The EMBO Journal vol.12 no.2 pp.501 – 511, 1993

Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters

Hong Du, Ananda L.Roy and Robert G.Roeder

Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021, USA

Communicated by L.Philipson

Earlier *in vitro* studies identified USF as a cellular factor which activates the adenovirus major late (Ad-ML) promoter by binding to an E-box motif located at position −60 with respect to the cap site. Purified USF contains 44 and 43 kDa polypeptides, and the latter was found (by cDNA cloning) to be a helix−loop−helix protein. In this report, we demonstrate a 25- to 30-fold stimulation of transcription via an upstream binding site by ectopic expression of the 43 kDa form of USF (USF^{43}) in transient transfection assays. More recent data have also revealed alternate interactions of USF^{43} at pyrimidine-rich (consensus YYAYTCYY) initiator (Inr) elements present in a variety of core promoters. In agreement with this observation, we show here that USF^{43} can recognize the initiator elements of the HIV-1 promoter, as well as those in the Ad-ML promoter, and that ectopic expression of USF^{43} can stimulate markedly the corresponding core promoters (TATA and initiator elements) when analyzed in transient co-transfection assays. Mutations in either Inr 1 or Inr 2 reduced the USF^{43}-dependent transcription activity *in vivo*. In addition, *in vitro* transcription assays showed that mutations in either or both of the Inr 1 and Inr 2 sequences of the HIV-1 and Ad-ML promoters could affect transcription efficiency, but not the position of the transcriptional start site. These results indicate that USF^{43} can stimulate transcription through initiator elements in two viral promoters, although the exact mechanism and physiological significance of this effect remain unclear.

Key words: Ad-ML and HIV-1 promoter/initiator element/transcription factor/USF

Introduction

The core promoters of genes transcribed by RNA polymerase II commonly contain either or both of two *cis*-acting DNA elements. They are the TATA element just upstream (−30 bp) of the transcriptional start site [reviewed in Breathnach and Chambon (1981)] and an initiator element encompassing the cap site (Grosschedl and Birnstiel, 1980; Smale and Baltimore, 1989). The relative importance of these two elements for promoter function varies widely among different promoters. The function of the TATA element is more clearly defined as it directly interacts with the transcription factor IID (TFIID) (Sawadogo and Roeder, 1985a; Nakajima *et al.*, 1988) to form a complex that promotes functional interactions of basic transcriptional factors and RNA polymerase II [reviewed in Roeder (1991)]. Although the precise role of the initiator element is still largely unknown, a weak consensus sequence 5′-YYCAYYYYY3′ around the cap site was identified in several eukaryotic TATA-containing promoter genes and, in some cases, mutations in this region resulted in changes in the transcription start site and/or a reduction in promoter strength (Corden *et al.*, 1980; Grosschedl and Birnstiel, 1980). In a mouse ribosomal protein-coding gene promoter lacking a well-defined TATA box, a polypyrimidine initiator sequence critically defines the position of transcription initiation and contributes to promoter efficiency (Hariharan and Perry, 1990). The SV40 late promoter, which also lacks a TATA box, requires sequences at the start site, as well as upstream (centered at −31) and downstream (centered at +28) sequences, for accurate and efficient transcription of the core promoter (Ayer and Dynan, 1988). Studies of the terminal deoxynucleotidyltransferase (TdT) promoter, which lacks an apparent TATA element, have defined an initiator element (Inr) which is necessary and sufficient for a basal level of accurate initiation both *in vitro* and *in vivo*, and influences both the efficiency and accuracy of initiation. In one study of synthetic promoters *in vitro*, the TATA and initiator elements appeared functionally similar (Smale and Baltimore, 1989). Thus, the initiator element is functionally important in at least some promoters which lack the TATA element. However, a TdT type of initiator element is also found in several TATA-containing promoters, including the adenovirus major late (Ad-ML) promoter (Smale *et al.*, 1990; Roy *et al.*, 1991), the human immunodeficiency virus type I (HIV-1) promoter (this paper) and the rabbit β-globin promoter (Dierks *et al.*, 1983). The possible functions and interplays of Inr elements in both TATA-containing and TATA-deficient promoters are not yet fully understood, and only recently has any *trans*-acting protein(s) been implicated by either physical or functional interactions with the initiator elements (Roy *et al.*, 1991; Seto *et al.*, 1991).

