

Restriction Endonucleases, Simian Virus 40, and the New Genetics

Daniel Nathans

Some 35 years ago the study of heredity took a chemical turn when Avery and his colleagues discovered that deoxyribonucleic acid—DNA—is the “transforming principle” that converts bacteria from one genotype to another. Watson and Crick’s structural model of DNA then provided the basis for investigating its role as the hereditary material. There followed rapid advances in the biochemistry of DNA replication, gene expression, and regulation of genes. Recently, fresh impetus has been given to the study of genetic mechanisms, particularly in higher organisms, by several methodological developments that have opened a new approach to the analysis of chromosomes: site-specific cleavage of DNA by restriction endonucleases and electrophoretic fractionation of the resulting fragments; recombination, cloning, and amplification of DNA segments from any source; rapid methods for determining the nucleotide sequence of DNA; site-directed in vitro mutagenesis; synthesis of polydeoxynucleotides of predetermined sequence; and the ability to introduce cloned, functioning genes into prokaryotic or eukaryotic cells. As a result of these developments even chromosomes that are largely inaccessible to classical genetic methods can now be analyzed piece by piece in chemical detail. Genes and signals can be altered at preselected sites, and the functional effect of such alterations can be determined. And active, synthetic genes can be constructed in vitro by recombination or by chemical synthesis.

Many investigators have contributed to the “new genetics.” Contributions from my own laboratory resulted from our studies of a model eukaryotic chromosome, that of a small mammalian tumor virus. I became interested in tumor viruses in the mid-1960’s when I was asked to give a lecture on this subject to Johns Hopkins University medical students. Although I had been working with an RNA coliphage (a bacterial virus) for some years, I knew very little about ani-

mal viruses. As I reviewed the tumor virus literature, I was impressed by the fact that simple viruses had a profound and permanent effect on the growth of cells in culture or in a living animal. Here was a microcosm of regulatory mechanisms related to the development of the virus itself and to the growth of animal cells, including neoplastic cells. At least some of these mechanisms appeared approachable with the tools of molecular genetics that had been so successfully used with bacterial viruses. Of course all of this was appreciated by a number of people in the tumor virus field, but to me it was an exhilarating revelation. I decided to take a leave of absence in order to explore experimental approaches to understanding viral tumorigenesis, and gradually to wind up my work on the RNA bacteriophage.

At the beginning of 1969 I went to the Weizmann Institute of Science in Israel, where I worked with Ernest Winocour and Leo Sachs and had a chance to read and think without interruption. During that spring I received a letter from a colleague in Baltimore, Hamilton Smith, telling me about the enzyme he had discovered in the bacterium *Hemophilus influenzae* that had the biochemical properties of a restriction endonuclease. This aroused my interest immediately in the possibility that restriction endonucleases were “trypsins and chymotrypsins for DNA,” and prompted me to review the literature on bacterial restriction and modification, beginning with the initial observations of Luria and Human (1) and Bertani and Weigle (2). From the incisive work of Arber and his colleagues on the molecular genetics of DNA restriction and modification (3), and the biochemical characterization of purified restriction enzymes by Meselson and Yuan (4) and Smith and Wilcox (5), it seemed likely (as first suggested by Arber) that restriction enzymes could be used to digest DNA molecules into specific fragments, just as specific proteolytic enzymes are used to fragment proteins. If

the genomes of DNA tumor viruses could be dissected in this way, and if individual fragments of viral DNA could be isolated, one might be able to determine by chemical mapping which segments of the genome were responsible for the various biological activities of the virus, an approach analogous to that Shimura and I had taken earlier to determine the location of genes along the RNA of a bacterial virus (6).

I had already decided that the small papovavirus, simian virus 40 (SV40) (7), was the most tractable tumor virus to work with. This virus is a nonenveloped, icosahedral particle with a diameter of about 40 nanometers (Fig. 1). Its genome is a ring of duplex DNA with only about 5000 nucleotide pairs (8)—equivalent to a few genes—present as a typical eukaryotic minichromosome (9, 10). Despite its paucity of genetic information, SV40 seemed to have all the biological properties of immediate interest: it grew in the nucleus of monkey cells in culture (11), and it caused heritable changes in the growth of rodent cells; that is, it “transformed” them to tumorigenicity (12). As an initial experiment, I planned to survey the known restriction endonucleases for their ability to cleave SV40 DNA. On my return to Baltimore in the summer of 1969, DNA in hand, our dissection of the SV40 chromosome began.

Cleavage of SV40 DNA by Restriction Endonucleases

For our initial survey of restriction endonucleases Stuart Adler, working with me in the summer and fall of 1969, prepared restriction enzymes from *Escherichia coli* strains B (13), K (4), and K(P1) (14), and he obtained *Hemophilus* enzyme from Smith. To our delight, the *E. coli* B enzyme and the P1 enzyme each cleaved the SV40 DNA circle once, yielding full-length linear molecules, and the *Hemophilus* enzyme cleaved SV40 DNA several times (15). However, the *E. coli* K enzyme did not attack SV40

Copyright© 1978 by the Nobel Foundation.

The author is Professor of Microbiology at the Johns Hopkins University School of Medicine, Baltimore, Maryland 21205. This article is the lecture he delivered in Stockholm, Sweden, 8 December 1978, when he received the Nobel Prize in Physiology or Medicine, a prize he shared with Dr. Werner Arber and Dr. Hamilton O. Smith. Minor corrections have been made by the author. The article is published here with the permission of the Nobel Foundation and will also be included in the complete volume of *Les Prix Nobel en 1978* as well as in the series *Nobel Lectures* (in English) published by the Elsevier Publishing Company, Amsterdam and New York. Dr. Arber’s lecture appeared in the 27 July issue of *Science*, page 361, and that of Dr. Smith in the 3 August issue, page 455.

DNA at all. We tentatively concluded that SV40 DNA has no sequences recognized by the K enzyme, that the B and P1 enzymes opened the SV40 circle at a unique site specific for each enzyme, and that the *Hemophilus* endonuclease cut the viral DNA at several specific sites. Later we were surprised to find that the *Eco* B (16) restriction endonuclease, a complex adenosine triphosphate (ATP) and *S*-adenosylmethionine-dependent enzyme (13) ("class I" enzyme), does not break SV40 DNA at a specific site, even though the enzyme cuts each molecule once (17, 18). Therefore, *Eco* B and similar enzymes would not be useful for our purpose. However, Smith's *Hind* endonuclease, a structurally simpler enzyme, not dependent on ATP or *S*-adenosylmethionine (19), was shown by Kelly and Smith to break DNA at a specific nucleotide sequence (20), and by

Kathleen Danna and me to generate specific, electrophoretically separable fragments from SV 40 DNA (21) (Fig. 2). Smith's enzyme turned out to be a mixture of two different restriction endonucleases (*Hind* II and *Hind* III) (22, 23) each of which gave a characteristic electrophoretic pattern of fragments from SV40 DNA. Subsequently, newly discovered cleavage site-specific restriction enzymes of the *Hemophilus* type ("class II" enzymes), more than a hundred of which are now known (24), were used to cut SV40 DNA, each yielding its own distinctive digest pattern when the fragments were visualized by electrophoresis in acrylamide or agarose gels. Thus, digestion of DNA by class II restriction enzymes followed by gel electrophoresis appeared to yield homogeneous fragments derived from specific regions of the genome.

