reaction buffer; pH was monitored and maintained at 7.5 throughout the reaction. Calcium was not critical for the inhibition of GA protein by lipids since addition of I m*M* EGTA or 50 μ*M* $CaCl_2$ did not alter the GTPase activity of GA protein in the presence or absence of phosphatidic acid.

24. J. C. Lacal and S. A. Aaronson, *Mol. Cell. Biol.* **6**, 4214 (1986).
25. We thank J. K. Smith for contributions made during the formative stage of this study; J. B. Gibbs for his gift of purified GA protein; and M. J. Newman and H. R. Kaback for critical review of this manuscript.

19 September 1988; accepted 23 November 1988

Switch Protein Alters Specificity of RNA Polymerase Containing a Compartment-Specific Sigma Factor

LEE KROOS,* BARBARA KUNKEL, RICHARD LOSICK†

During sporulation in *Bacillus subtilis*, expression of developmental genes *spoIVCB* and *cotD* is induced in the mother cell compartment of the sporangium at morphological stages IV and V, respectively. A 27-kilodalton RNA polymerase sigma factor called σ^K (or σ^{27}) has been found that causes weak transcription of *spoIVCB* and strong transcription of *cotD*. A 14-kD protein was also discovered that changes the specificity of σ^K-containing RNA polymerase, greatly stimulating *spoIVCB* transcription and markedly repressing *cotD* transcription. Both σ^K and the 14-kD protein are products of genes known to be required for expression of specific genes in the mother cell. Thus, σ^K directs gene expression in the mother cell and it is proposed that inactivation or sequestering of the 14-kD protein switches the temporal pattern of gene expression during the transition from stages IV to V of development.

A HALLMARK OF THE PROCESS OF endospore formation in the Gram-positive bacterium *Bacillus subtilis* is the formation of a sporangium composed of two compartments, the mother cell and the forespore (*1*). The compartments each receive a chromosome generated during the last vegetative round of DNA replication, but then undergo divergent developmental fates. The mechanisms governing temporal and spatial control of gene expression in the mother cell chamber of the sporangium have not been well studied. The sporulation genes *spoIIIC* and *spoIVCB*, the spore coat protein genes *cotA*, *cotC*, and *cotD*, and the germination gene *gerE* are examples of genes that are known or are inferred to be transcribed selectively in the mother cell (*2–6*). These genes fall into three temporal classes that are switched on at successive developmental stages. We report the use of cloned copies of *spoIVCB* (*7*)—which is switched on at stage IV (*3*)—and *cotD* (*5*)—which is switched on at stage V (*6*)—as templates for reconstructing compartment-specific gene transcription in vitro.

Run-off transcripts from linear templates bearing promoters for the mother cell–expressed genes *cotD* and *spoIVCB* were generated by RNA polymerase that had been partially purified (*8*) from sporulating cells of *B. subtilis* (Fig. 1). The transcripts were of the expected sizes (within 5 bases) for initiation at the known in vivo start sites of the *cotD* (*6*) and *spoIVCB* (*3*) genes, and the length of the run-off RNA's varied predictably with templates that extended for different distances downstream from the start sites (Fig. 1).

To identify the factor or factors responsible for *cotD* and *spoIVCB* transcription, the RNA polymerase was fractionated by gradient elution from a DNA-cellulose column (Fig. 2). The *cotD*-transcribing activity (A) eluted at slightly higher salt concentration (reaching a peak in fractions 24 to 26) than did *spoIVCB*-transcribing activity (Fig. 2B), which peaked in fraction 22. Moreover, longer exposure (*9*) of the same gel showed that the enzyme activity eluting at low salt (fractions 14 to 16) had almost exclusively *spoIVCB*-transcribing activity, whereas RNA polymerase in the higher salt-eluting fractions (fractions 30 to 34) had almost exclusively *cotD*-transcribing activity.

