Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of Escherichia coli

(molecular hybrids/DNA joining/viral transformation/genetic transfer)

DAVID A. JACKSON*, ROBERT H. SYMONS†, AND PAUL BERG

Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305

Contributed by Paul Berg, July 31, 1972

ABSTRACT We have developed methods for covalently joining duplex DNA molecules to one another and have used these techniques to construct circular dimers of SV40 DNA and to insert a DNA segment containing lambda phage genes and the galactose operon of E. coli into SV40 DNA. The method involves: (a) converting circular SV40 DNA to a linear form, (b) adding single-stranded homodeoxypolymeric extensions of defined composition and length to the 3' ends of one of the DNA strands with the enzyme terminal deoxynucleotidyl transferase (c) adding complementary homodeoxypolymeric extensions to the other DNA strand, (d) annealing the two DNA molecules to form a circular duplex structure, and (e) filling the gaps and sealing nicks in this structure with E. coli DNA polymerase and DNA ligase to form a covalently closed-circular DNA molecule.

Our goal is to develop a method by which new, functionally defined segments of genetic information can be introduced into mammalian cells. It is known that the DNA of the transforming virus SV40 can enter into a stable, heritable, and presumably covalent association with the genomes of various mammalian cells (1, 2). Since purified SV40 DNA can also transform cells (although with reduced efficiency), it seemed possible that SV40 DNA molecules, into which a segment of functionally defined, nonviral DNA had been covalently integrated, could serve as vectors to transport and stabilize these nonviral DNA sequences in the cell genome. Accordingly, we have developed biochemical techniques that are generally applicable for joining covalently any two DNA molecules.† Using these techniques, we have constructed circular dimers of SV40 DNA; moreover, a DNA segment containing λ phage genes and the galactose operon of E. coli has been covalently integrated into the circular SV40 DNA molecule. Such hybrid DNA molecules and others like them can be tested for their capacity to transduce foreign DNA sequences into mammalian cells, and can be used to determine whether these new nonviral genes can be expressed in a novel environment.

MATERIALS AND METHODS

DNA. (a) Covalently closed-circular duplex SV40 DNA [SV40(I)] labeled with [3H]dTMP, 5 x 10^6 cpm/μg, free from SV40 linear or oligomeric molecules [but containing 3–5% of nicked double-stranded circles—SV40(II)] was purified from SV40-infected CV-1 cells (Jackson, D., & Berg, P., in preparation). (b) Closed-circular duplex λgal DNA labeled with [3H]dT (2.5 x 10^6 cpm/μg), was isolated from an E. coli strain containing this DNA as an autonomously replicating plasmid (see ref. 3) by equilibrium sedimentation in CsCl–ethidium bromide gradients (4) after lysis of the cells with detergent. A more detailed characterization of this DNA will be published later. Present information indicates that the λgal (λd–120) DNA is a circular dimer containing tandem duplications of a sequence of several λ phage genes (including CI, O, and P) joined to the entire galactose operon of E. coli (Berg, D., Mertz, J., & Jackson, D., in preparation). DNA concentrations are given as molecular concentrations.

Enzymes. The circular SV40 and λgal DNA molecules were cleaved with the bacterial restriction endonuclease RI (Yoshimori and Boyer, unpublished; the enzyme was generously made available to us by these workers). Phage λ-exonuclease (given to us by Peter Lobban) was prepared according to Little et al. (5), calf-thymus deoxynucleotidyl terminal transferase (terminal transferase), prepared according to Kato et al. (6), was generously sent to us by F. N. Hayes; E. coli DNA polymerase I Fraction VII (7) was a gift of Douglas Brutlag; and E. coli DNA ligase (8) and exonuclease III (9) were kindly supplied by Paul Modrich.

Substrates. [α-32P]deoxynucleoside triphosphates (specific activities 5–10 Ci/μmol) were synthesized by the method of Symons (10). All other reagents were obtained from commercial sources.