Cleavage Map of the SV40 Chromosome

To use fragments generated by restriction of SV40 DNA for mapping viral functions, we needed to locate the precise positions of restriction sites in the viral DNA; that is, to construct a "cleavage map" for each restriction enzyme. This was accomplished by Danna and George H. Sack, who first determined the size of fragments in a given digest and then their order in the circular SV40 genome (21, 23). The size of each fragment was determined initially by its relative yield or by electron microscopic length measurements, and later, by electrophoretic mobility relative to standards. The order of fragments in the viral genome was determined by electrophoretic analysis of isolated partial digest products, and by sequential digestion with different restriction enzymes

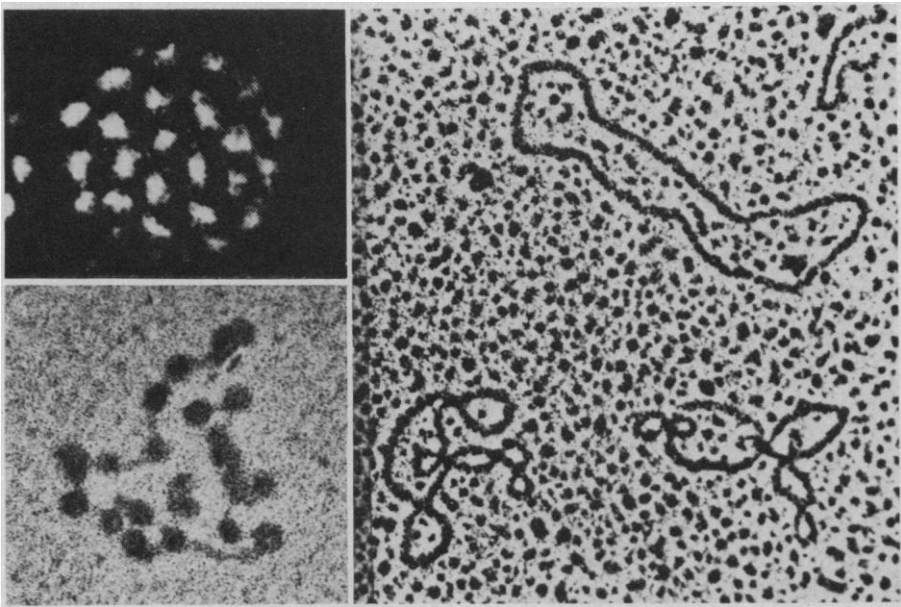
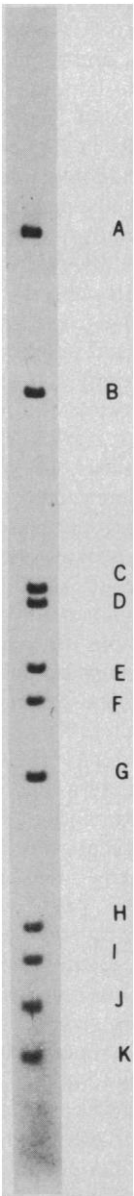


Fig. 1. (Upper left) Electron micrograph of an SV40 virus particle (75); lower left) an SV40 chromosome with typical nucleosomes (10); (right) free SV40 DNA. The DNA micrograph shows two molecules of form I DNA (covalently closed, circular duplex DNA) and one molecule of form II DNA (nicked, relaxed circular duplex DNA) (76).

Table 1. Order of *Hin* fragments: analysis of partial digestion products and *Hpa* fragments. Partial *Hin* digest fragments or *Hpa* fragments were recovered from electrophoresis gels and were redigested to completion with *Hind* II plus III. The redigestion products were identified by electrophoretic mobility (23).

| Initial fragment (% of SV40 DNA) | <i>Hin</i> digest products | Overlapping fragment order | | | | | | | | | | |
|-------------------------------------|-------------------------------|----------------------------|---|---|---|---|---|---|---|---|---|---|
| 12 | G,J | | J | G | | | | | | | | |
| 12 | F,K | F | | | | | | | | | | K |
| 13 | E,K | | | | | | | | | E | | K |
| 22 | B,G | | | G | B | | | | | | | |
| 40 | B,F,G,J,K | F | J | G | B | | | | | | | K |
| 43 | B,F,G,H,I,J | F | J | G | B | I | H | | | | | |
| 51 | A,C,D,E | | | | | | | A | C | D | E | |
| 20 (<i>Hpa</i> -C) | B,I | | | B | I | | | | | | | |
| 37 (<i>Hpa</i> -B) | A,H,C | | | | | H | A | C | | | | |
| 40 (<i>Hpa</i> -A) | D,E,F,G,J,K | F | J | G | | | | | | D | E | K |
| | Order: | F | J | G | B | I | H | A | C | D | E | K |

Fig. 2. Autoradiogram of ³²P-labeled SV40 DNA after digestion with the *Hind* enzyme of Smith and electrophoresis from top to bottom in 4 percent polyacrylamide gel (21). The largest fragment is near the top (A), and the smallest is near the bottom of the gel (K).



(Table 1). Our initial cleavage map was based on sites of cleavage by *Eco* RI (25), *Hind* II plus *Hind* III, and *Hpa* I plus *Hpa* II (26) (Fig. 3a). The single *Eco* RI site (18, 27) was designated the zero coordinate, and map units were expressed as fractional genome length from that site in an arbitrary direction around the cleavage map. With this map as a reference, sites of cleavage of SV40 DNA by other restriction enzymes have been localized in a number of laboratories, yielding the detailed map shown in Fig. 3b (28). As is seen in Fig. 3b, the circular SV40 genome can be opened at any one of several different sites by single-cut enzymes, and small or large fragments can be prepared from virtually any part of the molecule.

