5. M. N. Goldstein, K. Hamm, E. Amrod, *Science* **151**, 1555 (1966); D. Kessel and I. Wodinsky, *Cancer Res.* **28**, 938 (1968); K. Danø, *Biochim. Biophys. Acta* **323**, 466 (1973); W. A. Bleyer, S. A. Frisby, V. T. Oliverio, *Biochem. Pharmacol.* **24**, 633 (1975); P. D. Minor and D. H. Roscoe, *J. Cell Sci.* **17**, 381 (1975); M. Inaba and R. K. Johnson, *Biochem. Pharmacol.* **27**, 2123 (1978); T. Skovsgaard, *ibid.*, p. 1221; W. T. Beck, T. J. Mueller, L. R. Tanzer, *Cancer Res.* **39**, 2070 (1979).
6. R. H. F. Peterson and J. L. Biedler, *J. Supramol. Struct.* **9**, 289 (1978); R. H. F. Peterson *et al.*, *Cancer Res.* **43**, 222 (1983).
7. R. L. Juliano and V. Ling, *Biochim. Biophys. Acta* **455**, 152 (1976).
8. W. T. Beck, T. J. Mueller, L. R. Tanzer, *Cancer Res.* **39**, 2070 (1979).
9. D. Garman and M. S. Center, *Biochem. Biophys. Res. Commun.* **105**, 157 (1982).
10. J. R. Riordan *et al.*, *Nature (London)* **316**, 817 (1985).
11. M. B. Meyers and J. L. Biedler, *Biochem. Biophys. Res. Commun.* **99**, 228 (1981).
12. J. L. Biedler, H. Riehm, R. H. F. Peterson, B. A. Spengler, *J. Nat. Cancer Inst.* **55**, 671 (1975).
13. J. L. Biedler and R. H. F. Peterson, *Molecular Actions and Targets for Cancer Chemotherapeutic Agents*, A. C. Sartorelli, J. Lazo, J. R. Bertino, Eds. (Academic Press, New York, 1981), pp. 453–482.
14. M. B. Meyers, B. A. Spengler, T-D. Chang, P. W. Melera, J. L. Biedler, *J. Cell Biol.* **100**, 588 (1985).
15. O. Brison, F. Ardeghir, G. R. Stark, *Mol. Cell. Biol.* **2**, 578 (1982).
16. K. T. Montgomery, J. L. Biedler, B. A. Spengler, P. W. Melera, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5724 (1983); R. W. Michitsch, K. T. Montgomery, P. W. Melera, *Mol. Cell. Biol.* **4**, 2370 (1984).
17. J. L. Biedler *et al.*, *Proc. Am. Assoc. Cancer Res.* **21**, 292 (1980).
18. DNA was dissolved in 100 n*M* tris-HCl, *p*H 7.4, 0.2*N* NaOH and 6.6× SSC and heated to 80°C for 10 minutes. After neutralization with 2*M* tris-HCl, it was spotted onto nitrocellulose. Filters were dried, incubated in hybridization buffer without the labeled probe, and then hybridized (Fig. 2).
19. Y.-H. Chien and I. B. Dawid, *Mol. Cell. Biol.* **4**, 507 (1984).
20. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1976).
21. P. W. Melera *et al.*, *Mol. Cell. Biol.* **4**, 38 (1984).
22. M. Grunstein and D. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961 (1975).
23. M. V. Norgard, *Anal. Biochem.* **113**, 34 (1981).
24. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977).
25. J. M. Chirgwin *et al.*, *Biochemistry* **18**, 5294 (1979).
26. P. S. Thomas, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5201 (1980).
27. I. B. Roninson *et al.*, *Nature (London)* **309**, 626 (1984).
28. J. L. Biedler, T-D. Chang, P. W. Melera, K. W. Scotto, B. A. Spengler, *J. Cell. Biochem. Suppl.* **10A** (Liss, New York, in press).
29. We are indebted to P. Borst and colleagues for providing p-glycoprotein cDNA clone cp28 (A. M. Van der Blick, T. Van der Velde-Koerts, V. Ling, P. Borst, *Mol. Cell. Biol.*, in press) with which our cDNA clone p5L-18 cross-hybridizes under normal stringency conditions. Clone cp28 in turn cross-hybridizes with the p-glycoprotein encoding cDNA pCHP1 that was isolated from a λgt11 expression library with a p-glycoprotein specific monoclonal antibody (*10*).
30. F. M. Sirotnak *et al.*, *J. Cell. Physiol.* **126**, 266 (1986).
31. Supported in part by NIH grants CA-08748, CA-28595, and CA-09207. We thank J. Davide for assistance and O. M. Rosen for a critical reading of the manuscript.

