antigens would provide some assurance against potentially undesirable contamination (*11*).

Numerous subspecies of African Green monkeys (*C. aethiops*), also termed grivets, vervets, or guenons, are found throughout most of sub-Saharan Africa. In limited serologic surveys conducted to date, other African nonhuman primates such as the chimpanzee (*Pan troglodytes*), baboon (*Papio* sp.), patas monkey (*E. patas*), and colobus monkey (*C. polykomos*) were seronegative to STLV-III viral proteins. It is possible that other nonhuman primates from specific geographic locales in Africa are infected with an STLV-III. In Africa, the African Green monkeys are somewhat gregarious and are regarded as agricultural pests. They have been considered as reservoirs or vectors of certain other viruses that sometimes cause disease in humans including Ebola fever, Marburg disease, and African yellow fever.

$STLV\text{-}III_{mac}$ was isolated from macaques with an immunodeficiency syndrome. Serologic studies of the diseased macaques indicate that the possession of STLV-III antibodies is closely associated with this syndrome; healthy macaques have been seronegative (*12*). This is in contrast to the situation with wild-caught healthy African Green monkeys that are frequently seropositive. In preliminary studies $STLV\text{-}III_{mac}$ inoculation induced seroconversion in six out of six inoculated rhesus macaques (*12*). This was associated with the development of an acute and fatal immunosuppressive disease in four of the six macaques (*13*).

To date, there has been no evidence of disease in any African Green monkeys that have evidence of exposure to $STLV\text{-}III_{AGM}$. Understanding the biology of an HTLV-III–related virus in this primate species may help us understand the specific viral alterations or viral-host interactions that are involved in the pathogenicity of this family of T-lymphotropic retroviruses and perhaps provide a new approach in the development of an AIDS vaccine.

References and Notes

1. I. Miyoshi *et al.*, *Lancet* **1982-II**, 658 (1982); G. Hunsmann *et al.*, *Int. J. Cancer* **32**, 329 (1983); W. Saxinger *et al.*, in *Human T-Cell Leukemia Viruses*, R. C. Gallo, M. Essex, L. Gross, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 1984), p. 323; M. Hayami *et al.*, *Int. J. Cancer* **33**, 179 (1984); H-G. Guo *et al.*, *Science* **223**, 1195 (1984).
2. T. Homma *et al.*, *Science* **225**, 716 (1984).
3. M. D. Daniel *et al.*, *ibid.* **228**, 1201 (1985).
4. P. J. Kanki *et al.*, *ibid.*, p. 1199.
5. J. S. Allan *et al.*, *ibid.* **230**, 810 (1985).
6. N. Clumeck *et al.*, *N. Engl. J. Med.* **310**, 492 (1984); P. Van de Perre *et al.*, *Lancet* **1984-II**, 62 (1984); P. Piot *et al.*, *ibid.*, p. 65 (1984).
7. P. J. Kanki *et al.*, *Lancet* **1985-I**, 1330 (1985).
8. A. F. Gazdar *et al.*, *Blood* **55**, 409 (1980); B. J. Poiesz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6815 (1980).
9. M. Popovic *et al.*, *Science* **224**, 497 (1984).
10. F. Barin *et al.*, *ibid.* **228**, 1094 (1985); J. Allan *et al.*, *ibid.* **228**, 1091 (1985).
11. Weekly Epidemiological Record, in press.
12. P. Kanki and M. Essex, in preparation.
13. N. Letvin *et al.*, *Science* **230**, 71 (1985).
14. J. K. Laemmli, *Nature (London)* **227**, 680 (1970).

15. We thank G. Dreesman, R. Kurth, J. Watson, M. Kessler and the Franklin Park Zoo for primate sera, R. C. Gallo for the reference cell cultures, M. G. Sarngadharan for monoclonal antibody to HTLV-III p24, and J. Groopman for human reference sera from patients with AIDS or ARC. This research was supported by NIH research grants CA37466 and CA-18216. P.K. was supported by National Research Service Award ST 32 CA9382.

4 September 1985; accepted 8 October 1985

Association of Crossover Points with Topoisomerase I Cleavage Sites: A Model for Nonhomologous Recombination

Abstract. *Nonhomologous DNA recombination is frequently observed in somatic cells upon the introduction of DNA into cells or in chromosomal events involving sequences already stably carried by the genome. In this report, the DNA sequences at the crossover points for excision of SV40 from chromosomes were shown to be associated with eukaryotic topoisomerase I cleavage sites in vitro. The precise location of the cleavage sites relative to the crossover points has suggested a general model for nonhomologous recombination mediated by topoisomerase I.*

PETER BULLOCK*
Department of Molecular Biology, University of California, Berkeley 94720, and
Department of Microbiology and Immunology, School of Medicine, University of Washington, Seattle 98195
JAMES J. CHAMPOUX
Department of Microbiology and Immunology, School of Medicine, University of Washington, Seattle
MICHAEL BOTCHAN
Department of Molecular Biology, University of California, Berkeley

*Present address: Imperial Cancer Research Fund, London, England.