Proteins from the gradient elution were displayed by electrophoresis on an SDS-polyacrylamide slab gel (Fig. 2C). Several proteins showed an elution pattern that was approximately coincident with the distribution of *cotD*-transcribing activity, but reconstruction experiments showed that only one, a 27-kD polypeptide, directed *cotD* transcription. When, for example, proteins in gel slices in the 25- to 30-kD size range were eluted, renatured, and added to *B. subtilis* core RNA polymerase (Fig. 3A), only renatured protein corresponding to the 27-kD polypeptide activated *cotD* run-off transcription by core enzyme (Fig. 3B). Evidently, the 27-kD protein is a sigma factor for the *cotD* gene.

Because *spoIVCB* is expressed earlier than *cotD* during sporulation and because *spoIVCB*-transcribing activity eluted at lower salt concentration compared to *cotD*-transcribing activity (Fig. 2, A and B), another sigma factor may have been responsible for transcription of *spoIVCB*. However, proteins of many different sizes were tested (*9*), but only protein from the same 27-kD gel slice that stimulated *cotD* transcription (Fig. 3B) caused core RNA polymerase to transcribe from the *spoIVCB* promoter (Fig. 3C) (although the run-off transcription was weaker than that observed with *cotD*). One possible explanation was that the 27-kD gel slice contained two different proteins, one a sigma factor for *cotD* transcription that was

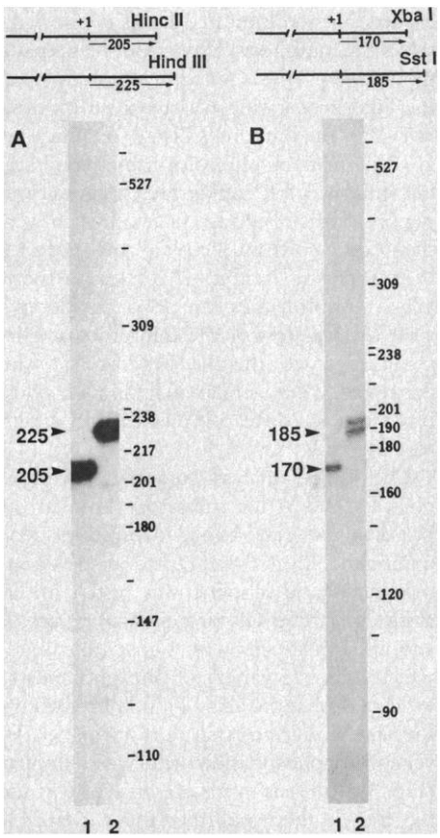

Fig. 1. Run-off transcription from templates containing the *cotD* and *spoIVCB* promoters. Linearized plasmid templates (2 μg) (top) were transcribed in 40-μl reaction mixtures (*17*) with RNA polymerase partially purified from sporulating cells of *B. subtilis* (*8, 18*). The products of run-off transcription were displayed by electrophoresis in a 5 percent polyacrylamide slab gel containing 8*M* urea and detected by autoradiography. (**A**) Run-off transcripts from the *cotD* promoter-containing template linearized with Hinc II (lane 1) or Hind III (lane 2). (**B**) Transcripts from the *spoIVCB* promoter-containing template linearized with Xba I (lane 1) or *Sst*I (lane 2).

Department of Cellular and Developmental Biology, The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138.

*Present address: Department of Biochemistry, Michigan State University, East Lansing, MI 48824.
†To whom correspondence should be addressed.

more abundant in fractions eluting at higher salt concentrations and the other a sigma factor for *spoIVCB* transcription that was more abundant in fractions eluting at lower salt concentrations. This possibility was ruled out when protein was eluted and renatured from gel slices of 27-kD protein from low- and high-salt eluting fractions that had been separately subjected to gel electrophoresis. Renatured 27-kD protein from the low- and the high-salt eluting fractions provoked identical patterns of strong *cotD* and weak *spoIVCB* transcription when added to core enzyme and challenged with an equimolar mixture of promoter-containing templates (*9*). Thus, a single 27-kD species of sigma factor, which we call σ^K (or σ^{27}), directed transcription of both *cotD* and *spoIVCB* DNA's.