13 December 1985; accepted 1 April 1986

Activation of the AIDS Retrovirus Promoter by the Cellular Transcription Factor, Sp1

KATHERINE A. JONES, JAMES T. KADONAGA, PAUL A. LUCIW, ROBERT TJIAN

The nature and position of transcriptional control elements responsible for the expression of genes encoded by the retrovirus associated with acquired immune deficiency syndrome (AIDS) have not been precisely defined. In this study it is shown that the mammalian Sp1 transcription factor binds to promoter sequences within the AIDS retrovirus long terminal repeat (LTR) and activates RNA synthesis five- to eightfold in reconstituted reactions in vitro. Experiments in which regions of DNA were protected from added reagents by specifically bound proteins (footprinting) indicated that the upstream promoter region of the AIDS virus LTR lies between −45 and −77 (relative to the RNA start site, +1) and contains three tandem, closely spaced Sp1 binding sites of variable affinity. Base-substitution mutations targeted to one or all three Sp1 binding sites were found both to eliminate the binding of Sp1 and to cause up to a tenfold reduction in transcriptional efficiency in vitro. These findings suggest that one important component of the AIDS virus transcriptional control region interacts with a cellular transcription factor, Sp1, and that this factor must function in conjunction with transcriptional elements located downstream of the RNA cap site to mediate the response of the LTR to viral *trans*-activation.

TRANSCRIPTION OF RETROVIRAL GENOMES is mediated by promoter elements that are located within the viral long terminal repeat (LTR) segments, and can be strongly regulated by viral *trans*-activator proteins. Because these elements and those of other viruses of eukaryotic cells rely on the cellular RNA polymerase II transcriptional apparatus, they are considered excellent models for analysis of the mechanism by which RNA synthesis is initiated and regulated. A combined genetic and biochemical approach to the study of promoter recognition has proved particularly useful in defining the role of specific RNA polymerase II transcription elements. For example, it has recently been shown that certain viral and cellular genes bind a cellular transcription factor designated Sp1, and these specific protein–DNA interactions contribute directly to promoter strength in vitro (*1–4*) as well as in vivo (*5, 6*). These studies have further revealed the following biochemical properties of Sp1 (*7*): (i) A promoter of Sp1 recognizes an asymmetric decanucleotide sequence (G/T GGGCGG-PuPuPy) with an affinity determined by the match of a given sequence to the consensus. (ii) A single Sp1 binding site is sufficient to activate transcription, and both binding and transcriptional activation are independent of the orientation of the binding site relative to the RNA start site. (iii) Sp1-responsive promoters generally contain multiple Sp1 binding sites located within 40 to 150 bp upstream of the RNA start site, and Sp1 binding sites are sometimes intermingled with those for other promoter-specific transcription factors. In the latter case, the coordinate interaction of two or more different factors may be required for transcriptional activity of the promoter (*3*). (iv) Weaker binding sites may be positioned at optimal locations adjacent to the TATA box, and these may be equivalent to or even more important for promoter strength than higher affinity Sp1 sites located at more distal positions. (v) Genes that are recognized by Sp1 are expressed or regulated in vivo by a variety of distinct mechanisms, including enhancer activation [for example, the SV40 early promoter (*5, 8*)], viral *trans*-activation of promoter elements [for example, the thymidine kinase gene of herpes simplex virus (HSV) (*9*)], and induction through upstream response elements [for example, the HSV immediate-early and human metallothionein genes (*10*)]. Here we report that Sp1 is also an important component of the promoter for a lymphocytopathic RNA virus, the retrovirus [variously termed ARV-2, LAV, or HTLV-III (*11, 12*)] associated with the acquired immune deficiency syndrome (AIDS) and related syndromes (*13, 14*).

Promoter and enhancer elements within the LTR segments of retroviruses characteristically generate transcripts of full genomic length that may be differentially spliced to yield a variety of mRNA species (*15*). Inspection of the AIDS retrovirus LTR re-

K. A. Jones, J. T. Kadonaga, R. Tjian, Department of Biochemistry, University of California, Berkeley, CA 94720.
P. A. Luciw, Chiron Research Laboratories, Chiron Corporation, Emeryville, CA 94608.

vealed that sequences upstream of the RNA start site harbor a decanucleotide that is identical to some of the previously characterized Sp1 binding sites. This suggested to us that Sp1 might mediate the promoter function of this virus. Accordingly, we evaluated the ability of Sp1 to bind the LTR of a subclone of ARV-2 (*11*) by means of experiments in which regions of DNA were protected from deoxyribonuclease I (DNase I) digestion (footprint analysis) and dimethyl sulfate (DMS) methylation (*16*) by specific binding of proteins. The Sp1 protein fractions used in all of the experiments reported here were obtained by extensive fractionation of a HeLa cell nuclear extract, yielding Sp1 fractions of an estimated 85% homogeneity (*17*) (Fig. 1C).

The DNase I footprint experiments revealed that Sp1 binds to the LTR and protects a 40-bp region of DNA (−43 to −83) from nuclease digestion (Fig. 1A). The affinity of Sp1 for the binding region was comparable to that observed for some of the sites in the SV40 promoter. The extended pattern of footprint protection indicated that the LTR contains two or three tandem Sp1 binding sites, because binding to a single site is known to protect only 18 to 20 bp of DNA (*7*). To delineate the binding region further, we analyzed the guanine residues protected by the binding of Sp1 from DMS methylation. A complex pattern of enhancement and suppression of methylation at the N7 position of individual guanines was observed in the presence of Sp1 (Fig. 1B). In contrast to the protection pattern previously observed on the SV40 early promoter (*18*), protection of guanines on both strands of the ARV-2 LTR was observed. Therefore Sp1 appeared to form close contacts with both strands of the DNA helix, consistent with the recent finding that Sp1 does not bind efficiently to single-stranded DNA (*19*).