Chromosomal excision of SV40 from nontandem inserts proceeds via imprecise, illegitimate (nonhomologous) recombination events. As a result, single integrated copies of SV40 yield heterogeneous populations of excision products. (*1*). Nevertheless, excisional recombination is not random because a given cell line gives rise to a reproducible population of discrete excised forms (*2*). SV40 integration is not specific with respect to either the viral or cellular sequences; thus, different cell lines generate different populations of excised forms because the crossover points used during the excisional recombination events involve unique combinations of viral and cellular DNA's. Our analysis of the parental DNA's used during SV40 excision from the proviral locus present within the rat cell line 14B (*1*) suggested that the nonrandom nature of these illegitimate recombination events was dependent on features either at, or very close to, the crossover points (*2*). The crossover point features we identified included small homologies [2 to 3 base pairs (bp) long] at the point of strand exchange and sequences that were very similar to those reported to be eukaryotic topoisomerase I (topo I) cleavage sites (*3*, *4*). The observation that potential topo I cleavage sites were associated with the excision crossover points was interesting in light of reports that this enzyme can catalyze the ligation in vitro of heterologous DNA fragments lacking any sequence complementarity (*4–6*). That eukaryotic topo I is available to catalyze illegitimate recombination in vivo is suggested by studies that demonstrated that this abundant enzyme is constitutively expressed in somatic cells (*7*) and is associated with chromatin (*8–10*).

Eukaryotic topo I cleavage sites have sequence features in common; the great majority of these sites contain the sequence CTT or GTT immediately 5′ to the cleavage site (*3*, *4*, *11*). However, the specificity is not absolute, as (for example) is the case with a restriction endonuclease. Also, although single-stranded DNA is infrequently cleaved by topo I relative to duplex DNA, it is nonetheless a substrate for this enzyme (*11*). Therefore, the possible association of topo I cleavage sites with SV40 excisional recombination crossover points had to be tested in vitro and with both single-stranded and duplex DNA's. This report indicates that the excisional recombination crossover points in duplex DNA, but not in single-stranded DNA, are associated with rat liver topo I cleavage sites in a statistically significant manner.

Eukaryotic topo I preferentially interacts with certain sequences on DNA, and these sequences can be identified by stopping the nicking-closing reaction with sodium dodecyl sulfate (SDS) (*12*). Under these conditions, nicks are gener-

ated that have 5′ termini ending in hydroxyl groups and 3′ termini containing phosphates to which topo I molecules are covalently bound (*13*, *14*). In the absence of SDS, the covalent attachment of topo I molecules to 3′ phosphates on the broken strand is believed to store the energy required for resynthesis of phosphodiester bonds during the nicking-closing reaction (*13*, *15*) or, possibly under special conditions, the resynthesis of a phosphodiester bond at a different location during recombination.

A physical map of the single viral insertion in cell line 14B, together with flanking chromosomal sequences, is presented in Fig. 1. By means of restriction endonuclease cleavage, four plasmids were constructed that contained sets of particular crossover points that are used during viral excision (F1, F5, F6, and F12). The relative positions of these crossover sites were previously determined (*2*). The plasmids were used to generate fragments (*16*), asymmetrically labeled at their 3′ termini, that contained only one crossover point. These fragments were used as substrates for the nicking reaction with rat liver topo I to determine if topo I specifically interacted with the DNA sequences involved in excisional recombination. Upon interruption of the nicking-closing reaction and subsequent heat denaturation to produce single strands, a 3′-labeled strand produces two populations of molecules. The first population contains the 3′-labeled end but different topo I–generated 5′ termini ending in hydroxyl groups. The second population contains unlabeled DNA fragments with topo I covalently bound to the 3′ phosphates.

The identity of the bonds cleaved by topo I can be established by sequence gel electrophoresis of the topo I cleavage products in lanes adjacent to lanes containing the same fragment cleaved by the base-specific sequencing reactions of Maxam and Gilbert (*17*, *18*). Moreover, to correctly identify the bonds cleaved by topo I in this procedure, it is necessary to take into account that bands generated by Maxam and Gilbert sequencing arise as a result of the degradation of individual nucleotides. Thus, bands in the sequencing lanes are one nucleotide shorter than corresponding