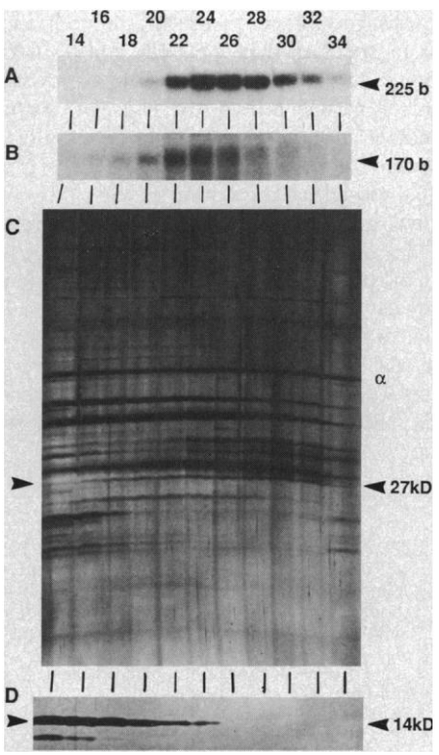

Fig. 2. Fractionation of *cotD*- and *spoIVCB*-transcribing activities. RNA polymerase was partially purified (*8, 18*) and eluted from a DNA-cellulose column with a linear gradient of 0.5 to 1.3*M* KCl. Enzyme in 10-μl samples from the gradient elution fractions was tested for transcribing activity, as described in Fig. 1, with an equimolar mixture (2-μg total) of *cotD* template linearized with Hind III (**A**) and *spoIVCB* template linearized with Xba I (**B**). Exposure of the gel to x-ray film was four times longer in (B) than in (A) to compensate for the lower level of *spoIVCB* transcription. Proteins in the gradient elution fractions were subjected to SDS–polyacrylamide gel electrophoresis (PAGE) and stained with silver. (**C**) A 12.5 percent polyacrylamide gel displaying intermediate-to-large polypeptides. (**D**) An 18 percent polyacrylamide gel (*20*) showing proteins of low molecular size. The amount of protein placed on the gel was five times greater in (C) than in (D).

Another possible explanation for the partial separation of transcribing activities was that the relative preference of σ^K-RNA polymerase ($E\sigma^K$) for the two templates was influenced by an auxiliary transcription factor in the fractions from the DNA-cellulose column, and that the salt concentration for elution of this factor was different from that of $E\sigma^K$. Such a factor was found when protein from the gradient was subjected to electrophoresis in a gel system that resolves proteins of small molecular size. The transcription factor can be seen in Fig. 2D as a small, highly abundant polypeptide (from 8 to 14 kD depending on the gel system used, but we refer to it here as 14 kD) that was preferentially eluted at the low-salt end of the gradient. When purified by preparative gel electrophoresis and added to $E\sigma^K$, the 14-kD protein altered the relative levels of transcription of *cotD* and *spoIVCB*, and inhibited run-off RNA synthesis from the former and greatly stimulated run-off transcription of the latter (Fig. 4A). Thus, the presence of the 14-kD factor, which we call a switch protein, and its elution earlier than σ^K from DNA-cellulose explains the partial separation of *cotD* and *spoIVCB* transcribing activities observed in Fig. 2.

The stimulation of *spoIVCB* transcription and the inhibition of *cotD* RNA synthesis are independent effects of the switch protein because each effect could be observed in the absence of the other template (Fig. 4B). Thus, in reactions that contained only the *spoIVCB*-bearing template, the addition to core enzyme of both σ^K and the 14-kD protein (lane 7) provoked a strong run-off transcription signal, whereas the addition to core enzyme of σ^K alone (lane 6) had only a modest effect in stimulating run-off transcription. [The addition to core enzyme of the 14-kD protein alone (lane 5) had no effect, an indication that the 14-kD protein is not itself a sigma factor.] Conversely, in reactions that contained only the *cotD*-bearing template the addition to $E\sigma^K$ of increasing quantities of 14-kD protein (Fig. 4, lanes 2 to 4, with lane 2 displaying the effect of the highest quantity of protein added) inhibited *cotD* RNA synthesis. Other characterization of the 14-kD protein, which is highly abundant in sporulating cells, indicates that it oligomerizes, and gel retardation and template competition experiments (*9*) suggest that it exerts its effects on the selectivity of transcription by binding to the DNA templates.

