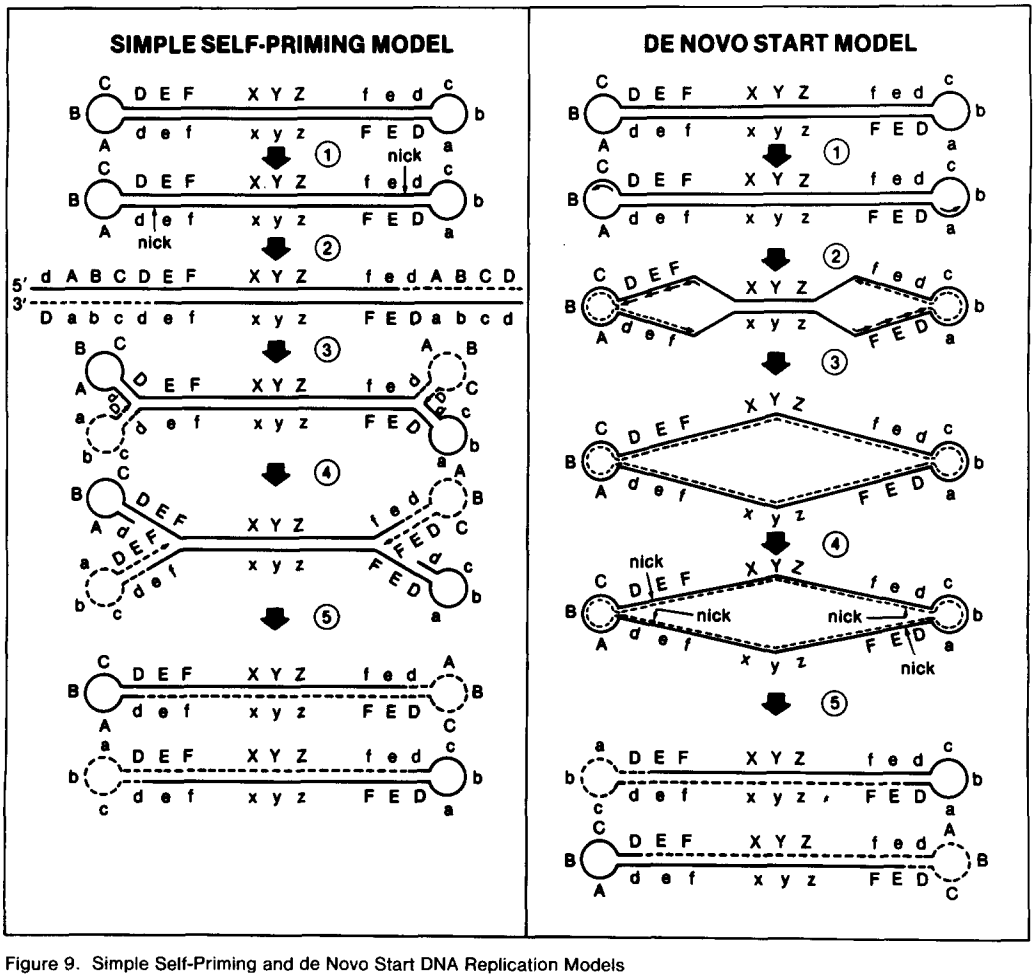

Figure 9. Simple Self-Priming and de Novo Start DNA Replication Models

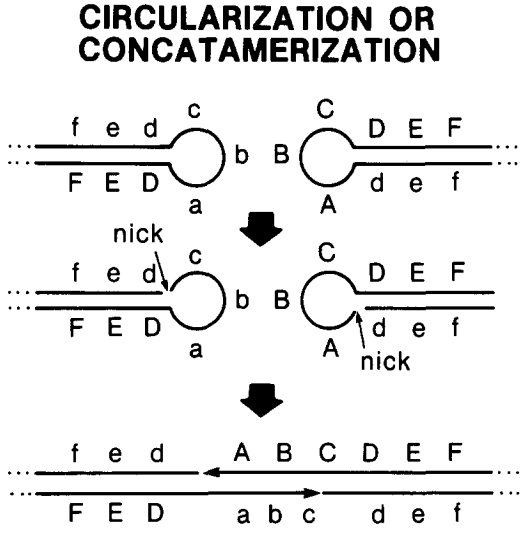

Figure 10. Circularization or Concatamerization Models

is that every molecule must have an F loop at one end and an S loop at the other. Nevertheless, transfer of loops from one end to the other could occur by homologous recombination within the inverted terminal repetition. By contrast, all possible structural variations including F or S loops at both ends could be replicated by the first two models considered. Terminal restriction fragments obtained from any concatameric replicative intermediates of the first two models would appear as head-to-head or tail-to-tail dimers, whereas the ends of a circular molecule would appear as head-to-tail dimers.