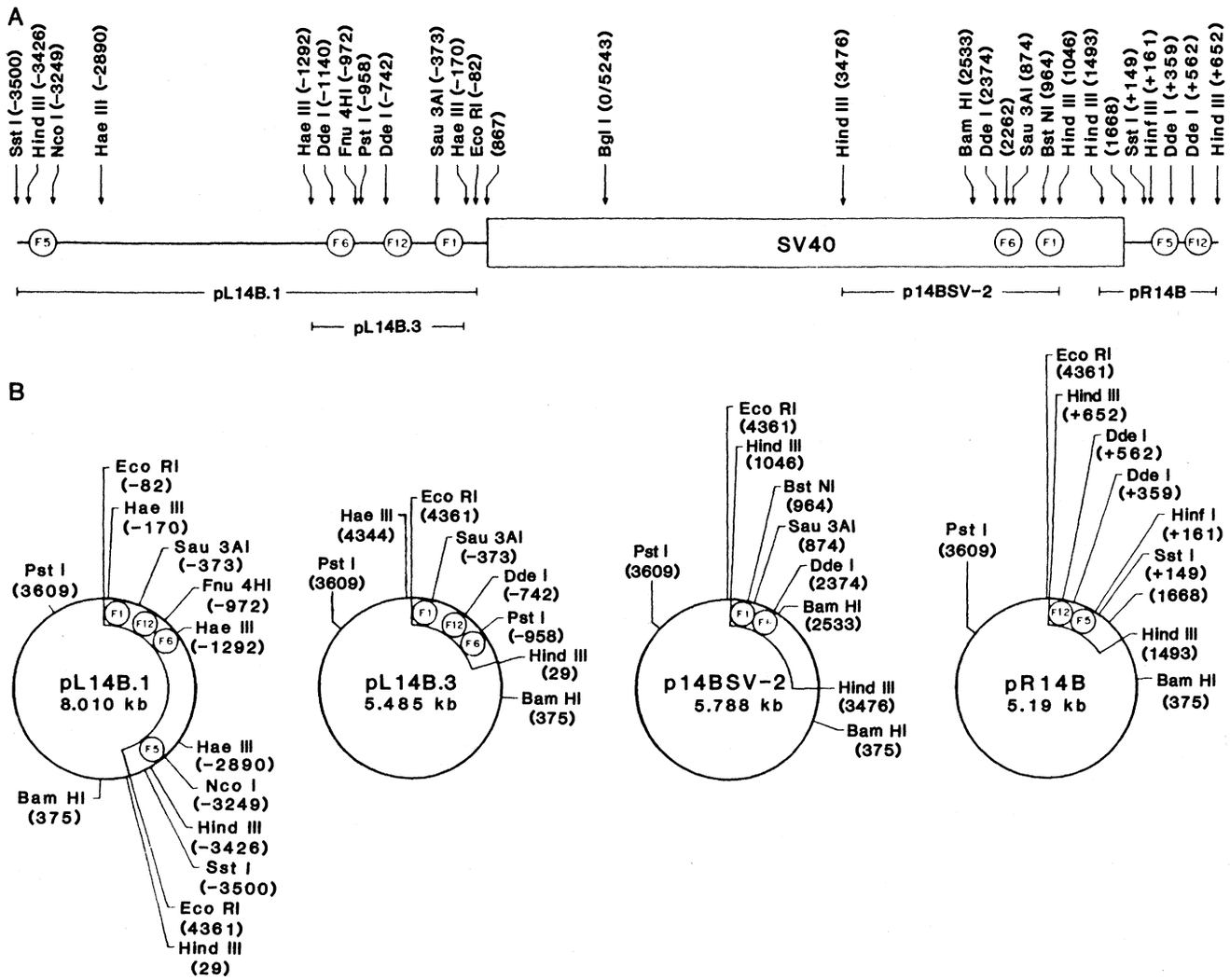

Fig. 1. (A) A physical map of the single SV40 DNA copy integrated in the 14B cell line. Genomic sequences to the left of the virus-cell junction at 867 are designated by negative numbers and those to the right of the virus-cell junction at 1668 are designated by positive numbers. The BB numbering system for the viral DNA sequences was used (*29*). The circles superimposed on the physical map designate pairs of sites at which nonhomologous recombination occurred during excision of the forms cloned in the Fusion (F) series (*1*, *2*). The boundaries of the 14B genomic sequences that were subcloned into pBR322 to create plasmids pL14B.1, pL14B.3, p14BSV-2, and pR14B are delineated by the lines below the map. (B) The orientation of the inserts within the four 14B subclones. Plasmid pBR322 sequences are numbered according to Sutcliffe (*30*). The steps taken to isolate the asymmetrically labeled fragments that contain a particular crossover point from these plasmids are as described (*16*). In all cases, 3′ ends of fragments were labeled with ^{32}P by repair synthesis with Klenow polymerase (*3*). Labeled fragments were separated on 3.5 percent acrylamide gels (*31*) and freed of acrylamide as described (*32*).

bands in lanes containing the topo I cleavage products. However, since fragments generated by these sequencing reactions terminate in 5′ phosphates, while topo I cleavage products terminate in 5′ hydroxyl groups, a difference in charge-to-mass ratios can distort the relative band mobilities of small (≤50 bp) fragments (*19*). Possible ambiguities in the assignment of the bonds cleaved by topo I were avoided by treating sequencing reaction products with bacterial alkaline phosphatase prior to electrophoresis (*20*).