Nucleotide Sequence Map

The ultimate chemical map of a DNA molecule is its nucleotide sequence. The availability of small specific fragments and corresponding cleavage maps made such analysis feasible. Soon after our isolation of restriction fragments of SV40 DNA, two groups became interested in carrying out nucleotide sequence analysis of the fragments—S. M. Weissman's laboratory in New Haven and W. Fiers' in Ghent. We were pleased to cooperate in the initial phase of their important work. At the outset, RNA transcripts of isolated fragments were used for sequencing. Later, with the development of rapid DNA sequencing methods by Sanger and Coulson (29) and by Maxam

and Gilbert (30), direct sequencing of DNA fragments completed the analysis much sooner than originally expected, to the benefit of all investigators in this field. I am grateful to Sherman Weissman and Walter Fiers for making their sequence data freely available as the work progressed. The nucleotide sequence map of Weissman's group (31), and of Fiers *et al.* (32), consisting of 5226 nucleotide pairs, provides exact positions for each of the restriction sites in the cleavage map, and allows precise localization of genes and signals in the SV40 genome, as illustrated below.

Functional Map of SV40

The cleavage map and later the sequence map of the SV40 genome served as a framework for identifying functional elements of the viral DNA, for example, the origin and terminus of DNA replication, templates for viral messenger RNA's, and the positions of structural genes.

The origin and terminus of SV40 DNA replication were localized with respect to restriction sites in the DNA by pulse-labeling experiments (33) analogous to those of Dintzis on the rate and direction of globin biosynthesis (34), and by electron microscopic analysis of replicating SV40 DNA (35). The labeling experiments were carried out by exposing SV40-infected cells to ^3H -labeled thymidine for a period approximating that required for one round of viral DNA replication. DNA molecules whose replica-

tion was completed during this interval were isolated and digested with a restriction endonuclease, and the amount of radioactivity in each restriction fragment was determined. If there is a unique replication origin and terminus, fragments derived from the segment of the molecule synthesized last will be most highly labeled, and fragments derived from that segment synthesized first will have the least radioactivity. From the results (Fig. 4) we could infer that SV40 DNA replication does begin at a unique site, approximately at map coordinate 0.67, proceeds bidirectionally around the circular genome, and terminates about 180° from the origin at about map coordinate 0.17. Similar experiments with SV40 deletion mutants indicate that, whereas the origin is at a fixed position and therefore must be determined by a structural feature of the DNA, the termination point is not fixed but appears to represent the junction of the two growing forks opposite the origin (36).

Viral messenger RNA's (mRNA) were mapped in collaboration with George Khoury and Malcolm Martin (37) and by Sambrook *et al.* (38) by hybridization to restriction fragments of SV40 DNA. In summary, viral mRNA present in infected cells prior to the onset of viral DNA replication ("early" RNA) was derived from about half the genome (between map coordinates 0.17 and 0.67) by counterclockwise transcription. "Late" mRNA, that is, the RNA that appears after the onset of viral DNA replication, was derived from the other half of the genome by clockwise transcription (Fig.

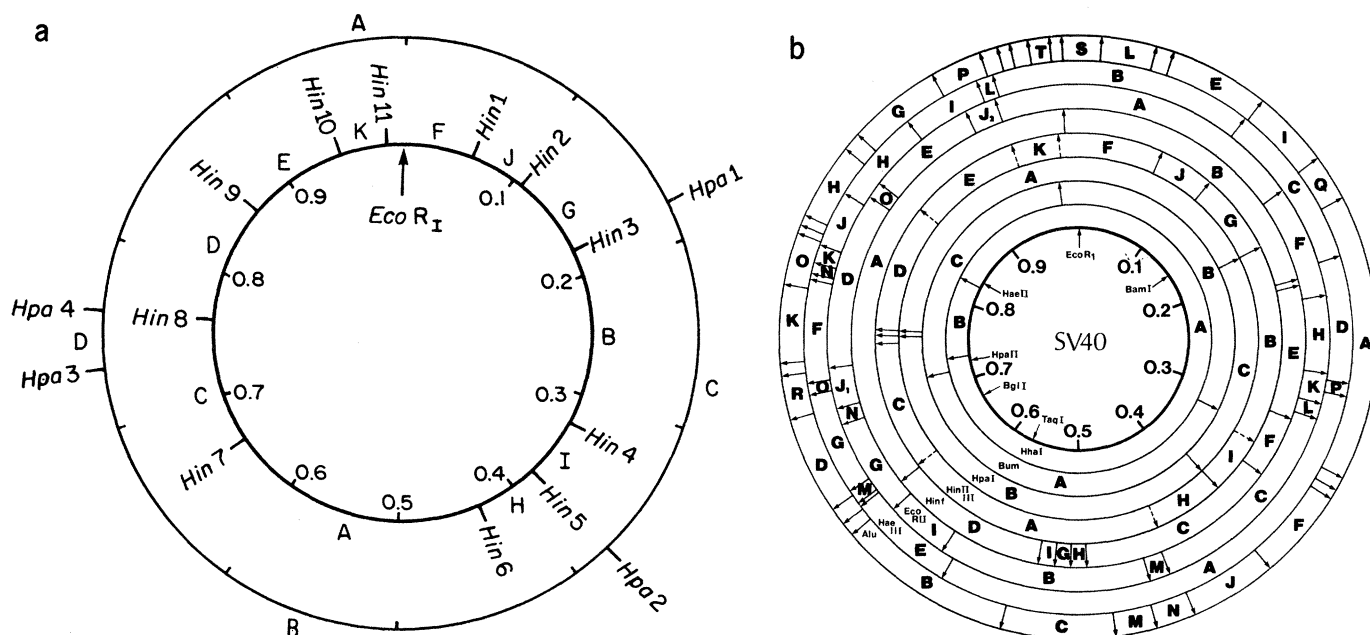


Fig. 3. Cleavage maps of SV40 DNA. (a) The initial cleavage map (23); (b) a more recent map. [See (28) and (63) for references to positioning of cleavage sites.]

5a). Viral mRNA's from transformed cells were derived from the early genomic segments plus adjacent late regions (39). A particularly intriguing finding was the position of the replication origin between the start of the early and late genome regions, suggesting a regulatory coupling between replication and transcription (Fig. 5a).

Recent major refinements in the analysis of mRNA (40, 41) have allowed much more precise mapping of individual SV40 messengers, down to the nucleotide sequence level in some cases, including the nucleotide positions of segments spliced out of initial gene transcripts (40). Some of these more recent findings are summarized in the map shown in Fig. 5b (41).