To study the in vivo function of the newly discovered transcription factors, we sought to identify the structural genes for the 14-kD and σ^K proteins by determining their partial NH_2-terminal amino acid sequences. The proteins were purified by preparative polyacrylamide gel electrophoresis, electroblotted onto membranes, and subjected directly to sequential Edman degradation in an automated gas-phase sequenator (*10*). The 14-kD protein was purified from pooled fractions 14 to 18 from the DNA-cellulose gradient elution (Fig. 2) by electro-

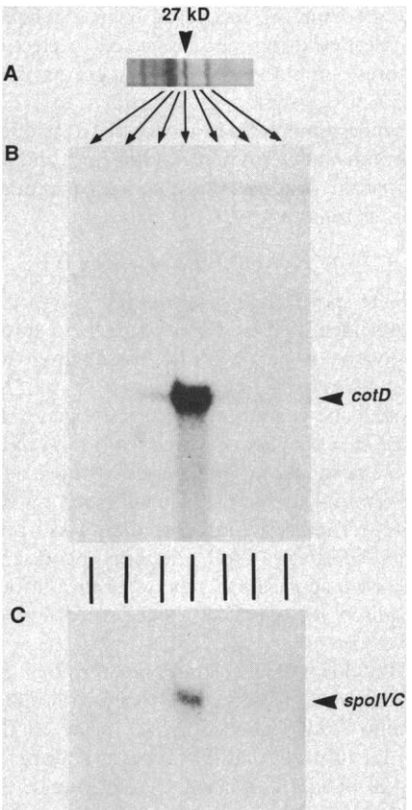

Fig. 3. Reconstruction of *cotD*- and *spoIVCB*-transcribing activities with a 27-kD protein and core RNA polymerase. Proteins were eluted from gel slices, and renatured (*21*), incubated on ice for 10 minutes with *B. subtilis* core RNA polymerase prepared as described (*22*) and tested for transcribing activity as described for Fig. 1. (**A**) A section of polyacrylamide gel in which protein in fraction 24 of the experiment of Fig. 2 was subjected to electrophoresis (with the direction of protein migration from left to right) to display proteins in the vicinity of the 27-kD protein. Elution and renaturation was performed on slices from a similarly prepared gel but with 50 times as much protein and stained with Coomassie blue. (**B**) Autoradiogram of the 225-base region of a gel displaying the products of run-off transcription from *cotD* template linearized with Hind III by core RNA polymerase supplemented with renatured protein from the slices corresponding to those indicated in (A). (**C**) Autoradiogram of the 170-base region of a gel of run-off transcripts from *spoIVCB* template linearized with Xba I by core RNA polymerase supplemented with renatured protein from the same gel slices as used in (B). Approximately 15 ng (if we assume complete recovery from the gel) of 27-kD protein was added to 60 ng of core RNA polymerase in the reactions that produced run-off transcripts.

phoresis in the gel system described in the legend to Fig. 2D. It yielded the following 34 residue-long NH_2-terminal sequence (*11*):

MHDYIKERTIKIGKYIVETKKTVR-VIAKEFGVSK

The σ^K factor was purified in a two-step procedure in which the protein in pooled fractions 19 to 34 was subjected to electrophoresis in a 10 to 20 percent polyacrylamide gradient gel, recovered from a gel slice by electroelution, and subjected to electrophoresis in a second gradient gel prior to sequence analysis, which yielded an NH_2-terminal amino acid sequence of 13 residues, then an undetermined residue (−), and six uncertain residues (indicated by parentheses) as follows:

YVKNNAFPQPLSS–(EEKKYL)

The sporulation gene *spoIIID* was previously identified as a candidate for a direct regulator of *spoIVCB* (*3*). Anticipating that it encodes the 14-kD protein or σ^K, we (*12*) cloned and sequenced the *spoIIID* gene, the predicted product of which was found to be a 93-residue-long polypeptide whose first 34 amino acids exactly conformed to the NH_2-terminal sequence of the 14-kD protein. Thus, the 14-kD protein is encoded by a known sporulation gene whose product is required for *spoIVCB* gene expression in vivo (*3*).