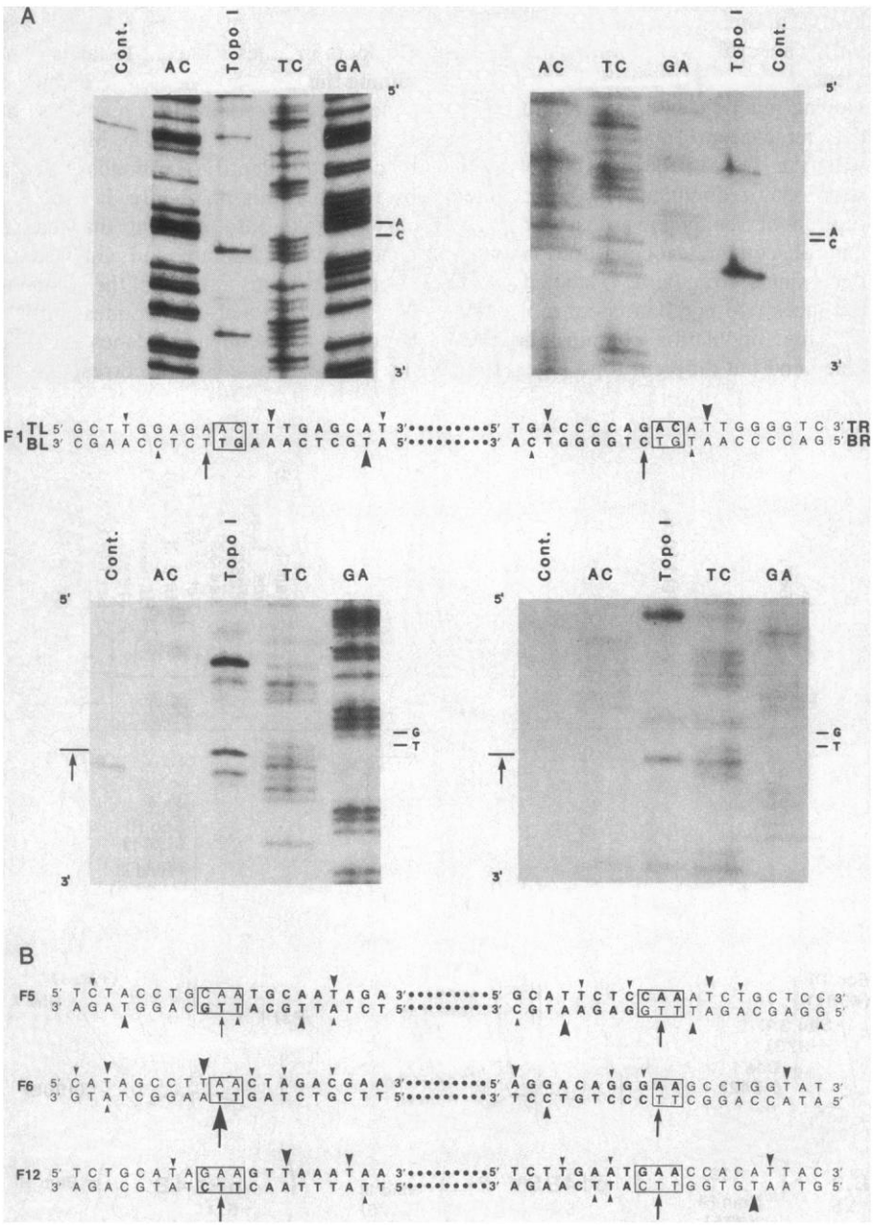

Fig. 2. Topo I nicking sites around crossover points. Sequences that define crossover points used during excisional recombination are enclosed in rectangles. Sequences in bold type were found in the excised forms; the ambiguous origin of the nucleotides at the crossover points (a result of the small homology at the crossover points) is indicated by both bold and nonbold type. Topo I cleavage sites that were nearest to the crossover points (positioned such that a simple ligation could account for the junction sequences found in the excised forms) are indicated by arrows with tails. Additional topo I cleavage sites within a 20-bp window of the crossover points, and the relative strengths of the cleavages, are designated by arrowheads of different sizes. The position of the arrow in the sequence indicates the nucleotide to which topo I becomes covalently linked upon severance of the bond; the phosphodiester bond cleaved by the enzyme is adjacent to the arrow towards the 3′ end of the DNA strand. (A) Examples of the data obtained for the pair of parental crossover points F1 (Fig. 1) are shown in the four panels; each panel is an analysis of cleavage sites on one strand of the duplex. The panels are arranged so that the analysis of the top right strand is at the top right and the bottom right strand is on the bottom right panel, and so on. The lanes marked "cont." show the background bands seen in the labeled strand without topoisomerase treatment; lanes marked "topo I" show the fragments created in the labeled strand by topo I cleavage; and the lanes AC, TC, and GA are the sequencing ladder markers dephosphorylated as described below. TL, top left strand; BL, bottom left strand; TR, top right strand; BR, bottom right strand (*16*). On the autoradiograms in the bottom right and left panels, the arrows mark the specific topo I breaks that are used in the model. (B) Summary of data from F5, F6, and F12.