To locate *structural genes* of SV40 on the cleavage map, Ching-Juh Lai and I (42) and Mathei *et al.* (43) determined the mutational sites of temperature-sensitive (ts) mutants of SV40, isolated and characterized by Tegtmeyer and Ozer (44, 45) and by Chou and Martin (46), by an adaptation of the "marker rescue" procedure devised for coliphage ϕ X174 (47) (Fig. 6). In this method, a single-strand circle of mutant DNA is annealed with a single-strand restriction fragment derived from wild-type SV40 DNA to form a partial heteroduplex. Inside infected monkey cells the partial heteroduplex is repaired to form a duplex circle that has a mismatched base pair if the fragment overlaps the mutational site. By mismatch correction or replication a wild-type genome is generated, and is scored by its ability to grow under conditions where the mutant virus does not grow. By using a series of restriction fragments with each mutant to be mapped, we could determine which fragment overlapped a given mutational site. Since the position of each fragment in the cleavage map was known, we could localize any given ts mutation, and hence the genes in which the mutation resides, in the viral chromosome.

Lai's results are summarized in Fig. 5a. All of the tsA mutants, which are defective in initiation of viral DNA replication and in transformation (45, 46, 48, 49), mapped in the early region of the genome between coordinate 0.20 and 0.43; tsB, C, and BC mutants, which are defective in a viral structural protein (44, 46, 50), mapped between coordinates 0.94 and 0.17; and tsD mutants, which are defective in a second viral structural protein (51), mapped between 0.86 and 0.94 map units. Extensive segments of the genome were mutationally "silent." From the mapping of viral mRNA's and identification of their *in vitro* translation products (52), and from an analysis of deletion mutants of SV40 (see below), it

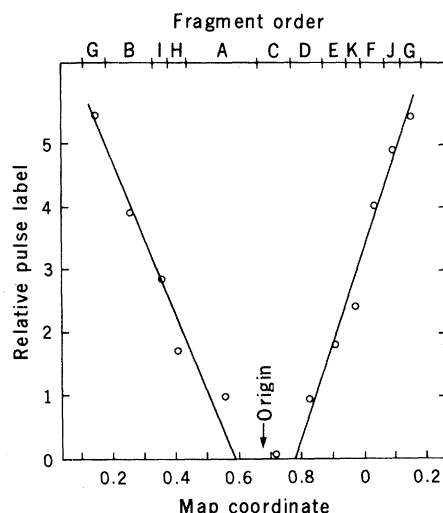


Fig. 4. Distribution of radioactivity in *Hin* fragments of pulse-labeled, newly completed molecules of SV40 DNA extracted from infected cells (33). The circular genome is shown in a linear form, with fragment G duplicated at the ends.

is now known that the A gene codes for SV40 tumor or T antigen (also known as the A protein), the B/C gene codes for the major virus structural protein (VPI), and the D gene codes for the overlapping minor virion proteins (VP2 and VP3). As is shown in Fig. 5b, nucleotide sequence data have subsequently allowed precise localization of each of these genes, including overlapping in-phase "late" sequences coding for the two minor structural proteins and overlapping in-phase "early" sequences coding for a second early protein, the so-called small t antigen (31, 32).

In vitro Construction of SV40 Mutants

The mutants just described were isolated by classical genetic techniques, namely, by random mutagenesis and selection of desired phenotypes. As is indicated in Fig. 5a, they covered only about half of the SV40 genome. With the advent of site-specific restriction endonucleases, it became possible to take a more active approach to mutational analysis of a DNA genome by creating mutations *in vitro* at preselected sites in the molecule. Site selection is based on restriction enzyme cleavage of one or both strands of the DNA, and mutations result from enzymatic or chemical modification at or near restriction sites. From DNA thus modified, individual mutants can be isolated without the need for phenotype selection. Given functional and chemical maps of the genome, interesting regions can be selected for perturbation to determine the effect of such changes on the function of genetic ele-

ments or gene products. In the case of SV40, to which these methods have been applied most extensively, a series of mutants has been generated with deletions or base substitutions at predetermined sites in the viral genome. These are proving useful in the identification and characterization of gene products and of regulatory signals in the DNA.

Constructed deletion mutants. In general, SV40 deletion mutants are constructed by enzymatic opening of the circular genome to form slightly shortened linear molecules, followed by transfection of cells with the linear DNA (53, 54). Some of the ways to form linear molecules missing a small segment of the genome are shown in Fig. 7. When used to transfect cultured monkey cells, such linear molecules form covalently closed circles within the cell that are missing nucleotide sequences at the joint (Fig. 7). The cyclization process itself (the enzymatic mechanism of which is not understood) leads to variable loss of nucleotides from the ends of the transfecting molecule, thus generating an array of "extended" deletion mutants (53). (To avoid the formation of extended deletions, linear molecules can be cyclized enzymatically *in vitro* prior to infection of cells.) If the overall loss of DNA does not remove a sequence essential for virus reproduction, deletion mutants can be isolated simply by selecting individual virus plaques arising in the infected cell monolayer (54). However, if the deletion of DNA leads to unconditional loss of function, the mutant must be isolated and propagated in the presence of a complementing "helper virus" (55). Once a mutant is cloned, that is, isolated in homogeneous form, the position and extent of the deletion in its DNA can be determined by restriction enzyme analysis and subsequent nucleotide sequence determination.

SV40 deletion mutants have been particularly useful in identifying structural gene products, such as the T antigens found in SV40-infected or SV40-transformed cells (56), in locating nonessential parts of the DNA (57), in more precise localization of the origin of replication (58), and in defining those regions required for cell transformation (59). Also important in localizing functions along the SV40 genome are (i) related experiments on transformation by restriction fragments of SV40 DNA (60) and on the activity of microinjected fragments (61) and (ii) the studies of adenovirus-SV40 hybrids containing SV40 DNA segments (62). As a result of these various investigations it became clear that the early region of the SV40 genome codes for the T antigens, as noted earlier, and that this

region (plus immediately adjacent sequences) is sufficient for viral DNA replication and for cell transformation.

SV40 mutants with base substitutions at preselected sites. Mutants of SV40 with single base pair changes at preselected restriction sites have been constructed by David Shortle, using local chemical mutagenesis, as illustrated in Fig. 8 (63). In this procedure, viral DNA is incised in one strand with a restriction enzyme, the "nick" is converted to a small gap with an exonuclease, and the bases exposed by the gap are then modified by reaction with a single-strand-specific mutagen, such as sodium bisulfite, which deaminates cytosine (C) to uracil (U). When such a bisulfite-treated gap is

repaired either in vitro or inside a cell, a U · A (A, adenine) pair is generated in place of the original C · G (G, guanine) pair. Thus, an entire base pair substitution occurs, and the mutation cannot be reversed by cellular enzymes. If (Fig. 8) there is only one site in the genome for the restriction enzyme used to make the initial scission, the mutagenized, repaired DNA can be exposed to the same enzyme to eliminate those molecules that have escaped mutagenesis within the restriction site. When such enzyme-resistant molecules are used to infect cell monolayers, the majority of resulting virus clones contain mutants that have lost the enzyme site (63). Recent extensions of the local mutagenesis procedure have

broadened the range of site selection considerably, so that many parts of a DNA molecule can be targeted for mutagenesis (64).