To our surprise, the NH_2-terminal sequence of σ^K matched perfectly with a 20–amino acid sequence (with serine at the undetermined residue) found at codons 21 to 40 of *spoIVCB* itself (*3, 13*), one of the genes whose transcription σ^K directed in vitro. As reported by Stragier *et al.* (*13*), *spoIVCB* is a truncated gene that encodes only the NH_2-terminus of σ^K. The COOH-terminus is specified by a separate gene *spoIIIC* to which *spoIVCB* becomes joined by a chromosomal rearrangement during sporulation. Thus, σ^K is encoded by a composite *spoIVCB-spoIIIC* gene, also called *sigK* (*13*).

The assignment of *spoIIID* as the structural gene for the 14-kD switch protein suggests that the *spoIIID* gene product is a temporal regulator of the mother cell line of gene expression, and that its inactivation or sequestering during the transition from stage IV (cortex formation) to stage V (coat formation) of sporulation causes a switch in the overall pattern of gene expression. In our scheme, *spoIIID* protein turns on the σ^K structural gene (*spoIVCB-spoIIIC*) and then, by acting in conjunction with $E\sigma^K$, activates the transcription of other stage IV genes, including genes involved in cortex formation. At the same time, the presence of the *spoIIID* gene product during stage IV prevents the transcription by $E\sigma^K$ of *cotD* and certain other coat genes. Next, during the transition to stage V, *spoIIID* protein is inactivated (or somehow rendered inaccessible to the mother cell chromosome), thereby turning off further transcription of stage IV genes and activating the transcription of *cotD* and certain other stage V genes.

Finally, we draw the following inferences from the discovery that σ^K is encoded by a composite of the *spoIVCB* and *spoIIIC* genes. (i) Transcription of *spoIVCB-spoIIIC* is in part autoregulatory, as its product directs its own transcription. Indeed, *spoIVCB* and *spoIIIC* mutants are known to be partially impaired in *spoIVCB* transcription (*3*). Furthermore, some *spoIVCB-spoIIIC* transcription must be directed, at least initially, by a holoenzyme form present earlier than $E\sigma^K$ in order to "prime the pump." Nevertheless, both categories of *spoIVCB-spoIIIC* RNA synthesis depend on the 14-kD protein because a *spoIIID* mutant is completely blocked in transcription from the *spoIVCB* promoter. (ii) The primary product of the σ^K structural gene is a proprotein bearing 20 extra residues at the NH_2-terminus of the sigma factor, as deduced from the observation that the NH_2-terminal sequence of mature σ^K commences at codon 21 of *spoIVCB*. If analogous to the sporulation sigma factor σ^E (*14*), the pro-σ^K may be an inactive precursor whose processing to the mature, active form is triggered by a landmark, morphological feature of the stage III to IV sporangium. (iii) σ^K is localized to the mother cell chamber of the sporangium, and hence it is a transcription factor specific to the mother cell. This follows from earlier work (*3*) showing that *spoIVCB* transcription is principally or exclusively confined to the mother cell and from the fact that *spoIVCB* and *spoIIIC* mutations selectively block gene expression in the mother cell (*4, 6*). Thus, σ^K is the mother cell counterpart to the recently discovered forespore sigma factor σ^G (*15*). Interestingly, transcription of the σ^G structural gene *spoIIIG* is also autoregulatory (*16*).

In summary, we have discovered two key regulatory components of the mother cell line of gene expression, and our data suggest a way in which the components govern temporal and spatial aspects of developmental gene expression during endospore formation.