For any pair of parental crossover points, it was necessary to assay both the top and bottom strands to assess the potential association of topo I cleavage with a particular recombination event. Therefore, four nicking reactions were required for each pair of crossover sites (Fig 2A) (*1, 2*). Strikingly, the exact crossover points used during F5, F6, and F12 excision were associated with topo I cleavage sites. In the case of F1, the parental crossover points are within one nucleotide of topo I cleavage sites. Thus, all pairs of excision crossover points are associated with topo I cleavage sites, and most are cleaved by topo I within the crossover points. Similar experiments performed with single-stranded substrates did not show any correlation between topo I cleavage sites and the crossover sites.

The probability for the association of topo I cleavage sites with pairs of crossover points (Fig. 2) can be calculated on the basis of the frequency of occurrence of break sites within the DNA. As the frequency of break sites within the 20-bp windows did not differ significantly from the frequency observed over the entire length of fragments studied (one break per seven nucleotides), the probability of breakage at any position was taken to be 0.14. Thus, for a pair of sites with 3 bp of homology, the probability of a break site at the same position within the homologous region is $1 - (48/49)^3 = 0.06$. For 2 bp of homology, the same calculation yields a value of 0.04. The probability that break sites would occur in any three pairs of three bases (F1 is treated as a three-base overlap for this calculation) and a single pair of two bases chosen at random is therefore $(0.06)^3(.04) \cong 10^{-5}$.

The arrangement and statistical significance of the topo I cleavage sites associated with the four pairs of crossover points has suggested a model for topo I–mediated excisional recombination (Fig. 3A). Eukaryotic topo I introduces transient breaks into DNA as an intermediate in the nicking-closing reaction. However, it probably does not bridge both ends of a transient break at all times (*21*). Therefore, a nearby break or gap in the DNA strand opposite to the one cleaved by topo I could lead to a double-stranded break if the number of base pairs between the two break sites were insufficient to hold the two ends together (Fig.

3A, step 2). Possible sources for a lesion in the strand of DNA opposite to the topo I break include additional topo I cleavages or nicks and gaps resulting from unscheduled rounds of proviral replication (*22*). The initial substrate for excisional recombination is therefore proposed to be chromosomal DNA that is structurally aberrant such that topo I breaks the DNA rather than catalyzing cycles of nicking and closing (Fig. 3A, step 1). Initiation of in situ "onion skin" viral replication caused by fusion to permissive cells (*22*, *23*) might create newly replicated DNA fragments with internal nicks to which topo I has ready access.

The suggestion that altered forms of chromatin, or perhaps regions of naked DNA, present structural features that promote recombination is also supported by the high rate of illegitimate recombination that accompanies transfection of both nonreplicative and replicative DNA into tissue culture cells (*24–26*). Fragments of double-stranded DNA released from the chromosome may be prone to recombination through the action of other enzymes; however, the molecular ends produced by topo I cleavage would be already primed for cyclization. Double-stranded fragments resulting from two topo I–mediated breaks for the same strand would contain a strand with one end terminating in a 5′ hydroxyl and the other end terminating in a 3′ phosphate coupled to a topo I molecule (Fig. 3A, step 2). Such termini are substrates for eukaryotic topo I–catalyzed ligations, and it is proposed that the juxtaposition of the ends undergoing ligation is facilitated by pairing the small regions of homology at the crossover points. The arrangement of the topo I–generated ends undergoing ligation need only account for one strand of the junction sequence in the excision products, since a single strand could give rise to the sequences in the cloned products by serving as the template for the subsequent rounds of DNA replication or repair. The positions of the topo I cleavage sites associated with the crossover points are consistent with the proposal that only one strand is formed by a topo I ligation (Figs. 2 and 3B). No other reactions are required for F5, F6, and F12. The F1 junction sequence would be formed with either strand if, in addition to a topo I–catalyzed ligation, a single nucleotide were converted during excision (Fig. 3B).

Finally, since the 5′ hydroxyl ends are not protected by a bound topo I molecule (as are the 3′ ends), the initial position of the 5′ ends may be altered by exonucleolytic degradation. Such alterations could account for the changes in the sequences that have been described at certain recombinant joints and may indeed obscure the nature of the primary sequences involved (*27*, *28*). Furthermore, our model relates to replicating molecules but can also be a mechanism for recombination in any situation wherein gaps or nicks are present (such as transfection or chromosomal DNA being processed by repair enzymes).

We have previously noted that many, but not all, illegitimate recombination crossover points in viruses and chromosomes are associated with potential topo I recognition sites (*2*). The specifics of the simple model for recombination presented above need explicit testing; however, the data on topo I cleavage specificity suggest that this enzyme can play a role in catalyzing nonhomologous recombinations in somatic cells.