Regardless of which model is correct, the occurrence of a flip-flop sequence inversion demands that there be site-specific nicking and religation. Moreover, this nick must occur proximal to the last mismatched base. Possibly, the incomplete base-pairing and extraordinarily high AT content of the terminal segment facilitates the transition between palindrome and hairpin forms. The enzyme that nicks the DNA has not been identified. Although the DNAases packaged within the virus particle have single-strand specificity in vitro (Pogo and O'Shea, 1977; Rosemond-Hornbeak et al., 1974), the possibility that this activity is modulated in vivo to cut at a specific site must be considered. Alternatively, a multifunctional enzyme similar to the gene A protein of ϕX174 (Kornberg, 1981) might be involved.

Experimental Procedures

Isolation of DNA Fragments

DNA was obtained from purified vaccinia virus (strain WR) (Garon et al., 1978), λ or f1 recombinants containing the terminal Eco RI fragment of this virus (Wittek et al., 1980; Venkatesan and Moss, 1981) or a newly constructed plasmid recombinant designated pAG5. The latter was derived by digesting the λ recombinant with Eco RI and Sal I restriction endonucleases, then ligating the fragment containing the terminal segment of the vaccinia genome into Eco RI- and Sal I-digested pBR322. Escherichia coli HB101 cells were transformed, and the identity of the recombinant plasmid was confirmed by restriction endonuclease analysis and agarose gel electrophoresis. Recombinant plasmids were purified by ethidium bromide–cesium chloride centrifugation as previously detailed (Venkatesan and Moss, 1981). Standard procedures were used for restriction endonuclease digestions and agarose gel electrophoresis (Wittek and Moss, 1980), and DNA fragments were purified by the glass extraction procedure of Vogelstein and Gillespie (1979) or by electrophoresis onto DEAE paper (Winberg and Hammarskjöld, 1980).

End-Labeling of DNA

DNA fragments (2–20 pmole) were treated with calf intestinal alkaline phosphatase, and the 5′ ends were labeled with γ-^{32}P-ATP and T4 polynucleotide kinase as described by Maxam and Gilbert (1980). For recessed 3′ ends, the Klenow fragment of E. coli DNA polymerase was used to fill in the first position with the appropriate α-^{32}P-dNTP essentially as described by Yang and Wu (1979). For overhanging 3′ ends, terminal deoxynucleotidyl transferase was used to add α-^{32}P-cordycepin monophosphate (Tu and Cohen, 1980).

Strand Separation of 70-Base-Pair Repeats

The multimolar 70 bp Hinf I or Taq I fragments were purified by electrophoresis on 10% polyacrylamide gels. After 5′-end-labeling, the strands were separated by hybridization to an excess of single-stranded DNA from the phage recombinant f1 SV1 (Venkatesan and Moss, 1981) in 0.6 M sodium chloride and 0.06 M sodium citrate at 65°C for about 18 hr. The hybridization products were resolved as three radioactively labeled bands by electrophoresis in 2% agarose gels. Sequence analysis confirmed that the upper band represented the single-stranded phage DNA hybridized to the complementary strand of the 70 bp repeat, and that the slower of the two lower molecular weight bands was the 70 bp duplex and the faster was the other strand of the 70 bp repeat.

DNA Sequencing

The chemical method of Maxam and Gilbert (1980) was used to sequence DNA restriction fragments labeled at a single 5′ or 3′ end. The four reactions used were G, G+A, T+C and C. The G+A reaction was carried out at 30°C instead of 20°C. The products were resolved on 40 or 80 cm, 8%, 12% or 20% polyacrylamide gels in 7 M urea (Sanger and Coulson, 1978). Nucleotide sequences were stored and analyzed by a computer program (Queen and Korn, 1980).