Centrifugations. Alkaline sucrose gradients were formed by diffusion from equal volumes of 5, 10, 15, and 20% sucrose solutions with 2 mM EDTA containing, respectively, 0.2, 0.4, 0.6, and 0.8 M NaOH, and 0.8, 0.6, 0.4, 0.2 M NaCl. 100-μl samples were run on 3.8-ml gradients in a Beckman SW56 Ti rotor in a Beckman L2-65B ultracentrifuge at 4° and 55,000 rpm for the indicated times. 2- to 10-drop fractions were collected onto 2.5-cm diameter Whatman 3MM discs, dried without washing, and counted in PPO–dimethyl POPOP–toluene scintillator in a Nuclear Chicago Mark II
Insertion of the Galactose Operon into SV40 DNA

**Formation of Hydrogen-Bonded Circular DNA Molecules.** [3H]dA and -dT DNAs were mixed at concentrations of 0.15 nM each in 0.1 M NaCl-10 mM Tris-HCl (pH 8.1)-1 mM EDTA. The mixture was kept at 51° for 30 min, then cooled slowly to room temperature.

**RESULTS**

**General approach**

Fig. 1 outlines the general approach used to generate circular, covalently-closed DNA molecules from two separate DNAs. Since, in the present case, the units to be joined are themselves circular, the first step requires conversion of the circular structures to linear duplexes. This could be achieved by a double-strand scission at random locations (see Discussion) or, as we describe in this paper, at a unique site with R 1 restriction endonuclease. Relatively short (50-100 nucleotides) poly(dA) or poly(dT) extensions are added on the 3'-hydroxyl termini of the linear duplexes with terminal transferase; prior to

**Conversion of SV40(II) DNA to Unit Length Linear DNA [SV40(LR1)] with R1 Endonuclease.** [3H]SV40 DNA (18.7 nM) in 100 mM Tris-HCl buffer (pH 7.5)-10 mM MgCl₂-2 mM 2-mercaptoethanol was incubated for 30 min at 37° with an amount of R 1 previously determined to convert 1.5 times this amount of SV40 (I) to linear molecules [SV40(LR1)]; Na EDTA (30 mM) was added to stop the reaction, and the DNA was purified in 67% ethanol.

**Removal of 5'-Terminal Regions from SV40(LR1) with λ Exonuclease.** [3H]SV40(LR1) (15 nM) in 67 mM K-glycinate (pH 9.5), 4 mM MgCl₂, 0.1 mM EDTA was incubated at 0° with λ-exonuclease (20 μg/ml) to yield [3H]SV40(LR1exoxo) DNA. Release of [3H]dTMP was measured by chromatographing aliquots of the reaction on polyethyleneimine thin-layer sheets (Brinkmann) in 0.8 M NH₄HCO₃ and counting the dTMP spot and the origin (undegraded DNA).

**Addition of Homopolymeric Extensions to SV40(LR1exo) with Terminal Transferase.** [3H]SV40(LR1exo) (50 nM) in 100 mM K-cacodylate (pH 7.0), 8 mM MgCl₂, 2 mM 2-mercaptoethanol, 150 μg/ml of bovine serum albumin, [α-32P]dNTP (0.2 mM for dATP, 0.4 mM for dTTP) was incubated with terminal transferase (30-60 μg/ml) at 37°. Addition of [32P]dNMP residues to SV40 DNA was measured by spotting aliquots of the reaction mixture on DEAE-paper discs (Whatman DE-81), washing each disc by suction with 50 ml (each) of 0.3 M NH₄-formate (pH 7.8) and 0.25 M NH₄HCO₃ and then with 20 ml of ethanol. To determine the proportion of SV40 linear DNA molecules that had acquired at least one "functional" (dA)ₙ tail, we measured the amount of SV40 DNA (3H counts) that could be bound to a Whatman GF/C filter (2.4-cm diameter) to which 150 μg of polyuridylic acid had been fixed (13). 15-μl Aliquots of the reaction mixture were mixed with 5 ml of 0.70 M NaCl-0.07 M Na citrate (pH 7.0)-2% Sarkosyl, and filtered at room temperature through the poly(U) filters, at a flow rate of 3-5 ml/min. Each filter was washed by rapid suction with 50 ml of the same buffer at 0°, dried, and counted. Control experiments showed that 98-100% of [3H]oligo(dA)ₙS bound to the filters under these conditions. When the ratio of [32P]dNMP to [3H]DNA reached the value equivalent to the desired length of the extension, the reaction was stopped with EDTA (30 mM) and 2% Sarkosyl. The [3H]SV40(LR1exo)-[32P]dA or -dT DNA was purified by neutral sucrose gradient zone sedimentation to remove unincorporated dNTP, as well as any traces of SV40(I) or SV40(II) DNA.
removal of a short sequence (30–50 nucleotides) from the 5'-phosphoryl termini by digestion with λ exonuclease facilitates the terminal transferase reaction. Linear duplexes containing (dA)₆ extensions are annealed to the DNA to be joined containing (dT)₆ extensions at relatively low concentrations. The circular structure formed contains the two DNAs, held together by two hydrogen-bonded homopolymeric regions (Fig. 1). Repair of the four gaps is mediated by E. coli DNA polymerase with the four deoxynucleosidetriphosphates, and covalent closure of the ring structure is effected by E. coli DNA ligase; E. coli exonuclease III removes 3'-phosphoryl residues at any nicks inadvertently introduced during the manipulations (nicks with 3'-phosphoryl ends cannot be sealed by ligase).