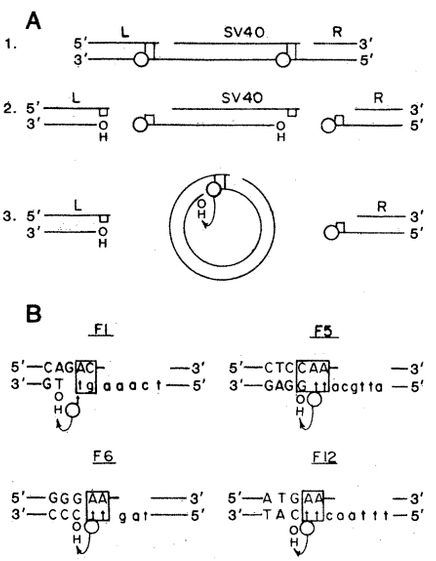

Fig. 3. A model for topo I–mediated excisional recombination. (A) In step 1, topo I molecules (open circles) bind to sequences in the vicinity of the crossover points (rectangles). Nicking-closing activity in the vicinity of a break in the opposite strand (indicated by the gaps) would result in a double-stranded break (step 2). Topo I–mediated formation of phosphodiester bonds and repair of the opposite strand would result in cyclization of the excised DNA-containing viral sequences (step 3). A topo I–catalyzed ligation of the left (L) and right (R) chromosomal ends would repair the chromosome, leaving a simple deletion. As onion skin replication initiates this process, a simple deletion would not result from one such event. However, a series of such deletion events could yield a simple deletion (*33*). (B) The cyclization reaction described in step 3 of the model is drawn with sequences suggested by our in vitro data. Sequences to the left of the proviral insert are written in lower case letters, and SV40 and right chromosomal sequences are written in uppercase letters. Sequences within rectangles define the parental crossover points. Except for a single conversion event shown in F1 (note the G-T mismatch), the suggested topo I–mediated ligations would yield the junction sequences found in the excision products.

References and Notes

1. M. Botchan, J. Stringer, T. Mitchison, J. Sambrook, *Cell* **20**, 143 (1980).
2. P. Bullock, W. Forrester, M. Botchan, *J. Mol. Biol.* **174**, 55 (1984).
3. M. D. Been, R. R. Burgess, J. J. Champoux, *Nucleic Acids Res.* **12**, 3097 (1984).
4. K. A. Edwards, B. D. Halligan, J. L. Davis, N. L. Nivera, L. F. Liu, *ibid.* **10**, 2565 (1982).
5. M. D. Been and J. J. Champoux, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2883 (1981).
6. B. D. Halligan, J. L. Davis, K. A. Edwards, L. F. Liu, *J. Biol. Chem.* **257**, 3995 (1982).
7. J. J. Champoux, L. S. Young, M. D. Been, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 53 (1979).
8. W. Keller, U. Muller, I. Eicken, I. Wendel, H. Zentgraf, *ibid.* **42**, 227 (1977).
9. B. L. McConaughy, L. S. Young, J. J. Champoux, *Biochim. Biophys. Acta* **655**, 1 (1981).
10. K. Javahesian and L. F. Liu, *Nucleic Acids Res.* **11**, 461 (1983).
11. M. D. Been and J. J. Champoux, *J. Mol. Biol.* **180**, 515 (1984).
12. J. J. Champoux, *Proc. Natl. Acad Sci. U.S.A.* **73**, 3488 (1976).
13. ———, *ibid.* **74**, 3800 (1977).
14. ———, *J. Mol. Biol.* **118**, 441 (1978).
15. J. C. Wang, *ibid.* **55**, 523 (1971).
16. We describe in detail how the labeled ends were obtained for the four F1 reactions (Fig. 2) *Top left strand (TL)*: Plasmid pL14B.3 is a recombinant plasmid containing the genomic fragment extending from the Hae III sites at −170 to −1292, blunt end ligated into the pBR322 C1a I site. A 1159-bp fragment, labeled at the Eco RI site, was isolated after cleavage of the pBR322 Hind III site. *Bottom left strand (BL)*: An 830-bp fragment extending from the Pst I site at −958 to the pBR322 Hae III site at 4344 was isolated from plasmid pL14B.3, digested with Sau 3A, and labeled. A 248-bp fragment extending from the labeled Sau 3A site at −373 to the Hae III site at 4344 was then isolated. (In order to remove a fragment that would have comigrated with the desired band in the preparative gel, we used Rsa I to cleave an Rsa I site at −669.) *Top right strand (TR)*: Plasmid p14BSV-2 contains the 1425-bp fragment bordered by the SV40 Hind III sites at 1046 and 3476 inserted into the pBR322 Hind III site. This plasmid was restricted with Eco RI and the resulting ends were labeled at their 3′ termini. After cleavage of the Bam HI site at 2533, a labeled 525-bp fragment was isolated. *Bottom right strand (BR)*: Upon cleavage of plasmid p14BSV-2 with Bam HI and Pst I, a 1277-bp fragment extending from the Bam HI site at 2533 to the pBR322 Pst I site at 3609 was isolated. This fragment was cleaved with Bst NI and labeled; after recleavage with Eco RI, a DNA fragment extending from the Bst NI site at 964 to the Eco RI site was isolated. For the other reactions we list the coordinates of the fragments used, the restriction sites that mark their ends, and the plasmid source. The 3′ label position is shown by an asterisk.
    F5: TL *Nco I (−3249) → Eco RI
    BL *Hind III (−3426) → Hae III (−2890)
    Plasmid source (TL and BL); pL14B.1
    TR *Eco RI → Sst I
    Plasmid source; pR14B
    BR *Hinf I (+161) → Hae III (+484)
    Plasmid source; pR14B
    F6: TL *Fnu4H1 (−972) → Hae III (−1292)
    BL *Dde I (−1140) → Sau 961 (−878)
    Plasmid source (TL and BL); pL 14B.1
    TR *Sau 3A (874) → Aha III (2363)
    BR *Dde I (2374) → Hind III (1046)
    Plasmid source (TL and BL); p14BSV-2
    F12: TL *Sau 3A (−373) → Alu (−705)
    Plasmid source; pL14B.1
    BL *Dde I (−742) → pBR Hae III (4344)
    Plasmid source; pL14B.3
    TR *Hind III (+652) → Sst (+149)
    BR *Dde I (+562) → Hind III (+652)
    Plasmid source (TR and BR); pR14B
17. A. Maxam and A. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
18. D. A. Bencini, G. A. O'Donovan, J. R. Wild, *Biotechniques* **1**, 4 (1984).
19. D. P. Tapper and D. A. Clayton, *Nucleic Acids Res.* **9**, 6787 (1981).
20. Rat liver topo I was purified as described [Martin *et al.*, *Methods Enzymol.* **100**, 137 (1983)]. Experimental conditions for topo I breakage of end-labeled duplex fragments are essentially