The methylation-protection experiments delineated a 35-bp binding region, and we noted three potential recognition elements that are directly repeated within this region: (I) 5′-GGGGAGTGGC-3′, −46 to −55; (II) 5′-TGGGCGGGAC-3′, −57 to −66; (III) 5′-GAGGCGTGGC-3′ −68 to −77 (Fig. 1, arrows). Guanines within each element were protected from DMS in the presence of Sp1, which suggests that the ARV-2 LTR binds three tandem Sp1 protomers. An analysis of the binding site occupancy at different protein concentrations indicated that the distal site (III) has a greater affinity for Sp1 than the other two sites and is comparable in magnitude to the high-affinity binding sites of GC boxes 3 and 5 of the SV40 promoter (*2*). Elements I and III are the first binding sites we have characterized that deviate from the Sp1 recognition sequence at positions within the core GC box (GGGCGG).

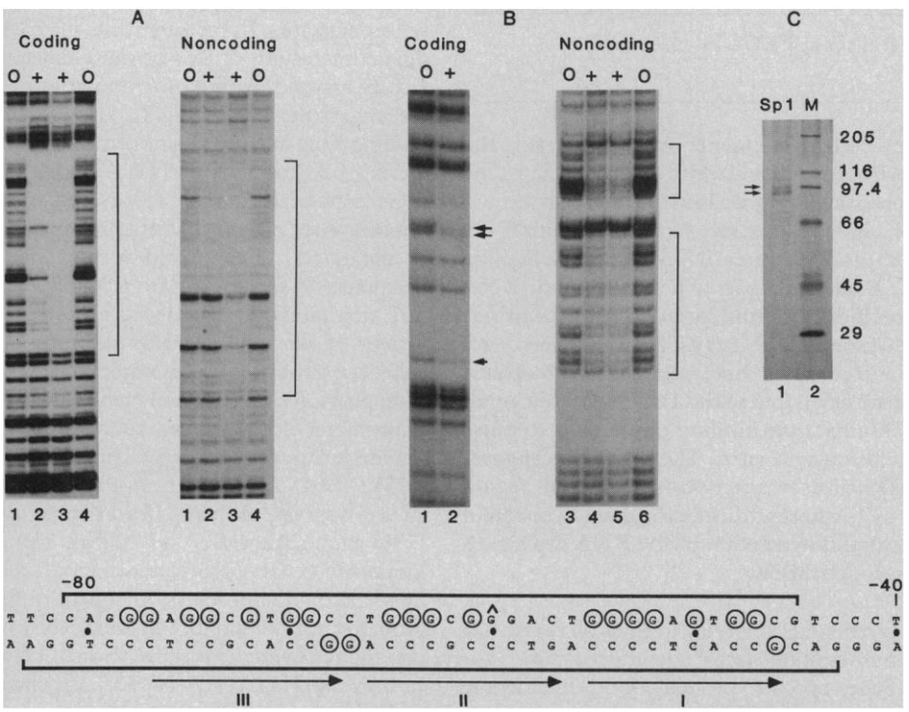

Fig. 1. Binding of the human Sp1 transcription factor to sequences in the ARV-2 LTR. (A) DNase I footprint experiments. ARV-2 DNA probes were end-labeled at the Bgl II site (+21) at the 5′ (noncoding strand) or 3′ (coding strand) ends. Lanes 1 and 4 show the partial nuclease digestion pattern obtained in the absence of Sp1. Lanes 2 and 3 contain 10 n*M* and 20 n*M*, respectively, of DNA affinity–purified Sp1 monomer (*17*) (50 ng and 100 ng per 50-μl reaction). The footprint protocol is otherwise as described (*3*), except that competitor DNA was omitted. Sp1 protein stocks were stored at approximately 10 ng/μl in 25 m*M* Hepes (K^+), *p*H 7.8, 0.5*M* KCl, 12.5 m*M* $MgCl_2$, 1 m*M* dithiothreitol, 20% glycerol, 0.1% Nonidet P-40 at −80°C. (B) DMS methylation-protection. DNA probes were 5′ end-labeled at the Bgl II site (+21, coding) or the Taq I site (−117, noncoding). Reactions contained 0 n*M* Sp1 (lanes 1, 3, and 6), 10 n*M* (50 ng per 50-μl reaction, lanes 2 and 4), or 40 n*M* (200 ng per 50-μl reaction, lane 5) Sp1 (monomer concentration). Protocol is as described (*18*). The sequence of the Sp1 binding region in the ARV-2 LTR is shown at the bottom; protected guanine residues are circled; capped guanine residue is enhanced in the presence of Sp1. Arrows, Sp1 recognition elements discussed in the text. (C) Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of Sp1 (arrows, lane 1), purified by DNA-affinity chromatography and characterized as described (*17*). The bands were visualized by staining with silver; M (lane 2), molecular size markers.