**Principal steps in the procedure**

Circular SV40 DNA Can Be Opened to Linear Duplexes by R₁ Endonuclease. Digestion of SV40(I) DNA with excess R₁ endonuclease yields a product that sediments at 14.5 S in neutral sucrose gradients and appears as a linear duplex with the same contour length as SV40(II) DNA when examined by electron microscopy [(18); Jackson and Berg, in preparation; see Table 1]. The point of cleavage is at a unique site on the SV40 DNA, and few if any single-strand breaks are introduced elsewhere in the molecule (18); moreover, the termini at each end are 5'-phosphoryl, 3'-hydroxyl (Mertz, J., Davis, R., in preparation). Digestion of plaque-purified SV40 DNA under our conditions yields about 87% linear molecules, 10% nicked circles, and 3% residual supercoiled circles.

**Addition of Oligo(dA) or (dT) Extensions to the 3'-Hydroxyl Termini of SV40 (L₁R₁).** Terminal transferase has been used to generate deoxyhomopolymeric extensions on the 3'-hydroxyl termini of DNA (7); once the chain is initiated, chain propagation is statistical in that each chain grows at about the same rate (12). Although the length of the extensions can be controlled by variation of either the time of incubation or the amount of substrate, we have varied the time of incubation to minimize spurious nicking of the DNA by trace amounts of endonuclease activity in the enzyme preparation; we have so far been unable to remove or selectively inhibit these nuclease (Jackson and Berg, in preparation).

Incubation of SV40(L₁R₁) with terminal transferase and either dATP or dTTP resulted in appreciable addition of mononucleotidyl units to the DNA. But, for example, after addition of 100 residues of dA per end, only a small proportion of the modified SV40 DNA would bind to filter discs containing poly(U) (13). This result indicated that initiation of terminal nucleotidyl addition was infrequent with SV40(L₁R₁), but that once initiated those termini served as preferential primers for extensive homopolymer synthesis.

Lobban and Kaiser (unpublished) found that P22 phage DNA became a better primer for homopolymer synthesis after incubation of the DNA with λ exonuclease. This enzyme removes, successively, deoxymononucleotides from 5'-phosphoryl termini of double-stranded DNA (15), thereby rendering the 3'-hydroxyl termini single-stranded. We confirmed their finding with SV40(L₁R₁) DNA; after removal of 30–50
nucleotides per 5'-end (see Methods), the number of SV40(LR1) molecules that could be bound to poly(U) filters after incubation with terminal transferase and dATP increased 5- to 6-fold. Even after separation of the strands of the SV40(LRexo)-

The weight-average length of the homopolymer extensions was 50–100 residues per end. Zone sedimentation of [3H]-

Hydrogen-Bonded Circular Molecules Are Formed by An

Covelyently Closed-Circular DNA Molecules Are Formed by Incubation of Hydrogen-Bonded Complexes with DNA Poly

TABLE 1. Relative lengths of SV40 and λdegal-120

Input 32P label derived from the oligo(dA) and -(dT) tails sediments with the 3H label present in the SV40 DNA, in the position expected of a covalently closed-circular SV40 dimer (70–75 S). About the same amount of labeled DNA bands in a CsCl-ethidium bromide gradient at a buoyant density characteristic of covalently closed-circular DNA (Fig. 4).