those of Been *et al.* (*3*). However, upon termination of the topo I reactions with SDS, butanol (rather than ethanol) was used to precipitate the reactions. The modifications of the Maxam and Gilbert (*17*) sequencing reactions described by Bencini *et al.* (*18*) were used to isolate markers. The 3′-labeled sequencing markers were dephosphorylated as follows. After butanol precipitation of the piperidine reactions, the samples were dried and resuspended in 47 μl of H_2O, 2.5 μl of 1*M* tris (*p*H 8.0), and 0.5 μl of bacterial alkaline phosphatase (International Biotechnologies Inc., 24 unit/ml). After incubation at 65°C for 60 minutes, 2 volumes of 1 percent SDS and 1 ml of butanol were used to precipitate the dephosphorylated markers. Polyacrylamide-urea sequencing gels (8 percent; 40 cm long and 8.3 mm thick) were run for various times at 1600 V.

21. M. D. Been and J. J. Champoux, *Nucleic Acids Res.* **8**, 6129 (1980).
22. M. Botchan, W. Topp, J. Sambrook, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 709 (1979).
23. J. Miller, P. Bullock, M. Botchan, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7534 (1984).
24. M. P. Calos, J. S. Lebkowski, M. R. Botchan, *ibid.* **80**, 3015 (1983).
25. C. T. Wake, T. Gudewiczm, T. Porter, A. White, J. H. Wilson, *Mol. Cell. Biol.* **4**, 387 (1984).
26. K. R. Folger, K. Thomas, M. R. Capecchi, *ibid.* **5**, 59 (1985).
27. J. H. Wilson, P. B. Berget, J. M. Pipas, *ibid.* **2**, 1258 (1982).
28. M. Woodworth-Gutai, *Virology* **109**, 353 (1981).
29. J. Tooze, Ed., *DNS Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982).
30. J. G. Sutcliffe, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 77 (1979).
31. T. Maniatis, A. Jeffrey, H. Vande Sande, *Biochemistry* **17**, 3787 (1975).
32. D. R. Smith and J. M. Calvo, *Nucleic Acids Res.* **8**, 2255 (1980).
33. V. Coulantoni, L. Pauley, G. Della Valle, C. Basilico, *J. Virol.* **43**, 618 (1982).
34. We thank F. Dean and M. Been for many helpful discussions and N. Cozzarelli for comments on the manuscript. Supported by American Cancer Society grant MV-91 and grant CA 30490 from the National Cancer Institute (M.B.) and by grant PCM 8300944 from the National Science Foundation (J.C.).