Acknowledgments

We thank Alan Gershowitz for constructing pAG5 plasmid recombinant, Norman Cooper for preparing plasmid DNA, James Rose and Kenneth Berns for helpful discussions, Richard J. Feldman, Elaine Winters and Edwin Sebring for assisting us in the use of the computer programs and Jeanne Carolan for typing the manuscript.

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

Received October 27, 1981

References

Astell, C. R., Smith, M., Chou, M. B. and Ward, D. C. (1979). Structure of the 3′ hairpin termini of four rodent parvovirus genomes: nucleotide sequence homology at origins of DNA replication. Cell *17*, 691–703.

Bateman, A. J. (1975). Simplification of palindromic telomere theory. Nature *253*, 379–380.

Berns, K. I., Hauswirth, W. W., Fife, K. H. and Lusby, E. (1979). Adeno-associated virus DNA replication. Cold Spring Harbor Symp. Quant. Biol. *43*, 781–786.

Bourguignon, G. J., Tattersall, P. J., and Ward, D. C. (1976). DNA of minute virus of mice: self-priming, non-permuted, single-stranded genome with a 5′-terminal hairpin duplex. J. Virol. *20*, 290–306.

Cavalier-Smith, T. (1974). Palindromic base sequences and replication of eukaryotic chromosomes. Nature *250*, 467–470.

Esteban, M., Flores, L. and Holowczak, J. A. (1977). Model for vaccinia virus DNA replication. Virology *83*, 467–473.

Esteban, M. and Holowczak, J. A. (1977a). Replication of vaccinia DNA in mouse L cells. I. *In vivo* DNA synthesis. Virology *78*, 57–75.

Esteban, M. and Holowczak, J. A. (1977b). Replication of vaccinia DNA in mouse L cells. III. Intracellular forms of viral DNA. Virology *82*, 308–322.

Forte, M. A. and Fangman, W. L. (1976). Naturally occurring crosslinks in yeast chromosomal DNA. Cell *8*, 425–431.

Forte, M. A., and Fangman, W. L. (1979). Yeast chromosomal DNA molecules have strands which are cross-linked at their termini. Chromosoma *72*, 131–150.

Gafford, L. G., Mitchell, E. B., Jr. and Randall, C. C. (1978). Sedimentation characteristics and molecular weights of three poxvirus DNAs. Virology *89*, 229–239.

Garon, C. F., Barbosa, E. and Moss, B (1978). Visualization of an inverted terminal repetition in vaccinia virus DNA. Proc. Nat. Acad. Sci. USA *75*, 4863–4867.

Geshelin, P. and Berns, K. I. (1974). Characterization and localization of the naturally occurring crosslinks in vaccinia virus DNA. J. Mol. Biol. *88*, 785–796.

Goddard, J. M. and Cummings, D. J. (1977). Mitochondrial DNA replication in Paramecium aurelia. Cross-linking of the initiation end. J. Mol. Biol. *190*, 327–344.

Hauswirth, W. W. and Berns, K. I. (1977). Origin and termination of adeno-associated virus DNA replication. Virology *78*, 488–499.

Hauswirth, W. W. and Berns, I. I. (1979). Adeno-associated virus DNA replication: nonunit-length molecules. Virology *93*, 57–68.

Kornberg, A. (1981). DNA Replication. (San Francisco: W. H. Freeman Company).

Mackett, M. and Archard, L. C. (1979). Conservation and variation in orthopoxvirus genome structure. J. Gen. Virol. *45*, 683–701.

Maxam, A. M. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

Menna, A., Wittek, R., Bachmann, P. A., Mayr, A. and Wyler, R. (1979). Physical characterization of a stomatitis papulosa virus genome: a cleavage map for the restriction endonucleases *Hind*III and *Eco*RI. Arch. Virol. *59*, 145–156.

Moss, B., Winters, E., and Cooper, N. (1981). Instability and reiteration of DNA sequences within the vaccinia virus genome. Proc. Nat. Acad. Sci. USA *78*, 614–618.

Moyer, R. W. and Graves, R. L. (1981). Replicating poxvirus DNA in cytoplasmic virosomes is concatameric. Abst. Fifth Inter. Cong. Virol., Strasbourg, France, p. 336.

Ortin, J., Enjuanes, L. and Vinuela, E. (1979). Cross-links in African swine fever virus DNA. J. Virol. *31*, 579–583.