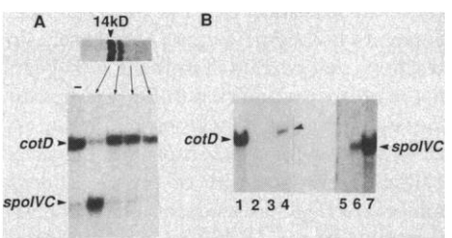

Fig. 4. A 14-kD protein switches the specificity of transcription by σ^K-RNA polymerase. (**A**) Gel purification of the 14-kD protein. Proteins were eluted from gel slices, renatured (*21*), and added to transcription reactions immediately after the addition of $E\sigma^K$ (reconstituted as in Fig. 3). At the top is a section of a Coomassie blue–stained protein gel prepared as described for Fig. 2D (except with 250 times more protein) in which protein in fraction 16 of the experiment of Fig. 2 was subjected to electrophoresis (with the direction of protein migration from left to right) to display protein in the vicinity of the 14-kD protein. (Bottom) Autoradiogram showing 225-base *cotD* transcript and 170-base *spoIVCB* transcript from an equimolar mixture (2-μg total) of templates linearized with Hind III and Xba I, respectively, produced by $E\sigma^K$ supplemented with the indicated renatured proteins. (−) indicates that no gel-eluted protein had been added. Approximately 240 ng (if we assume complete recovery from the gel) of 14-kD protein was present in the reaction that produced the altered pattern of transcription. (**B**) The stimulatory and inhibitory effects of the 14-kD protein are separable. (Lanes 1 to 4) Autoradiographs of 225-base *cotD* transcript produced by $E\sigma^K$ with no additions (lane 1) or with the addition of the 14-kD protein at 240 ng (lane 2), 80 ng (lane 3), and 27 ng (lane 4). (Lanes 5 to 7) Autoradiographs of 170-base *spoIVCB* transcript produced by core RNA polymerase (60 ng) to which 14-kD protein alone (240 ng) had been added (lane 5), to which σ^K alone (15 ng) had been added (lane 6), or to which σ^K and 14-kD protein had been added (lane 7).

REFERENCES AND NOTES

1. P. Piggot and J. Coote, *Bacteriol. Rev.* **40**, 908 (1976); R. Losick and P. Youngman, in *Microbial Development*, R. Losick and L. Shapiro, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1984), pp. 63–88; R. Losick, P. Youngman, P. J. Piggot, *Annu. Rev. Genet.* **20**, 625 (1986).
2. S. M. Turner, J. Errington, J. Mandelstam, *J. Gen. Microbiol.* **132**, 2995 (1986).
3. B. Kunkel, K. Sandman, S. Panzer, P. Youngman, R. Losick, *J. Bacteriol.* **170**, 3513 (1988).
4. K. Sandman, L. Kroos, S. Cutting, P. Youngman, R. Losick, *J. Mol. Biol.* **200**, 461 (1988); S. Cutting and J. Mandelstam, *J. Gen. Microbiol.* **132**, 3013 (1986); S. Cutting, S. Panzer, R. Losick, *J. Mol. Biol.*, in press.
5. W. Donovan, L. Zheng, K. Sandman, R. Losick, *J. Mol. Biol.* **196**, 1 (1987).
6. L. Zheng and R. Losick, unpublished data.
7. *spoIVCB* is the *B* gene of the *spoIVC* locus, which consists of two separately transcribed *A* and *B* genes [(*3*); R. Farquhar and M. D. Yudkin, *J. Gen. Microbiol.* **134**, 9 (1988).
8. C. W. Cummings and W. G. Haldenwang, *J. Bacteriol.* **170**, 5863 (1988).
9. L. Kroos and R. Losick, unpublished data.
10. M. W. Hunkapillar, E. Lujan, F. Ostrander, L. E. Hood, *Methods Enzymol.* **91**, 227 (1983); P. Matsudaira, *J. Biol. Chem.* **262**, 10035 (1987).
11. Single letter abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
12. B. Kunkel, L. Kroos, H. Poth, P. Youngman, R. Losick, unpublished data.