28 May 1985; accepted 21 August 1985

Secretion of a Bacterial Cellulase by Yeast

Abstract. *Gene fusions were constructed between a yeast expression plasmid and a* Cellulomonas fimi *DNA fragment encoding an endo-1,4-β-D-glucanase or carboxymethylcellulase. Yeast transformed with the recombinant plasmids secreted carboxymethylcellulase activity. Secretion of active enzyme was greatly increased when the leader of a secreted yeast protein, the K1 toxin, was inserted immediately upstream of and in frame with the bacterial cellulase sequence. This is the first step in constructing a functional cellulase complex in* Saccharomyces cerevisiae. *It also provides an excellent system for the detailed examination of the determinants of protein secretion because of the ease with which secreted cellulase can be detected.*

NIGEL SKIPPER
MARGARET SUTHERLAND
R. WAYNE DAVIES
Division of Molecular Biology, Allelix, Inc., Mississauga, Ontario, Canada L4V 1P1
DOUGLAS KILBURN
ROBERT C. MILLER, JR.
ANTHONY WARREN
RAYMOND WONG
Department of Microbiology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

The microbial utilization of crystalline cellulose as a source of glucose for growth requires a group of enzymes, the cellulases. These include an extracellular endo-1,4-β-D-glucanase and exo-1,4-β-D-glucanase and an intracellular 1,4-β-glucosidase (*1*). The glucanases are not produced by the yeast *Saccharomyces cerevisiae*, an organism widely used to ferment glucose to ethanol and carbon dioxide. We introduced a DNA sequence encoding a naturally secreted bacterial cellulase into a yeast expression vector and report that yeast transformed with the vector secreted the bacterial cellulase efficiently if a yeast secretion signal was incorporated into the construction.

The gene for an endo-1,4-β-D-glucanase (carboxymethylcellulase or CM-cellulase) has been cloned on a 5-kilobase (kb) DNA fragment from the cellulolytic bacterium *Cellulomonas fimi* (*2*). The gene is expressed in *Escherichia coli* C600 and the CM-cellulase is largely periplasmic (*2*). A subclone containing a 2.4-kb fragment behaves in a similar manner. This subclone, plasmid pEC2.2, was used for the experiments with yeast. The 2.4-kb fragment may lack the sequence encoding the NH_2-terminal region of the CM-cellulase; it is expressed as a fusion protein when inserted in the tetracycline-resistance gene of plasmid pBR322. Plasmids pOP and pN3 (Fig. 1) were used to examine CM-cellulase expression in yeast. Plasmid pOP contains a modified complementary DNA (cDNA) copy of a yeast double-stranded RNA sequence that encodes preprotoxin (*3*, *4*). The cDNA is contained between the alcohol dehydrogenase 1 promoter and the iso-1–cytochrome c terminator. Expression of the cDNA from the alcohol dehydrogenase 1 promoter in yeast has been described; the cells produce extracellular K1 toxin, and a toxin immunity determinant (Fig. 1) (*4*). Protoxin, the first detectable precursor of K1 toxin in yeast cells (*5*), enters the yeast secretory apparatus. The pre-region of preprotoxin is probably a signal peptide for cotranslational insertion of protoxin into the endoplasmic reticulum and as such might be useful for directing secretion of the CM-cellulase. The Bgl II restriction site in pOP (Fig. 1) provides a cloning site for insertion of the CM-cellulase gene a few codons downstream from this pre-region. Plasmid pN3 was a derivative of pOP in which a 0.17-kb Bal I fragment from the cDNA had been replaced by a synthetic oligonucleotide. This generated an Nco I restriction site, suitable for insertion of the CM-cellulase gene downstream from the $Lys^{232} \cdot Arg^{233}$ codons of preprotoxin (Fig. 1). Since the Tyr^{234} of preprotoxin is the NH_2-terminal residue of the β-polypeptide of secreted toxin (*5*), the $Lys^{232} \cdot Arg^{233}$ residues may specify cleavage of protoxin and so the release of the β-polypeptide or of any polypeptide substituted for β.

Figure 2 shows the structure of plasmids constructed to determine if transformed yeast cells can express CM-cellulase. From previous experiments, it was known that the translational reading frame in these constructions would put the open reading frame of the CM-cellulase in phase with the initiator ATG of preprotoxin. The CM-cellulase sequence was fused to either the first three (pL5.19), the first 52 (pK2.4), or the first 233 (pK1.3) codons of preprotoxin. In this way we could determine whether the bacterial enzyme could be secreted in the absence of a yeast leader peptide, in the presence of the preprotoxin leader peptide or pre-region, or in the presence of all the preprotoxin except for the COOH-terminal region that is the β-polypeptide (Fig. 1) of extracellular K1 toxin. In pL5.19 and pK1.3 there were four and three additional codons, respectively, between the yeast and bacterial sequences. This was due to the cloning manipulations.

Each plasmid was introduced into the tryptophan-dependent *S. cerevisiae* strains 2OB12 (*6*) and SX34-4D (*7*), by LiCl transformation (*8*). Tryptophan-independent cells were then streaked out on a selective medium containing carboxymethyl cellulose, and the colonies were assayed for extracellular CM-cellulase activity by the Congo red procedure (*9*). Each plasmid directed the synthesis of extracellular CM-cellulase (Fig. 3), although pL5.19 and pK1.3 gave much lower activity than pK2.4. Cells contain-