To obtain corroborating evidence for the structural arrangement of the Sp1 binding sites described above, we constructed two specific site-directed promoter substitution mutants. Previous studies on the interaction of Sp1 with the HSV thymidine kinase and SV40 early promoters revealed that base substitutions at positions 4, 6, 7, or 8 within the decanucleotide recognition sequence drastically reduce the binding affinity of Sp1 (*2–6*). An M13 oligonucleotide-directed mutagenesis procedure (*20*) was used to generate mutants dpm (III) and dpm (I-II-III) that contain G to T base transversion mutations at positions 3 and 4 of the decanucleotide recognition sequence of site III (bp −74, −75) and double base substitutions at the analogous positions of the proposed recognition elements for each of the three Sp1 sites (bp −52, −53, −63, −64, −74, and −75) (Fig. 2A). DNase I footprint analysis of the binding of Sp1 to these mutants is shown in Fig. 2B. The double base substitutions present in dpm (III) prevented binding to site III and resulted in a shortened (30 to 33 bp) footprint that was the consequence of Sp1 binding to elements I and II. Similarly, the multiple base substitutions in mutant dpm (I-II-III) interfered with the binding of Sp1 to each of the three sites. These results confirm that the AIDS virus LTR contains three distinct promoter domains that specify the binding of the cellular Sp1 transcription factor.

The location of the Sp1 binding sites upstream from the RNA start site suggested that the factor could contribute to the transcriptional activity of the AIDS virus promoter. We tested this possibility directly by means of transcription reconstitution experiments in vitro. Templates were constructed that contain the ARV-2 LTR promoter fused either to the bacterial chloramphenicol acetyltransferase (CAT) gene (*21*) at position +190 in the ARV-2 LTR, or to the

HSV thymidine kinase (TK) gene (*22*) at position +21 in the ARV-2 leader segment. The mutant promoters were also linked to the TK gene at position +21. RNA synthesized in vitro was detected by primer extension with synthetic primers complementary to CAT or TK gene sequences. As shown in Fig. 3, the ARV-2 promoter was efficiently recognized in vitro, and RNA synthesized from this promoter initiated accurately at positions used in vivo. Moreover, the transcriptional activity of the LTR-TK fusion gene suggested that sequences downstream of +21 are not required for promoter activity in vitro. We observed up to an eightfold activation in transcription upon the addition of Sp1 to fractions containing the cellular RNA polymerase II and other general transcription factors. Mutant dpm (I-II-III), which contains double base substitutions that eliminate the binding of Sp1 to each site (Fig. 2B), was not activated in the reconstitution experiment and was expressed at a level approximately tenfold below that of the wild-type promoter. The mutant that was deficient only in binding to the distal Sp1 site, dpm (III), was transcribed at a marginally reduced rate and was still activated by Sp1. These results indicate that sites I or II, or both sites, although weaker in affinity than site III, nevertheless play an important role in the activation of transcription from the AIDS promoter.

Expression of the AIDS retrovirus is typically restricted to T cells and certain other cells with the OKT4$^+$ surface receptor (*23*). In contrast, the AIDS virus promoter appears to be active in a variety of cell lines, including HeLa cells, when introduced by DNA-mediated transfection (*24*). Thus, both the viral promoter and enhancer elements appear to be of a general rather than cell type-specific nature. Because the tissue distribution of Sp1 is not well documented, however, it was important to confirm the presence of Sp1 in human T lymphocytes. We therefore prepared a whole-cell extract from a human T-cell line (Hut-78) that is receptive to viral infection and growth (*14, 25*). The extract was fractionated by chromatography on heparin-agarose, and the fraction obtained at the 0.2 to 0.4*M* KCl step was found to contain, on the basis of binding to the SV40 early promoter, at least as much Sp1 activity as the human HeLa cell line [for HeLa cells this number averages 50,000 molecules per cell (*17*)]. Thus the protein we have implicated as being important for LTR promoter activity is present in those cell lines that express LTR-linked genes.

Sp1 binding sites I and III do not contain the usual GGGCGG hexanucleotide, but this finding was not surprising because both sites contain an eight out of ten match with the decanucleotide consensus sequence (*7*). It thus appears that there is some degree of flexibility for deviation from the consensus, yet it is also likely that the sensitivity for alteration of each position is variable. For example, single G to C transversion mutations at positions 4, 6, 7, or 8 drastically reduce the binding affinity of Sp1, even though the data reported here suggest that a G to T transversion at position 7 is acceptable. In addition, guanine residues appear to be highly preferred at positions 3, 4, 6, and 9, whereas purines appear to be equivalent at positions 2 and 8. The least critical residues may be those at positions, 1 and 10.

The multiplicity of promoter elements appears to be a common phenomenon among RNA polymerase II genes, although experiments in vitro indicate that a single binding site can be sufficient to activate transcription (*4*). The data presented here indicate that sites I or II or both sites of the AIDS virus promoter are primarily responsible for promoter strength, whereas binding site III may be unimportant for stimulation of RNA synthesis. In contrast, optimal promoter strength of the SV40 early and HSV TK genes depends on three upstream transcription factor binding sites (*2, 3*). These distinctions may reflect differences in the relative affinity and position of each binding site.