13. P. Stragier, B. Kunkel, L. Kroos, R. Losick, *Science* **243**, 507 (1989).
14. T. L. LaBell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1784 (1987); P. Stragier, C. Bonamy, C. Karmazyn-Campelli, *Cell* **52**, 697 (1988).
15. D. Sun *et al.*, *Genes Dev.* **3**, 141 (1989).
16. C. Karmazyn-Campelli *et al.*, *ibid.*, p. 150.
17. Transcription reactions were performed as described by H. L. Carter and C. P. Moran [*Proc. Natl. Acad. Sci. U.S.A.* **83**, 9438 (1986)] except that the polymerase was allowed to bind to template for 10 minutes at 37°C before the addition of nucleotides heparin (6 μg) was added 1 minute after the addition of RNA polymerase to prevent reinitiation, and, after the reactions were stopped, 30 μl of reaction mixture was subjected to electrophoresis. The plasmid containing the *cotD* promoter, pLRK100, was constructed by insertion of a 430-bp *B. subtilis* Eco RI–Hinc II DNA fragment from pBD156 (*5, 6*) into Eco RI– and Hinc II–digested pUC19 [C. Yanisch-Perron, J. Vieira, J. Messing, *Gene* **33**, 103 (1985)]. The in vivo transcription start site indicated for *cotD* at the top of the figure is based on unpublished results (*6*). The *spoIVCB* promoter-containing plasmid, pBK16, was constructed by insertion of an approximately 400-bp Pst I–Sac I *B. subtilis* DNA fragment from pBK9 (*3*) that had been made blunt-ended at the Sac I site by digestion with mung bean nuclease into Pst I– and Hinc II–digested pUC19. The indicated in vivo *spoIVCB* transcription start site results from (*3*).
18. The RNA polymerase used in the experiment of Fig. 1 was a fraction that eluted at high-salt concentrations from a DNA-cellulose column (provided by C. W. Cummings and W. G. Haldenwang). For Fig. 2 we used the following modification of their gradient elution procedure (*8*). *Bacillus subtilis* strain SC104 (S. Cutting and R. Losick, unpublished data) containing a *cotA-lacZ* translational fusion (which is similarly regulated to *cotD*) was harvested during sporulation when the gene fusion was substantially induced as monitored by assaying *cotA*-directed β-galactosidase synthesis. Cells were washed with harvest buffer [T. Linn, A. L. Greenleaf, R. Losick, *J. Biol. Chem.* **250**, 9256 (1975)] supplemented with 10 percent glycerol and then with buffer I (*19*) supplemented with 5 percent (v/v) phenylmethylsulfonyl fluoride (PMSF) (6 mg/ml in 95 percent ethanol) and stored at −70°C. Cells (56 g) were resuspended in 140 ml of buffer I containing PMSF, passaged twice through a French Pressure Cell (15,000 psi), sonicated for 1 minute (450 watts), and centrifuged for 90 minutes at 120,000*g* and 4°C. The clarified supernatant was subjected to heparin-agarose and DNA-cellulose column chromatography as described (*8*) except that the DNA-cellulose column (15 ml) was washed with 15 ml of buffer C containing 0.1*M* KCl and then with 15 ml of buffer C containing 0.5*M* KCl prior to elution with a linear gradient (120 ml) of 0.5*M* to 1.3*M* KCl in buffer C. Fraction 14 (Fig. 2) represents the beginning of the salt gradient. Fractions (4 ml) were collected and dialyzed into storage buffer (*19*) modified to contain 0.1*M* KCl and were stored at −20°C.
19. R. G. Shorenstein and R. Losick, *J. Biol. Chem.* **248**, 6163 (1973).
20. J. O. Thomas and R. D. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2626 (1975).
21. D. Hager and R. R. Burgess, *Anal. Biochem.* **109**, 76 (1980).
22. Core RNA polymerase was prepared from *B. subtilis* strain ML21, which contains deletion mutations of the sigma factor encoding-genes *sigB* and *sigH* (M. Lampe and R. Losick, unpublished), as described (*19*) except that a Bio-Rex 70 column was substituted [R. R. Burgess and J. J. Jendrisak, *Biochemistry* **14**, 4634 (1975)] for the phosphocellulose column.
23. We thank W. G. Haldenwang for invaluable assistance during the early stages of this work and P. Stragier for pointing out that the σ^K amino acid sequence matches that deduced for the *spoIVCB* product and for stimulating suggestions during the final stages of the project. Supported by NIH grant GM18568 (R.L.), a postdoctoral fellowship from the Helen Hay Whitney Foundation (L.K.), and a predoctoral fellowship from NSF (B.K.).