DNA isolated from the heavy band of the CsCl-ethidium bromide gradient contains primarily circular molecules, with a contour length twice that of SV40(II) DNA (Table 1) when viewed by electron microscopy. No covalently closed DNA is formed if either one of the linear precursors is omitted from the annealing step or if the enzymes are left out of the closure reaction. We conclude, therefore, that two unit-length linear SV40 molecules have been joined to form a covalently closed-circular dimer.

Covelyently closure of the hydrogen-bonded SV40 DNA di

Preparation of the Galactose Operon for Insertion into SV40

DNA. The galactose operon of E. coli was obtained from a λdegal DNA; λdegal is a covalently closed, supercoiled DNA molecule four times as long as SV40(II) DNA (Table 1). After complete digestion of λdegal DNA with the R1 endonuclease, linear molecules two times the length of SV40(II) DNA are virtually the exclusive product (Table 1). This population has a unimodal length distribution by electron microscopy and appears to be homogeneous by ultracentrifugal criteria (Jackson and Berg, in preparation). The R1 endonuclease seems, therefore, to cut λdegal circular DNA into two equal length linear molecules. Since one R1 endonuclease cleavage per λd mono-

Fig. 4. CsCl-ethidium bromide equilibrium centrifugation of the products analyzed in Fig. 4. Line A, dA-ended, plus dT-

![Graph](A)
same sites and, therefore, that each linear piece contains an intact galactose operon.

The purified λdegal (LRI) DNA was prepared for joining to SV40 DNA by treatment with λ-exonuclease, followed by terminal transferase and [32P]dTTP, as described for SV40-(LRI).

Formation of Covalently Closed-Circular DNA Molecules Containing both SV40 and λdegal DNA. Annealing of [3H]-SV40(LRIexo)-[32P](dA)60 and [3H]λdegal-180(LRIexo)-[32P](dT)85 incubated for 3 hr with and without (P+L+III). Centrifugation was for 60 min. Line A, dA-ended SV40, plus dT-ended λdegal-180 linears, plus (P+L+III) [3H]; line B, dT-ended λdegal-180 linears, plus dT-ended SV40 linears, plus (P+L+III) [32P]; line C, dA-ended SV40 linears, plus dT-ended λdegal-180 linears, without (P+L+III) [32P].

The arrows indicate the position in the gradient of supercoiled marker DNAs having the indicated multiple of SV40 DNA molecular size.

Omission of the enzymes from the reaction mixture prevents λdegal-SV40 DNA formation (Figs. 5 and 6). No covalently closed product is detectable (Fig. 5) if λdegal and SV40 linear molecules with identical, rather than complementary, tails are annealed and incubated with the enzymes. This result demonstrates directly that the formation of covalently closed DNA depends on complementarity of the homopolymeric tails.

We conclude from the experiments described above that λdegal DNA containing the intact galactose operon from E. coli, together with some phage λ genes, has been covalently inserted into an SV40 genome. These molecules should be useful for testing whether these bacterial genes can be introduced into a mammalian cell genome and whether they can be expressed there.

DISCUSSION

The methods described in this report for the covalent joining of two SV40 molecules and for the insertion of a segment of DNA containing the galactose operon of E. coli into SV40 are general and offer an approach for covalently joining any two DNA molecules together. With the exception of the fortuitous property of the R1 endonuclease, which creates linear DNA precursors, none of the techniques used depends upon any unique property of SV40 and/or the λdegal DNA. By the use of known enzymes and only minor modifications of the methods described here, it should be possible to join DNA molecules even if they have the wrong combination of hydroxyl and phosphoryl groups at their termini. By judicious use of generally available enzymes, even DNA duplexes with protruding 5'- or 3'-ends can be modified to become suitable substrates for the joining reaction.