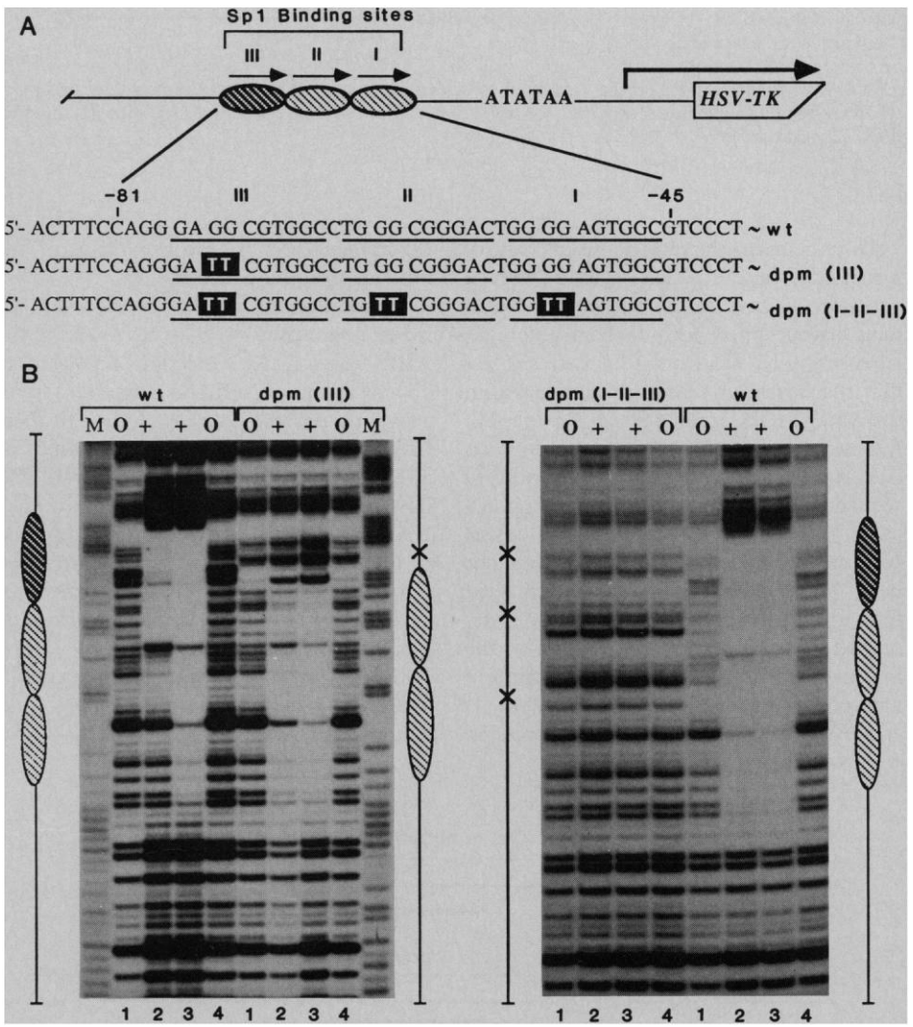

Fig. 2. Binding of the Sp1 transcription factor to site-specific mutants in the ARV-2 LTR. (A) (Top) Structure of the plasmid construct pLTR-TK (*32*) used in the oligonucleotide-directed mutagenesis procedure. Sp1 binding sites are shown with ellipses; dark diagonal striping marks the highest affinity site. (Bottom) Sequence of the wild-type (wt) and base-substitution mutants within the Sp1 binding domain. Mutations were confirmed by using the dideoxy sequencing technique (*33*) with an oligonucleotide primer complementary to the HSV-TK gene. (B) Autoradiograms of DNase I footprinting gels. Probes for the wild-type or mutant DNA's were 5′ end-labeled at the Bgl II site (+21, coding strand). Reactions that show the cleavage pattern in the absence of Sp1 are shown in lanes 1 and 4. Reactions in lanes 2 and 3 of the left panel contained 6 n*M* and 18 n*M* Sp1 (30 ng and 90 ng per 50-μl reaction) and reactions in lanes 2 and 3 of the right panel contained 8 n*M* and 16 n*M* of Sp1 (40 ng and 80 ng per 50-μl reaction). Ellipses indicate the regions of observed footprint protection, and X marks the location of double base substitution mutations. M, G + A sequence markers.

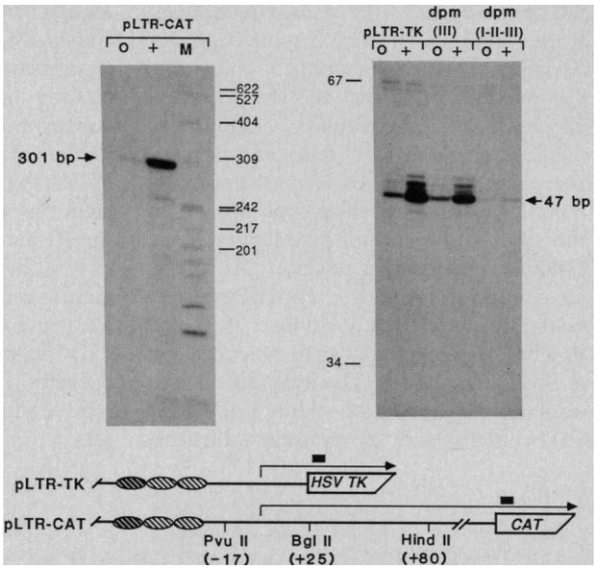

Fig. 3. Primer extension analysis of RNA synthesized in vitro from the ARV-2 LTR promoter. All reactions contained 36 μg of a protein fraction (C1. 225) containing the endogenous RNA polymerase II and 2 μg of a fraction, Sp2, that contains one or more general RNA polymerase II transcription factors (*34*). Final reaction volumes were 50 μl, including 140 ng of supercoiled DNA template (*32*), as indicated above each lane. The 0 and + lanes refer to Sp1 concentrations of 0 n*M* and 5 n*M*, respectively. Reactions were incubated for 45 minutes at 30°C, and 5′ end-labeled oligonucleotide primers were annealed for 45 minutes at 62°C; conditions were otherwise as described (*3*). Autoradiograms were exposed for approximately 30 minutes at −80°C with an intensifying screen, or without a screen for quantitation by densitometry. Arrows mark the sizes of cDNA products expected for accurately initiated RNA transcripts. M, Hpa II–digested pBR322 size markers.