USF was originally identified as a human cellular factor which bound to an Ad-ML promoter upstream sequence (ML-U or E box) centered at −63 to −52 and which stimulated transcription in *in vitro* assays (Carthew *et al.*, 1985; Miyamoto *et al.*, 1985; Sawadogo and Roeder, 1985a). A full-length cDNA encoding the 43 kDa component (USF^{43}) of natural USF, which contains both 44 and 43 kDa polypeptides (Sawadogo, 1988; Sawadogo *et al.*, 1988), has been cloned in this laboratory. The predicted amino acid sequence from the cDNA showed that USF^{43} is a member of the c-myc family of regulatory proteins, and that it contains three motifs (a basic region, a helix−loop−helix region and a leucine repeat) which are required for dimerization and efficient DNA binding (Gregor *et al.*, 1990). USF is present in a variety of cell types and species, (reviewed in Sirito *et al.*, 1992). USF interacts with the upstream ML-U site (E box) in a variety of promoters, including that of human growth hormone (Peritz *et al.*, 1988), mouse metallothionein I (Carthew *et al.*, 1987), rat

γ fibrinogen (Chodosh *et al.*, 1987), duck histone H5 (Bungert *et al.*, 1992) and *Xenopus* TFIIIA genes (Hall and Taylor, 1989; Scotto *et al.*, 1989). However, the role of USF in the regulation of these promoters is not yet clear. To date, the biological activity of USF has not been demonstrated *in vivo*.

In this report, we demonstrate: (i) that USF^{43} markedly stimulates reporter genes carrying multiple ML-U sites in co-transfection assays; (ii) that USF^{43} physically interacts with alternative sites (initiator elements) in both HIV-1 and Ad-ML promoters, in agreement with previous observations (Roy *et al.*, 1991); (iii) that ectopic expression of USF^{43} can significantly stimulate the basal activity of both HIV-1 and Ad-ML core promoters (lacking ML-U sites) through Inr sites; (iv) that the HIV-1 and Ad-ML promoters each contain two initiator elements, located at −3 to +8 (Inr 1) and at +35 to +60 (Inr 2), which contribute to promoter strength both *in vivo* and *in vitro*; (v) that the USF^{43}-induced transcription from these Inr-dependent promoters starts at the normal +1 site both *in vivo* and *in vitro*. Taken together, these data suggest that USF^{43} can serve not only as a classical upstream activator *in vivo*, but also that direct or indirect interactions with the initiator elements of a variety of promoters can lead to markedly enhanced levels of basal transcription *in vivo*.

Results

Cloned USF^{43} exhibits transcriptional stimulatory activity through the ML-U site in vivo

We first asked whether the cloned human USF^{43} protein can stimulate transcription *in vivo* through the ML-U site. The USF^{43} cDNA coding region was inserted into a eukaryotic expression vector, pCX, which contains the human cytomegalovirus early promoter region (−520 to +75) fused to the rabbit β-globin splicing and poly(A) signals (Gerster *et al.*, 1990). The derived USF expression vector is designated pCXUSF (Figure 1A). Expression of exogenous USF^{43} protein in transfected cells was confirmed by immunofluorescence using a rabbit antibody raised against the bacterially expressed USF^{43} (H.Kaulen, H.Du and R.G.Roeder, unpublished data). When co-transfected into HeLa cells with reporter genes containing multiple ML-U sites (Figure 1A), pCXUSF resulted in a significant stimulation of activity that was dependent on ML-U sites. As shown in Figure 1B, USF^{43} stimulated transcription ~4-fold on the promoter containing two ML-U sites (U_2E1bCAT) and ~24-fold on the promoter containing four ML-U sites (U_4E1bCAT), relative to the endogenous USF activity. Furthermore, the extent of activation was dependent on the USF^{43} concentration, as an increase in USF^{43} expression resulted in an increase in CAT activity (data not shown). Importantly, on the adenovirus E1b core promoter employed here, which contains no known initiator element, ectopic USF^{43} had no effect on transcription in the absence of upstream ML-U sites (Figure 1B). These experiments clearly demonstrated that the cloned USF^{43} is capable of stimulating transcription through the ML-U site *in vivo*.