Constructed Regulatory Mutants of SV40

The local mutagenesis method just described has been used by Shortle and by Daniel DiMaio to construct mutants with single base pair substitutions within regulatory sequences of the viral DNA in and around the origin of replication (65). In one set of experiments, SV40 DNA was nicked with restriction endonuclease *Bgl* I, which cuts the viral DNA once within a long symmetrical sequence

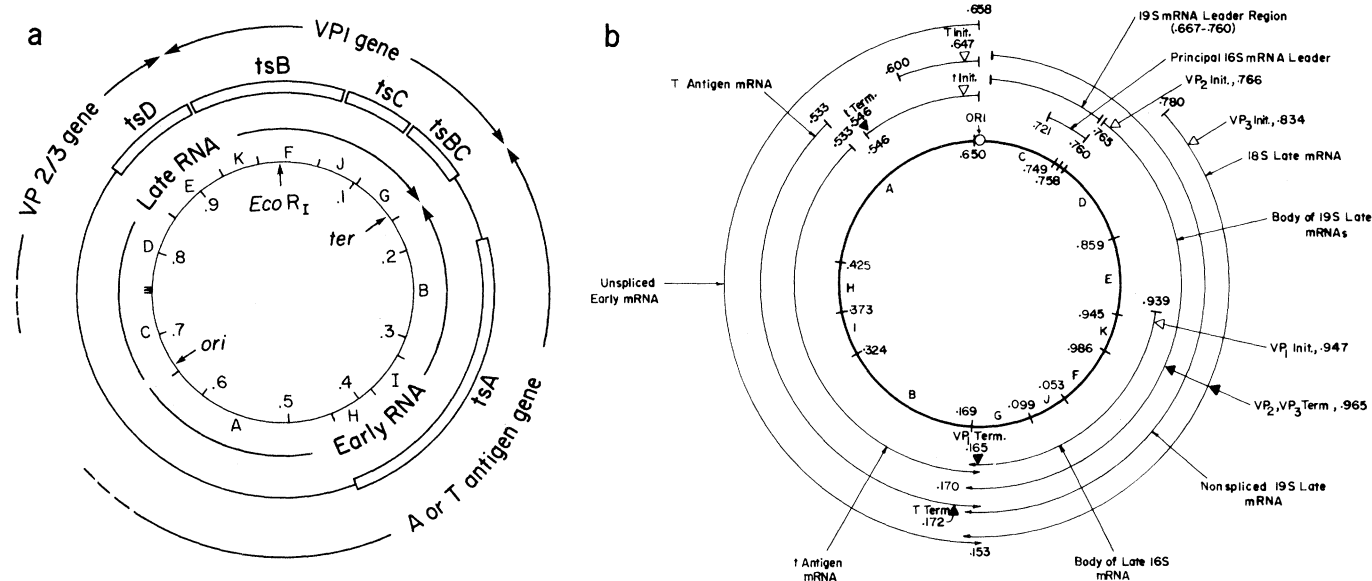


Fig. 5. Maps of viral functions relative to restriction sites: (a) an initial map localizing the origin (*ori*) and terminus (*ter*) of DNA replication; early and late mRNA, the direction of transcription, and mutational sites of *ts* mutants (77). *VP1*, major viral capsid protein and *VP2* and *VP3*, minor viral capsid proteins. (b) A recent map based on nucleotide sequence analysis of DNA and mRNA's (41). [The nucleotide position 1 of Reddy *et al.* (31) is near the origin of replication (here shown at the top of the circle)]. The nucleotide positions of segments coding for proteins are as follows (31, 32): *t* antigen, residues 5081 to 4559; *T* antigen, residues 5081 to 4837 and 4490 to 2612; *VP1*, residues 1423 to 2508; *VP2*, residues 480 to 1535; and *VP3*, residues 834 to 1535.

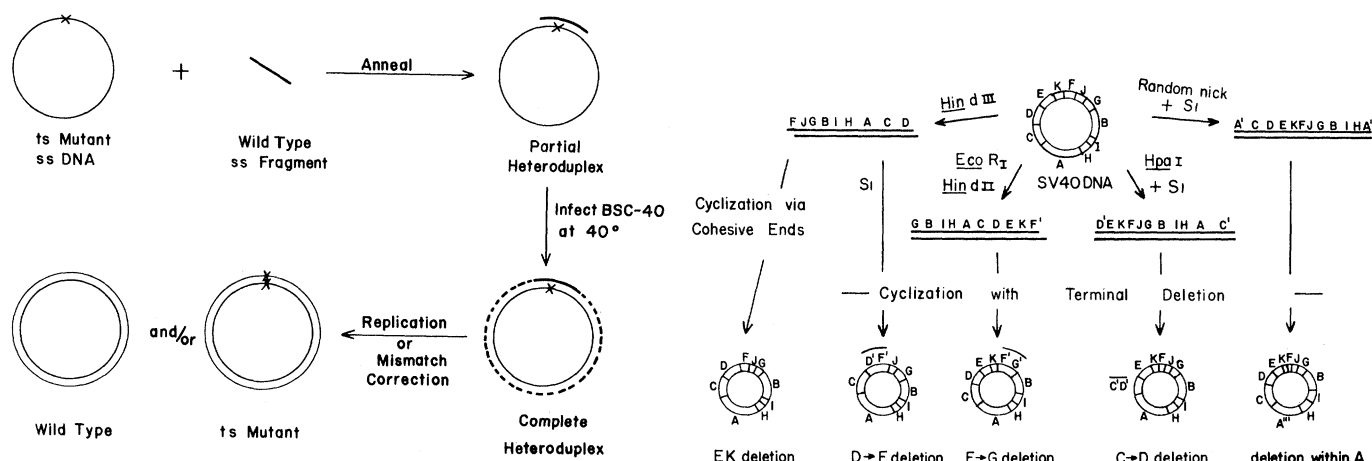


Fig. 6. (left) Mapping of SV40 mutants by marker rescue (42). See text for a description of each step. X, mutational site in the DNA. Fig. 7. (right) Generation of deletion mutants of SV40 by enzymatic excision of nucleotides from the DNA followed by cell-mediated cyclization with terminal deletion. Restriction enzymes are used to linearize the DNA at specific sites. Nonselective opening of the DNA (far right) can be used to produce a random set of deletions. *S1* refers to a single-strand specific nuclease (78), which nibbles the ends of linear DNA.

or "palindrome" at about map coordinate 0.67 (31), corresponding to the map position of the replication origin (Fig. 9). The *Bgl* I-nicked DNA was then gapped and locally mutagenized to generate *Bgl* I-resistant mutants. Nucleotide sequence analysis of several of these mutants revealed, in each case, a single base pair substitution within the palindrome (66) (Fig. 9). What is most interesting about these mutants is the effect of each base pair change on the rate of viral DNA replication (Table 2). A G · C to A · T change at position 5161, which forms the axis of symmetry of the palindrome, has no effect on the rate of DNA replication; a G · C to A · T change at position 5162 causes a marked decrease in the DNA replication rate; a C · G to

T · A change at position 5154 leads to a "cold-sensitive" replication phenotype (that is, reduced at low temperature); and a C · G to A · T change at position 5155 causes an increased rate of DNA replication. Appropriate tests indicate that the replication-defective mutants have abnormalities in a *cis* element controlling the rate of viral DNA replication. The mutational alterations therefore serve operationally to define the origin signal.