To understand the transcriptional control mechanism of the AIDS virus it will be necessary to characterize in detail the proteins that recognize both the basal and regulated promoter elements. Our data suggest that the upstream promoter region within the ARV-2 LTR extends from −45 to −77. A tenfold decrease in gene expression has been observed in vivo upon deletion of sequences that include the Sp1-binding domain [−104 to −44 (*26*)]. Nevertheless, because the viral enhancer elements have not been precisely mapped or distinguished from the upstream promoter elements described here, an accurate assessment of the importance of Sp1 for viral transcription will require the analysis of mutants dpm (I) and dpm (I-II-III) in transient expression experiments and by reconstruction of the mutants into the viral genome.

In addition to the basal promoter elements, the expression of genes linked to the AIDS virus LTR is strongly regulated in vivo by a virus-encoded 86 amino acid *trans*-activating protein (*24, 26*). Although there have been conflicting reports regarding the level at which this regulation is exerted (*26, 27*), a significant increase in steady-state RNA levels has been observed in response to the *trans*-activator protein in cotransfection experiments (*28*). Target sequences for *trans*-activation lie downstream of the Sp1 binding sites in a region between −17 and +80 (*26*), with a specific requirement for sequences downstream of the RNA start site, between +21 and +80 in the leader segment. The sequences required for viral gene expression under *trans*-activation conditions (that is, in infected cells) includes the Sp1 binding domain of the AIDS virus promoter as well as regulatory elements located downstream of the RNA cap site.

Although several examples of downstream enhancer elements have recently been described, it appears that transcriptional regulation mediated by the AIDS virus leader segment most closely resembles that of the adenovirus major late promoter (Fig. 4). Lewis and Manley (*29*) have shown that an element of the major late promoter (from the RNA cap site to position +33) responds to viral *trans*-activation, possibly by the adenovirus E1A protein. Moreover, untranslated leader sequences located further downstream of +33 are responsible for an additional five- to tenfold activation of the transcriptional initiation rate in adenovirus-infected cells (*30*). Although the manner by which these viral *trans*-activators recognize the target region is unknown, the observation that the E1A protein does not bind to DNA (*31*) suggests an indirect mechanism. As a model for this type of sequence-specific *trans*-activation, it will be informative to understand how the downstream viral regulatory elements function to modulate basal level transcription from these two viral promoters.

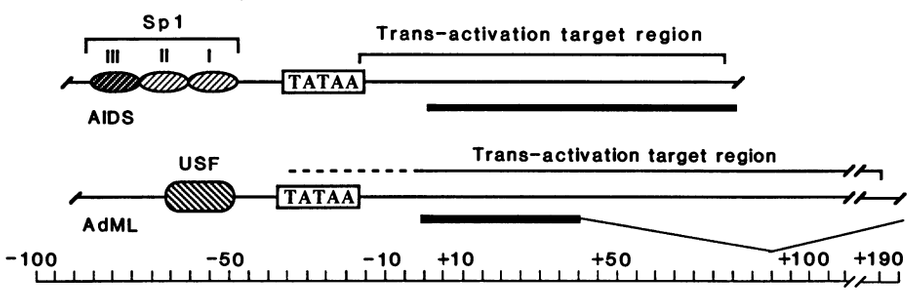

Fig. 4. Comparison of the promoter and regulatory domains of the ARV-2 and the major late transcription unit of adenovirus (Ad-2 ML). Ellipses mark the Sp1 binding sites on the ARV-2 LTR and the binding site on the adenovirus major late (AdML) gene for a different promoter-specific factor, designated USF (*35*). For simplicity, the footprint of the TATA binding protein of the AdML promoter is omitted. The relative position of a TATA homology of ARV-2 is shown, although a role for this sequence in directing AIDS transcription has not been established. Heavy lines below each diagram show untranslated leader segments of each RNA species. The *trans*-activation target sequences for the AIDS virus are as described (*24, 26*), and for the AdML promoter is a composite of the regions described by Lewis and Manley (*29*), and Mansour *et al.* (*30*).