One important feature of this method, which is different from some other techniques that can be used to join unrelated DNA molecules to one another (16, 19), is that here the joining is directed by the homopolymeric tails on the DNA. In our protocol, molecule A and molecule B can only be joined to each other; all AA and BB intermolecular joinings and all A and B intramolecular joinings (circularizations) are prevented. The yield of the desired product is thus increased, and subsequent purification problems are greatly reduced.

![Fig. 5. Alkaline sucrose gradient sedimentation of annealed [3H]SV40(LRIexo)-[32P](dA)60 and [3H]λdegal-180(LRIexo)-[32P](dT)85 incubated for 3 hr with and without (P+L+III). Centrifugation was for 60 min. Line A, dA-ended SV40, plus dT-ended λdegal-180 linears, plus (P+L+III) [3H]; line B, dT-ended λdegal-180 linears, plus dT-ended SV40 linears, plus (P+L+III) [32P]; line C, dA-ended SV40 linears, plus dT-ended λdegal-180 linears, without (P+L+III) [32P]. The arrows indicate the position in the gradient of supercoiled marker DNAs having the indicated multiple of SV40 DNA molecular size.](image1)

![Fig. 6. CsCl-ethidium bromide equilibrium centrifugations of joined [3H]SV40(LRIexo)-[32P](dA)60 and [3H]λdegal-180(LRIexo)-[32P](dT)85 DNA. The samples were those referred to in Fig. 5. Line A, dA-ended SV40 linears, plus dT-ended λdegal-180 linears, plus (P+L+III) [32P]; line B, the same mixture without (P+L+III) [32P].](image2)
For some purposes, however, it may be desirable to insert λdegal or other DNA molecules at other specific, or even random, locations in the SV40 genome. Other specific placements could be accomplished if other endonucleases could be found that cleave the SV40 circular DNA specifically. Since pancreatic DNase in the presence of Mn²⁺ produces randomly located, double-strand scissions (17) of SV40 circular DNA (Jackson and Berg, in preparation), it should be possible to insert a DNA segment at a large number of positions in the SV40 genome.

Although the λdegal DNA segment is integrated at the same location in each SV40 DNA molecule, it should be emphasized that the orientation of the two DNA segments to each other is probably not identical. This follows from the fact that each of the two strands of a duplex can be joined to either of the two strands of the other duplex (e.g., \( W^W C^\overline{C} \) or \( C^C W^\overline{W} \)). What possible consequences this fact has on the genetic expression of these segments remains to be seen.

We have no information concerning the biological activities of the SV40 dimer or the λdegal–SV40 DNAs, but appropriate experiments are in progress. It is clear, however, that the location of the R1 break in the SV40 genome will be crucial in determining the biological potential of these molecules; preliminary evidence suggests that the break occurs in the late genes of SV40 (Morrow, Kelly, Berg, and Lewis, in preparation).

A further feature of these molecules that may bear on their usefulness is the (dA-dT)₄ tracts that join the two DNA segments. They could be helpful (as a physical or genetic marker) or a hindrance (by making the molecule more sensitive to degradation) for their potential use as a transducer.

The λdegal–SV40 DNA produced in these experiments is, in effect, a trivalent biological reagent. It contains the genetic information to code for most of the functions of SV40, all of the functions of the E. coli λ galactose operon, and those functions of the λ bacteriophage required for autonomous replication of circular DNA molecules in E. coli. Each of these sets of functions has a wide range of potential uses in studying the molecular biology of SV40 and the mammalian cells with which this virus interacts.

We are grateful to Peter Lobban for many helpful discussions. D. A. J. was a Basic Science Fellow of the National Cystic Fibrosis Research Foundation; R. H. S. was on study leave from the Department of Biochemistry, University of Adelaide, Australia and was supported in part by a grant from the USPHS. This research was supported by Grant GM-13235 from the USPHS and Grant VC-23A from the American Cancer Society.