10 November 1988; accepted 22 December 1988

Amplification and Molecular Cloning of HTLV-I Sequences from DNA of Multiple Sclerosis Patients

E. PREMKUMAR REDDY,* MAGNHILD SANDBERG-WOLLHEIM, RICHARD V. METTUS, PHILLIP E. RAY, ELAINE DEFREITAS, HILARY KOPROWSKI

Techniques of gene amplification, molecular cloning, and sequence analysis were used to test for the presence of sequences related to human T-lymphotropic virus type I (HTLV-I) in peripheral blood mononuclear cells of six patients with multiple sclerosis (MS) and 20 normal individuals. HTLV-I sequences were detected in all six MS patients and in one individual from the control group by DNA blot analysis and molecular cloning of amplified DNAs. The viral sequences in MS patients were associated with adherent cell populations consisting predominantly of monocytes and macrophages. Molecular cloning and nucleotide sequence analysis indicated that these amplified viral sequences were related to the HTLV-I proviral genome.

ALTHOUGH EARLY DATA INDICATED that neurologic disorders of wild mice are caused by mouse retrovirus infection (*1*), no direct correlation has been established between human neurological disorders and retroviruses. However, recent evidence suggests that the human immunodeficiency virus (HIV-1) as well as HTLV-I and HTLV-II are associated with neurological disorders (*2–11*).

We previously (*4*) reported the presence of antibodies that react with the HTLV-I *gag* (p24) protein in samples of serum and cerebrospinal fluid (CSF) from patients with multiple sclerosis (MS) in Sweden and in Florida. Sequences of HTLV-I were detected by in situ hybridization analysis in the lymphocytes of cultures of CSF cells obtained from one-third of the patients (*4*). In three patients with progressive chronic encephalomyelopathy, we noted three different patterns of reactivity in relation to HTLV-I infection (*8, 9*): (i) the presence in serum of HTLV-I antibodies together with viral sequences, as determined by in situ hybridization and by the detection of viral antigen in lymphocytes; (ii) the absence of HTLV-I antibodies in the presence of viral sequences in CSF cells; and (iii) the presence of HTLV-I antibodies in the absence of viral sequences in peripheral blood lymphocytes (PBL) and CSF cells (*8, 9*). Subsequently, HTLV-I was isolated from lymphocyte culture (*10*) of one of these patients, and the proviral genomes were cloned (*11*) from these cells.

Since HTLV-I sequences can be detected by in situ hybridization assays in very few lymphocytes of MS patients, never in more than one in 10^4 or one in 10^5 cells (*4*), and since the absence of detectable antibodies to HTLV-I does not exclude HTLV-I infection, we used the polymerase chain reaction (*12, 13*) coupled with forced cloning of the amplified DNA to look for the presence of HTLV-I sequences in three males and three females with MS as defined according to McDonald and Halliday (*14*). The patients were from Sweden and all were diagnosed and treated at the Department of Neurology, University Hospital, University of Lund, Sweden. The age at onset of disease ranged from 11 to 34 years, and the duration, from 2 to 12 years (Table 1). All patients had a moderate mononuclear pleocytosis in CSF and five had oligoclonal immunoglobulin on isoelectrofocusing. The patient with no oligoclonal bands had multiple bilateral lesions on magnetic resonance imaging compatible with MS. At the time of the study, three patients had exacerbations and one of these was treated with corticosteroids. None of the patients had ever received a blood transfusion. Negative controls were ten healthy subjects from Sweden who had no antibodies for HTLV-I and ten healthy blood donors from Philadelphia, Pennsylvania.

Blood samples were obtained from the cubital vein, and the peripheral blood mononuclear cells (PBMC) were processed (*15*). Results of ELISA assays with the use of disrupted HTLV-I virions revealed the presence of antibodies in serum samples from two of the six patients at the time their PBMC were tested for the presence of HTLV-I sequences (Table 1). To amplify HTLV-I sequences in the cellular DNA extracted from PBMC, we used primer pairs from the *gag* and *env* region (*16*). These were 25 bases long and rich in G-C content to allow stable hybridization. The *gag* primers included the recognition sequence for Sma I or Pst I and the *env* primers contained the

E. P. Reddy, R. V. Mettus, P. E. Ray, E. DeFreitas, H. Koprowski, The Wistar Institute of Anatomy and Biology, 36th Street at Spruce, Philadelphia, PA 19104.
M. Sandberg-Wollheim, Department of Neurology, University of Lund, S-22185 Lund, Sweden.

*To whom correspondence should be addressed.