REFERENCES AND NOTES

1. W. S. Dynan, S. Sazer, R. Tjian, R. T. Schimke, *Nature (London)* **319**, 246 (1986); S. Ishii *et al.*, *Science*, in press.
2. D. Gidoni *et al.*, *Science* **230**, 511 (1985).
3. K. A. Jones, K. R. Yamamoto, R. Tjian, *Cell* **42**, 559 (1985).
4. K. A. Jones and R. Tjian, *Nature (London)* **317**, 179 (1985).
5. C. Benoist and P. Chambon, *ibid.* **290**, 304 (1981).
6. R. D. Everett, D. Baty, P. Chambon, *Nucleic Acids Res.* **11**, 2447 (1983); S. L. McKnight and R. Kingsbury, *Science* **217**, 316 (1982); S. L. McKnight, R. C. Kingsbury, A. Spence, M. Smith, *Cell* **37**, 253 (1984).
7. J. T. Kadonaga, K. A. Jones, R. Tjian, *Trends Biochem. Sci.* **11**, 20 (1986).
8. J. Banerji, S. Rusconi, W. Schaffner, *Cell* **27**, 299 (1981); P. Gruss, R. Dhar, G. Khoury, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 943 (1981); M. Fromm and P. Berg, *J. Mol. Appl. Genet.* **111**, 457 (1982).
9. S. P. Eisenberg, D. M. Coen, S. L. McKnight, *Mol. Cell. Biol.* **5**, 1940 (1985); A. El Kareh, A. J. M. Murphy, T. Fichter, A. Efstratiadis, S. Silverstein, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1002 (1985).
10. J. C. Lang, D. A. Spandidos, N. M. Wilkie, *EMBO J.* **3**, 389 (1984); M. Karin *et al.*, *Nature (London)* **308**, 513 (1984).
11. R. Sanchez-Pescador *et al.*, *Science* **227**, 484 (1985).
12. S. Wain-Hobson, P. Sonigo, O. Danos, S. Cole, M. Alizon, *Cell* **40**, 9 (1985); L. Ratner *et al.*, *Nature (London)* **313**, 277 (1985); M. A. Muesing *et al.*, *ibid.*, p. 450; B. Starcich *et al.*, *Science* **227**, 538 (1985).
13. F. Barré-Sinoussi *et al.*, *Science* **220**, 868 (1983); F. Brun-Vezinet *et al.*, *Lancet* **1984-I**, 1253 (1984); R. C. Gallo *et al.*, *Science* **224**, 500 (1984); M. G. Sarngadharan, M. Popovic, L. Bruch, J. Schüpbach, R. C. Gallo, *ibid.*, p. 506.
14. M. Popovic, M. G. Sarngadharan, E. Read, R. C. Gallo, *Science* **224**, 497 (1984).
15. P. A. Luciw, J. M. Bishop, H. E. Varmus, *Cell* **33**, 705 (1983).
16. D. Galas and A. Schmitz. *Nucleic Acids Res.* **5**, 3157

(1978); U. Siebenlist and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* 77, 122 (1980).
17. M. Briggs, J. T. Kadonaga, R. Tjian, unpublished data; J. T. Kadonaga and R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
18. D. Gidoni, W. S. Dynan, R. Tjian, *Nature (London)* **312**, 409 (1984).
19. J. T. Kadonaga and R. Tjian, unpublished data.
20. M. J. Zoller and M. Smith, *Nucleic Acids Res.* **10**, 6487 (1982).
21. C. M. Gorman, L. F. Moffat, B. H. Howard, *Mol. Cell. Biol.* **2**, 1044 (1982); C. M. Gorman, G. T. Merlino, M. C. Willingham, I. Pastan, B. Howard, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6777 (1982).
22. S. L. McKnight, *Nucleic Acids Res.* **8**, 5949 (1980).
23. D. Klatzman *et al.*, *Science* **225**, 59 (1984); A. G. Dalgleish *et al.*, *Nature (London)* **312**, 763 (1984); J. S. McDougal *et al.*, *Science* **231**, 382 (1986).
24. J. Sodroski, R. Patarca, C. Rosen, F. Wong-Staal, W. A. Haseltine, *Science* **229**, 74 (1985); S. K. Arya, C. Guo, S. F. Josephs, F. Wong-Staal, *ibid.* **227**, 69 (1985).
25. J. A. Levy *et al.*, *ibid.* **225**, 840 (1984).
26. C. A. Rosen, J. G. Sodroski, W. A. Hazeltine, *Cell* **41**, 813 (1985).
27. C. A. Rosen *et al.*, *Nature (London)* **319**, 555 (1986).
28. B. M. Peterlin, M. A. Walker, P. A. Luciw, personal communication.
29. E. D. Lewis and J. L. Manley, *Mol. Cell. Biol.* **5**, 2433 (1985).
30. S. L. Mansour, T. Grodzicker, R. Tjian, *ibid.*, in press.
31. B. Ferguson *et al.*, *ibid.* **5**, 2653 (1985).
32. The pLTR-TK and pLTR-CAT plasmids were constructed as follows. (i) Plasmid pLTR-1 was constructed by ligation of an Xho I–Nar I fragment of the ARV-2 LTR with a Sal I–Nar I digest of a pML [M. Lusky, L. Berg, H. Weiher, and M. Botchan, *Mol. Cell. Biol.* **3**, 1108 (1983)] derivative that contains a polylinker insert (Eco RI–Xho I–Bam HI–Stu I–Hind III–Nar I) at the Eco RI site. (ii) Plasmid pLTR-1 was digested with Bgl II and Hind III, gel-purified and ligated with a gel-purified Bgl II–Hind III insert of the HSV-TK gene (*22*) to generate pLTR-TK. (iii) Plasmid pLTR-CAT contains a Stu I–Bam HI insert of the CAT gene (*21*) into the analogous sites of the polylinker sequence of pLTR-1. The vector for site-directed mutagenesis, pLTR-M13, contains the Kpn I–Sph I fragment of pLTR-TK inserted into the analogous sites of mp 18 [C. Yanisch-Perron, J. Vieira, J. Messing, *Gene* **33**, 103 (1985)].
33. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
34. W. S. Dynan and R. Tjian, *Cell* **32**, 669 (1983).
35. M. Sawadago and R. G. Roeder, *ibid.* **43**, 165 (1985).
36. We thank P. J. Barr for synthesis of the oligonucleotides used in the mutagenesis procedure, U. Heberlein for advice on generating the site-directed mutants, and S. Mansour for many helpful discussions. K.A.J. is the recipient of an NIH postdoctoral fellowship and J.T.K. is a fellow of the Miller Institute. This work was funded by the National Cancer Institute.