Our interpretation of the altered rates of mutant DNA replication is based on previous evidence for the involvement of the SV40 T antigen in initiation of viral DNA replication (45), and the preferential binding of this protein to a segment of SV40 DNA, including that shown in

Fig. 9 (67). The postulated first step in the replication of SV40 DNA is the specific binding of T antigen to the origin signal (Fig. 10). From the properties of origin mutants it appears that a single base pair change in the signal alters the binding site, leading to a change in the amount of T antigen bound or in the activity of the complex. In the cold-sensitive mutants, binding may be less efficient at 32°C than at higher temperatures either because of a temperature-dependent change in the binding site or as a result of a change in the secondary structure of the T antigen. Recently, second-site revertants of one of the replication-defective origin mutants have been isolated. The mutation responsible for the reversion maps in the gene for T antigen. Therefore, these revertants may be producing T antigens that recognize the mutant origin sequence more efficiently than does the wild-type T antigen (Fig. 10). Such double mutants could represent new viral replicons useful for biochemical investigations of T antigen functions.

A striking feature of SV40 origin mutants is that many are conditionally de-

Table 2. Replication origin mutants of SV40. Properties of mutants with base substitutions at the origin of viral DNA replication (66). Nucleotide positions are shown in Fig. 9.

| Base pair change | Plaque morphology | DNA replication |
|--------------------|-------------------|-------------------|
| 5161 G · C → A · T | Wild type | Normal |
| 5162 G · C → A · T | Small | Decreased |
| 5154 C · G → T · A | Small at 32°C | Decreased at 32°C |
| | Wild type at 40°C | Normal at 40°C |
| 5155 C · G → A · T | Small, sharp | Increased |

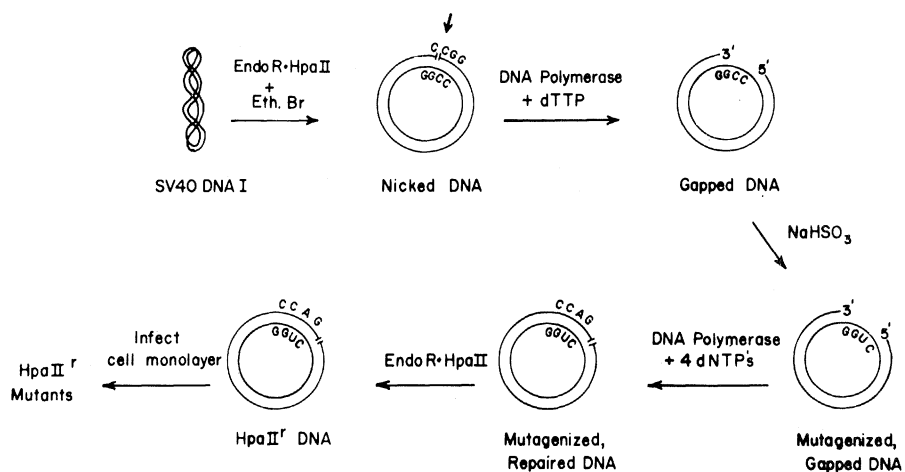


Fig. 8. Outline of the local mutagenesis procedure (63). See text for a description of each step. (Eth Br., ethidium bromide.)

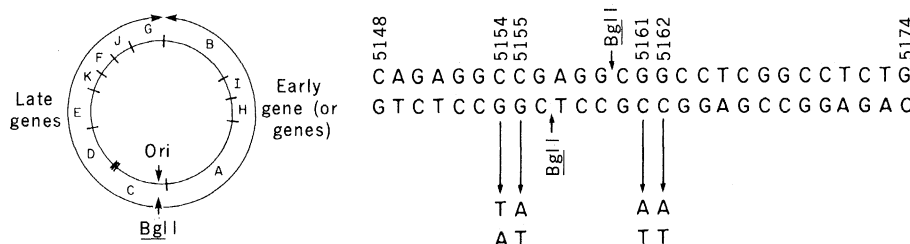


Fig. 9. Positions of base pair substitutions within the palindromic nucleotide sequence in DNA from *Bgl* I-resistant mutants of SV40 constructed by local mutagenesis, as described in the text. (left) The *Hin* cleavage map of the SV40 genome and the map position of the origin of DNA replication and of the *Bgl* I cleavage site. (right) The palindromic sequence at this map position, numbered as in (31). Nucleotide pair 5161 is the axis of symmetry of the palindrome, and the *Bgl* I cleavage sites are as indicated. The single base pair substitution in each of four phenotypically distinct mutants (at nucleotides 5154, 5155, 5161, or 5162) is indicated by an arrow.

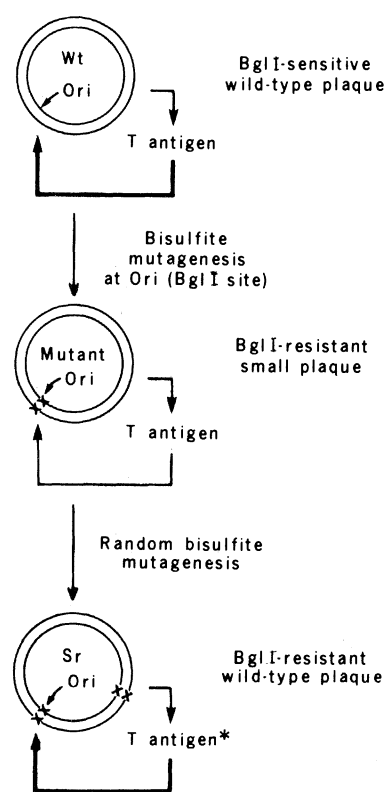


Fig. 10. Model of SV40 T antigen binding to the origin of replication in wild type (Wt), *ori* mutant, and second-site revertant (Sr) DNA. The thickness of the arrow from T antigen to *ori* reflects the hypothesized extent of binding. X, mutational site; T antigen*, mutant T antigen. On the right the phenotypes of wild-type, mutant, and second-type revertant viruses are noted.

fective. Especially frequent is the cold-sensitive phenotype. This kind of temperature dependence of specific DNA-protein interactions is well known from in vitro studies, as in the case of bacterial RNA polymerase binding to DNA promoter sequences (68), and may be a general property of regulatory protein-nucleic acid interactions that could be exploited to isolate mutants with sequence changes within many different controlling elements in DNA or RNA.