Parr, R. P., Burnett, J. W. and Garon, C. F. (1977). Structural characterization of the Molluscum contagiosum virus genome. Virology *81*, 247–256.

Pogo, B. G. T. (1977). Elimination of naturally occurring crosslinks in vaccinia virus DNA after viral penetration into cells. Proc. Nat. Acad. Sci. USA *74*, 1739–1742.

Pogo, B. G. T., and O'Shea, M. T. (1977). Further characterization of deoxyribonucleases from vaccinia virus. Virology *77*, 56–66.

Pogo, B. G. T. and O'Shea, M. T. (1978). The mode of replication of vaccinia virus DNA. Virology *84*, 1–8.

Pogo, B. G. T., O'Shea, M. and Freimuth, P. (1981). Initiation and termination of vaccinia virus DNA replication. Virology *108*, 241–248.

Queen, C. L. and Korn, L. J. (1980). Computer analysis of nucleic acids and proteins. Meth. Enzymol. *65*, 595–609.

Rosemond-Hornbeck, H., Paoletti, E. and Moss, B. (1974). Single-stranded deoxyribonucleic acid-specific nuclease from vaccinia virus. Purification and characterization. J. Biol. Chem. *249*, 3287–3291.

Salzman, L. A. and Fabisch, P. (1979). Nucleotide sequence of the self-priming 3′ terminus of the single-stranded DNA extracted from the parvovirus Kilham rat virus. J. Virol. *30*, 946–950.

Sanger, F. and Coulson, R. (1978). The use of thin acrylamide gels for DNA sequencing. FEBS Letters *87*, 107–110.

Smith, H. O. and Birnstiel, M. L. (1976). A simple method for DNA restriction site mapping. Nucl. Acids Res. *3*, 2387–2398.

Spear, I. S., Fife, K. H., Hauswirth, W. W., Jones, C. J. and Berns, K. I. (1977). Evidence for two nucleotide sequence orientations within the terminal repetition of adeno-associated virus DNA. J. Virol. *24*, 627–634.

Straus, S. E., Sebring, E. D. and Rose, J. A. (1976). Concatamers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. Proc. Nat. Acad. Sci. USA *73*, 742–746.

Tattersall, P. and Ward, D. C. (1976). Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. Nature *263*, 106–109.

Tu, C.-P., and Cohen, S. N. (1980). 3′-end labeling of DNA with [α-^{32}P]cordycepin-5′-triphosphate. Gene *10*, 177–183.

Venkatesan, S. and Moss, B. (1981). *In vitro* transcription of the inverted terminal repetition of the vaccinia virus genome: correspondence of initiation and cap sites. J. Virol. *37*, 738–747.

Venkatesan, S., Baroudy, B. M. and Moss, B. (1981). Distinctive nucleotide sequences adjacent to mulitple initiation and termination sites of an early vaccinia virus gene. Cell *25*, 805–812.

Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. Proc. Nat. Acad. Sci. USA *76*, 615–619.

Watson, J. D. (1972). Origin of concatameric T7 DNA. Nature New Biol. *239*, 197–201.

Winberg, G. and Hammarskjöld, M.-L. (1980). Isolation of DNA from agarose gels using DEAE-paper. Application to restriction site mapping of adenovirus type 16 DNA. Nucl. Acids Res. *8*, 253–264.

Wittek, R. and Moss, B. (1980). Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. Cell *21*, 277–284.

Wittek, R., Barbosa, E., Cooper, J. A., Garon, C. F., Chan, H. and Moss, B. (1980). The inverted terminal repetition in vaccinia virus DNA encodes early mRNAs. Nature *285*, 21–25.

Wittek, R., Menna, A., Schümperli, D., Stoffel, S., Müller, H. K. and Wyler, R. (1977). *Hind*III and *Sst*I restriction sites mapped on rabbit poxvirus and vaccinia virus DNA. J. Virol. *23*, 669–678.

Wittek, R., Menna, A., Müller, H. K., Schümperli, D., Boseley, P. G. and Wyler, R. (1978). Inverted terminal repeats in rabbit poxvirus and vaccinia virus DNA. J. Virol. *28*, 171–181.

Yang, R. C. A. and Wu, R. (1979). BK virus DNA sequence coding for the amino-terminus of the T-antigen. Virology *92*, 340–352.

Zuker, M. and Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucl. Acids Res. *9*, 133–148.