13 January 1986; accepted 13 March 1986

Defective Sialic Acid Egress from Isolated Fibroblast Lysosomes of Patients with Salla Disease

MARTIN RENLUND, FRANK TIETZE, WILLIAM A. GAHL*

Normal fibroblasts exposed to *N*-acetylmannosamine yielded lysosome-rich granular fractions loaded with free (unbound) sialic acid, whose velocity of egress increased with increasing initial loading. Fibroblast granular fractions of patients with Salla disease exhibited negligible egress of sialic acid, whether endogenous or derived from *N*-acetylmannosamine exposure. Salla disease represents the first disorder demonstrated to be caused by defective transport of a monosaccharide out of cellular lysosomes.

LYSOSOMAL STORAGE DISORDERS result from several different biochemical abnormalities. Lysosomal enzymes may be defective, enzyme activator proteins may be deficient, or newly synthesized enzymes may lack recognition markers essential for targeting them to the lysosome (*1*). Recently, another group of lysosomal storage disorders was described, characterized by defective transport of small molecules across the lysosomal membrane. The first disorder of this class was nephropathic cystinosis, a fatal disease in which the disulfide cystine fails to traverse the lysosomal membrane at a normal rate and, consequently, accumulates within lysosomes (*2*). Later, a disorder of vitamin B_{12} storage was ascribed to defective lysosomal transport of free vitamin B_{12} out of lysosomes (*3*). We now report that impaired lysosomal transport of a carbohydrate, *N*-acetylneuraminic acid (sialic acid; NANA), results in Salla disease, a disorder characterized by free sialic acid storage within lysosomes.

Patients with Salla disease, which is inherited in an autosomal recessive fashion, have moderate to severe psychomotor retardation, spasticity, and ataxia with an early onset and slow progression (*4*). The majority of the patients in whom the disease has been detected are in Finland. They excrete large amounts of free (unbound) NANA and store 10 to 30 times the normal amounts of this compound within several tissues and cultured fibroblasts (*5*). Electron microscopy of tissues reveals characteristic vacuolated cytoplasmic inclusions, that are apparently lysosomes swollen with high concentrations of free NANA (*4*). The amounts and distributions of NANA-containing membrane glycoproteins and gangliosides are normal in cells from patients with Salla disease ("Salla disease cells"), and a normal amount of acid neuraminidase activity cleaves NANA from glycoconjugates in Salla disease lysosomes (*5*). The fate of free NANA produced has not been rigorously determined, but *N*-acetylneuraminate pyruvate-lyase (E.C. 4.1.3.3.), a cytoplasmic enzyme, does cleave the neuraminic acid ring structure as the first step in NANA catabolism. In Salla disease fibroblasts this enzyme has normal activity (*5*), which led us to investigate whether free NANA transport out of lysosomes and into the cytoplasm is the defective process in the disorder.

Previous studies demonstrated that [^{3}H]NANA bound to low-density lipoprotein (LDL) was taken up normally by the lysosomes of Salla disease fibroblasts but was cleared at a much slower rate than in normal cells (*6*). In these studies the specific radioactivity of the labeled free NANA was diluted by large amounts of nonradioactive NANA present in the Salla cells but not in the normal cells. This made direct kinetic measurements of NANA egress in normal and Salla disease cells difficult to compare. In the present study we directly measured NANA egress by using lysosome-rich granular fractions and assays of nonradioactive NANA. Normal granular fractions were loaded with NANA to levels observed in Salla disease by exposure of intact fibroblasts to high concentrations of the NANA precursor, *N*-acetylmannosamine (ManNAc) (*7*).

Normal or Salla disease fibroblasts, grown to confluence in Eagle's minimum essential medium containing 10% fetal calf serum, 2 m*M* glutamine, streptomycin (100 μg/ml), and penicillin (100 U/ml), were exposed to 0 to 100 m*M* nonradioactive ManNAc for 3 to 14 days. The cells were harvested by trypsinization, washed once with cold phosphate-buffered saline, *p*H 7.4, and once in cold 0.25*M* sucrose, and disrupted by nitrogen cavitation [30 psi (2.0 atm), 10 minutes]. After centrifugation at 1200*g* for 5 minutes, the supernatant was centrifuged for 10 minutes at 17,000*g* to prepare a granular fraction enriched approximately threefold with respect to hexosaminidase

M. Renlund, Section on Human Biochemical Genetics, Human Genetics Branch, National Institute of Child Health and Human Development, Bethesda, MD 20892, and Children's Hospital, Helsinki University Central Hospital, 00290 Helsinki, Finland.
F. Teitze, Section on Developmental Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, Bethesda, MD 20892.
W. A. Gahl, Section on Human Biochemical Genetics, Human Genetics Branch, National Institute of Child Health and Human Development, Bethesda, MD 20892.

*To whom correspondence should be addressed at Building 10, Room 10N-318, National Institutes of Health, Bethesda, MD 20892.