Analysis of More Complex Chromosomes

The methods used to dissect the tiny genome of SV40 are directly applicable to more complex DNA molecules that can be isolated in homogeneous form (69): large viral chromosomes, plasmids, or DNA from cellular organelles. Even certain genes in mammalian DNA, whose complexity is some one million times that of the SV40 genome, have been mapped by restriction enzyme cleavage, by means of the sensitive detection method devised by Southern (70). However, the completely general application of restriction enzymes to the analysis of cellular chromosomes depends on recombinant techniques for cloning and amplifying individual DNA fragments from complex mixtures (71), and on the ability to introduce active genes back into living cells (72). These advances have opened the genome of every organism to the type of chemical and functional analysis I have described for SV40. Interesting findings have already emerged; for example, the discontinuity of genes in eukaryotes (73) and the mobility of gene segments during development (74); and experiments are under way to identify regulatory elements in cloned cellular DNA. In time it should be possible to make out the basic regulatory mechanisms used by plant and animal cells, and eventually to understand some of the complex genetic programs that govern the growth, development, and specialized functions of higher organisms, including man.

References and Notes

1. S. E. Luria and S. L. Human, *J. Bacteriol.* **64**, 557 (1952).
2. G. Bertani and J. J. Weigle, *ibid.* **65**, 113 (1953).
3. W. Arber and D. Dussoix, *J. Mol. Biol.* **5**, 18 (1962); D. Dussoix and W. Arber, *ibid.*, p. 37; W. Arber, *Annu. Rev. Microbiol.* **19**, 365 (1965).
4. M. Meselson and R. Yuan, *Nature (London)* **217**, 1110 (1968).
5. H. O. Smith and K. Wilcox, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **28**, 465 (1969).
6. Y. Shimura, H. Kaizer, D. Nathans, *J. Mol. Biol.* **38**, 453 (1968).
7. B. H. Sweet and M. R. Hilleman, *Proc. Soc. Exp. Biol. Med.* **105**, 420 (1960).
8. L. V. Crawford and P. H. Black, *Virology* **24**, 388 (1964).
9. J. E. Germond, B. Hirt, P. Oudet, M. Gross-Bellard, P. Chambon, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1843 (1975).
10. J. D. Griffith, *Science* **187**, 1202 (1975).
11. J. Minowada and G. E. Moore, *Exp. Cell Res.* **29**, 31 (1963).
12. H. M. Shein and J. F. Enders, *Proc. Natl. Acad. Sci. U.S.A.* **48**, 1164 (1962).
13. S. Linn and W. Arber, *ibid.* **59**, 1300 (1968).
14. A. Haberman, J. Heywood, M. Meselson, *ibid.* **69**, 3138 (1972).
15. S. P. Adler and D. Nathans, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **29**, 725 (1970).
16. The restriction enzyme nomenclature [H. O. Smith and D. Nathans, *J. Mol. Biol.* **81**, 419 (1973)] is based on a three-letter abbreviation of the name of the host organism followed by a strain designation and enzyme number where required (for example, *Eco* B for *E. coli* strain B; *Hind* III for *Hemophilus influenzae* strain d, enzyme III).
17. S. P. Adler and D. Nathans, *Biochim. Biophys. Acta* **299**, 177 (1973); K. Horiuchi and N. D. Zinder, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3220 (1972).
18. J. F. Morrow and P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3365 (1972).
19. H. O. Smith and K. W. Wilcox, *J. Mol. Biol.* **51**, 379 (1970).
20. T. J. Kelly, Jr., and H. O. Smith, *ibid.*, p. 393.
21. K. J. Danna and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2913 (1971).
22. P. H. Roy and H. O. Smith, *J. Mol. Biol.* **81**, 445 (1973).
23. K. J. Danna, G. H. Sack, Jr., D. Nathans, *ibid.* **78**, 363 (1973).
24. R. J. Roberts, *Gene* **4**, 183 (1978).
25. J. Hedgpeth, H. M. Goodman, H. Boyer, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3448 (1972).
26. R. Gromkova and S. H. Goodgal, *J. Bacteriol.* **109**, 987 (1972).
27. C. Mulder and H. Delius, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3215 (1972).
28. T. J. Kelly, Jr., and D. Nathans, *Adv. Virus Res.* **21**, 85 (1977).
29. F. Sanger and A. R. Coulson, *J. Mol. Biol.* **94**, 441 (1975); F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
30. A. Maxam and W. Gilbert, *ibid.*, p. 560.
31. V. B. Reddy, B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, S. M. Weissman, *Science* **200**, 494 (1978).
32. W. Fiers, R. Contreras, G. Haegeman, R. Rogiers, A. Van de Voorde, H. Van Heuverswyn, J. Van Herreweghe, G. Volckert, M. Ysebaert, *Nature (London)* **273**, 113 (1978).
33. K. J. Danna and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3097 (1972).
34. H. M. Dintzis, *ibid.* **47**, 247 (1961).
35. G. C. Fareed, C. F. Garon, N. P. Salzman, *J. Virol.* **10**, 484 (1972).
36. C. J. Lai and D. Nathans, *J. Mol. Biol.* **97**, 113 (1975).
37. G. Khoury, M. A. Martin, T. N. H. Lee, K. J. Danna, D. Nathans, *ibid.* **78**, 377 (1973); G. Khoury, M. A. Martin, T. N. H. Lee, D. Nathans, *Virology* **63**, 263 (1975).
38. J. Sambrook, B. Sugden, W. Keller, P. A. Sharp, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3711 (1973).
39. G. Khoury, P. Howley, D. Nathans, M. A. Martin, *J. Virol.* **15**, 433 (1975).
40. A. J. Berk and P. A. Sharp, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1274 (1978).
41. P. K. Ghosh, V. B. Reddy, J. Swinscoe, P. Lebowitz, S. M. Weissman, *J. Mol. Biol.* **126**, 813 (1978).
42. C. J. Lai and D. Nathans, *Virology* **60**, 466 (1974); *ibid.* **66**, 78 (1975).
43. N. Mathei, H. W. Boyer, H. M. Goodman, *J. Virol.* **16**, 754 (1975).
44. P. Tegtmeyer and H. J. Ozer, *ibid.* **8**, 516 (1971).
45. P. Tegtmeyer, *ibid.* **10**, 591 (1972).
46. J. Y. Chou and R. G. Martin, *ibid.* **13**, 1101 (1974).
47. C. A. Hutchison III and M. H. Edgell, *ibid.* **8**, 181 (1971); P. J. Weisbeck and J. H. Van de Pol, *Biochim. Biophys. Acta* **224**, 328 (1970).
48. G. Kimura and R. Dulbecco, *Virology* **52**, 529 (1973).
49. ———, *ibid.* **49**, 394 (1972).
50. R. M. Martin and J. Chou, *J. Virol.* **15**, 599 (1975).
51. J. A. Robb and R. G. Martin, *ibid.* **9**, 956 (1972).
52. C. L. Prives, H. Aviv, E. Gilboa, M. Revel, E. Winocour, *Cold Spring Harbor Symp. Quant. Biol.* **39**, 309 (1974); C. L. Prives, E. Gilboa, M. Revel, E. Winocour, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 457 (1977); J. F. Greenblatt, B. Allet, R. Weil, C. Ahmad-Zadeh, *J. Mol. Biol.* **108**, 361 (1976); E. Paucha, A. Mellor, R. Harvey, A. E. Smith, R. W. Hewick, M. D. Waterfield, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2165 (1978).
53. C. J. Lai and D. Nathans, *J. Mol. Biol.* **89**, 179 (1974).
54. J. Carbon, T. Shenk, P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1392 (1975).
55. W. W. Brockman and D. Nathans, *ibid.* **71**, 942 (1974); J. E. Mertz and P. Berg, *Virology* **62**, 112 (1974).
56. K. Rundell, P. Tegtmeyer, H. L. Ozer, C. J. Lai, D. Nathans, *J. Virol.* **21**, 636 (1977); L. V. Crawford, C. N. Cole, A. E. Smith, E. Paucha, P. Tegtmeyer, K. Rundell, P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 117 (1978).
57. T. E. Shenk, J. Carbon, P. Berg, *J. Virol.* **18**, 644 (1976).
58. K. N. Subramanian and T. Shenk, *Nucleic Acids Res.* **5**, 3635 (1978); D. DiMaio and D. Nathans, in preparation.
59. W. A. Scott, W. W. Brockman, D. Nathans, *Virology* **75**, 319 (1976); M. J. Sleight, W. C. Topp, R. Hanich, J. F. Sambrook, *Cell* **14**, 79 (1978); N. Bouck, N. Beales, T. Shenk, P. Berg, G. DiMayora, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2473 (1978).
60. F. L. Graham, P. J. Abrahams, C. Mulder, H. L. Heijneker, S. O. Warnaar, F. A. J. de Vries, W. Fiers, A. J. van der Eb, *Cold Spring Harbor Symp. Quant. Biol.* **39**, 637 (1974).
61. C. Mueller, A. Graessmann, M. Graessmann, *Cell* **15**, 579 (1978).
62. A. M. Lewis, Jr., J. H. Breeden, Y. L. Werwka, L. E. Schnipper, A. S. Levine, *Cold Spring Harbor Symp. Quant. Biol.* **39**, 651 (1974).
63. D. Shortle and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2170 (1978).
64. D. Shortle, J. Pipas, S. Lazarowitz, D. DiMaio, D. Nathans, in *Genetic Engineering*, J. K. Setlow and A. Hollander, Eds. (Plenum, New York, 1979).
65. D. Shortle and D. Nathans, *Cold Spring Harbor Symp. Quant. Biol.*, in press.
66. ———, *J. Mol. Biol.*, in press.
67. R. Tjian, *Cell* **13**, 165 (1978).
68. M. J. Chamberlin, in *RNA Polymerase*, R. Losick and M. Chamberlin, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1976).
69. D. Nathans and H. O. Smith, *Annu. Rev. Biochem.* **44**, 273 (1975).
70. E. Southern, *J. Mol. Biol.* **98**, 503 (1975).
71. D. A. Jackson, R. H. Symon, P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2904 (1972); P. E. Lobban and A. D. Kaiser, *J. Mol. Biol.* **78**, 453 (1973); S. N. Cohen, A. C. Y. Chang, H. W. Boyer, R. B. Helling, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3240 (1973); J. F. Morrow, S. N. Cohen, A. C. Y. Chang, H. W. Boyer, H. M. Goodman, R. B. Helling, *ibid.* **71**, 1743 (1974); P. C. Wensink, D. J. Finnegan, J. E. Donelson, D. S. Hogness, *Cell* **3**, 315 (1974).
72. D. D. Brown and J. B. Gurdon, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2064 (1977); K. Struhl and R. W. Davis, *ibid.*, p. 5255; D. H. Hamer, D. Davoli, C. A. Thomas, Jr., G. C. Fareed, *J. Mol. Biol.* **112**, 155 (1977); M. Wigler, A. Pellicer, S. Silverstein, R. Axel, *Cell* **14**, 725 (1978).
73. C. Brack and S. Tonegawa, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5652 (1977); D. M. Glover and D. S. Hogness, *Cell* **10**, 167 (1977); S. M. Tilghman, D. C. Tiemeier, J. G. Seidman, B. M. Peterlin, M. Sullivan, J. V. Maizel, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 725 (1978); P. K. Wellauer and I. B. David, *Cell* **10**, 193 (1977).
74. N. Hozumi and S. Tonegawa, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3628 (1976).
75. F. A. Anderer, H. D. Schlumberger, M. A. Koch, H. Frank, H. J. Eggers, *Virology* **32**, 511 (1967).
76. J. Vinograd and J. Lebowitz, *J. Gen. Physiol.* **49**, 103 (1966).
77. D. Nathans, *Harvey Lect.* **70**, 111 (1976).
78. T. Ando, *Biochim. Biophys. Acta* **114**, 158 (1966).
79. It is a pleasure to acknowledge the contributions of my associates who carried out the research described in this article and shared its excitement with me: Stuart P. Adler, Kathleen J. Danna, Theresa N. H. Lee, George H. Sack, Jr., William W. Brockman, Elena O. Nightingale, Ching-Juh Lai, Walter A. Scott, Nicholas Muzyczka, Mary W. Gutai, David R. Shortle, James Pipas, Sondra Lazarowitz, and Daniel DiMaio. I am grateful also to my colleagues, Hamilton O. Smith and Thomas J. Kelly, Jr., for their sound critique and generous help over a period of many years; to George Khoury, Malcolm A. Martin, and Peter Tegtmeyer for enjoyable collaborations; and to the U.S. National Cancer Institute, the American Cancer Society, and the Whitehall Foundation for research support.