USF^{43} activates the HIV-1 core promoter

We next tested the ability of USF^{43} to activate an HIV-1 core promoter with two ML-U sites inserted at position −40. When co-transfected with a CAT gene-containing promoter

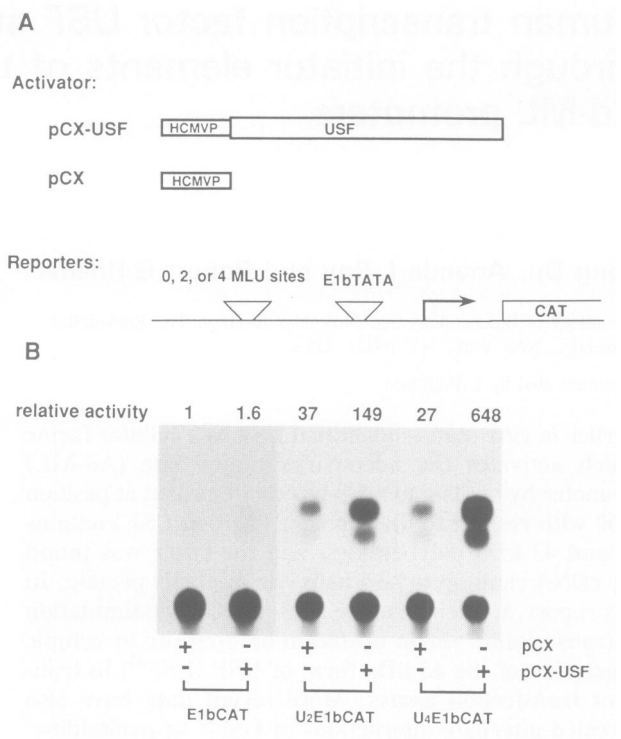


Fig. 1. Promoter activation *in vivo* by ectopic USF^{43} via ML-U sites. (**A**) Schematic diagram of the pCXUSF activator plasmid and reporter plasmids. (**B**) Ectopic USF^{43} activates transcription *in vivo* from an E1bCAT promoter with ML-U sites. HeLa cell monolayers were co-transfected with 5 μg of the reporter plasmids E1bCAT, U_2E1bCAT and U_4E1bCAT, and either 5 μg of the activator plasmid pCXUSF or 5 μg of the control plasmid pCX.

(MLU$_2$HCAT), pCXUSF resulted in a 15-fold increase in CAT activity relative to the control (Figure 2A). Again, this transcriptional stimulation was dependent on the USF^{43} concentration (Figure 2B). However, this USF^{43}-dependent stimulation was also observed for a control reporter plasmid (MLU2mHCAT) containing ML-U site mutations previously shown (H.Du and R.G.Roeder, unpublished results) to markedly reduce USF^{43} binding in mobility assays. This indicated that ectopic USF^{43}-dependent stimulation could occur independently of the ML-U sites, and this was confirmed by the analysis of two additional control reporter plasmids which had either two SP1 sites (SP1$_2$HCAT) or one NF1 site (NF1HCAT). These plasmids showed large (16- and 29-fold) responses to USF^{43} and, as with MLU$_2$HCAT, the level of transcriptional stimulation from the control reporters was dependent on the USF^{43} cDNA concentration (Figure 2C). In addition, no transcriptional stimulation was observed when these reporter plasmids were co-transfected with the control plasmid pCX (Figure 2A and C). These results strongly suggested that USF^{43} may interact directly or indirectly with the HIV-1 core promoter through a site other than the inserted ML-U sites.

Interaction of USF^{43} with the Inr 1 and Inr 2 sites of the HIV-1 and Ad-ML promoters

In order to test for an alternative binding site, we first looked for direct interactions between USF^{43} and the HIV-1 core promoter (−40 to +60). A sequence comparison of this promoter with the Ad-ML and TdT promoters revealed two regions in the HIV-1 promoter with sequence homologies to the initiator element first described in the TdT promoter

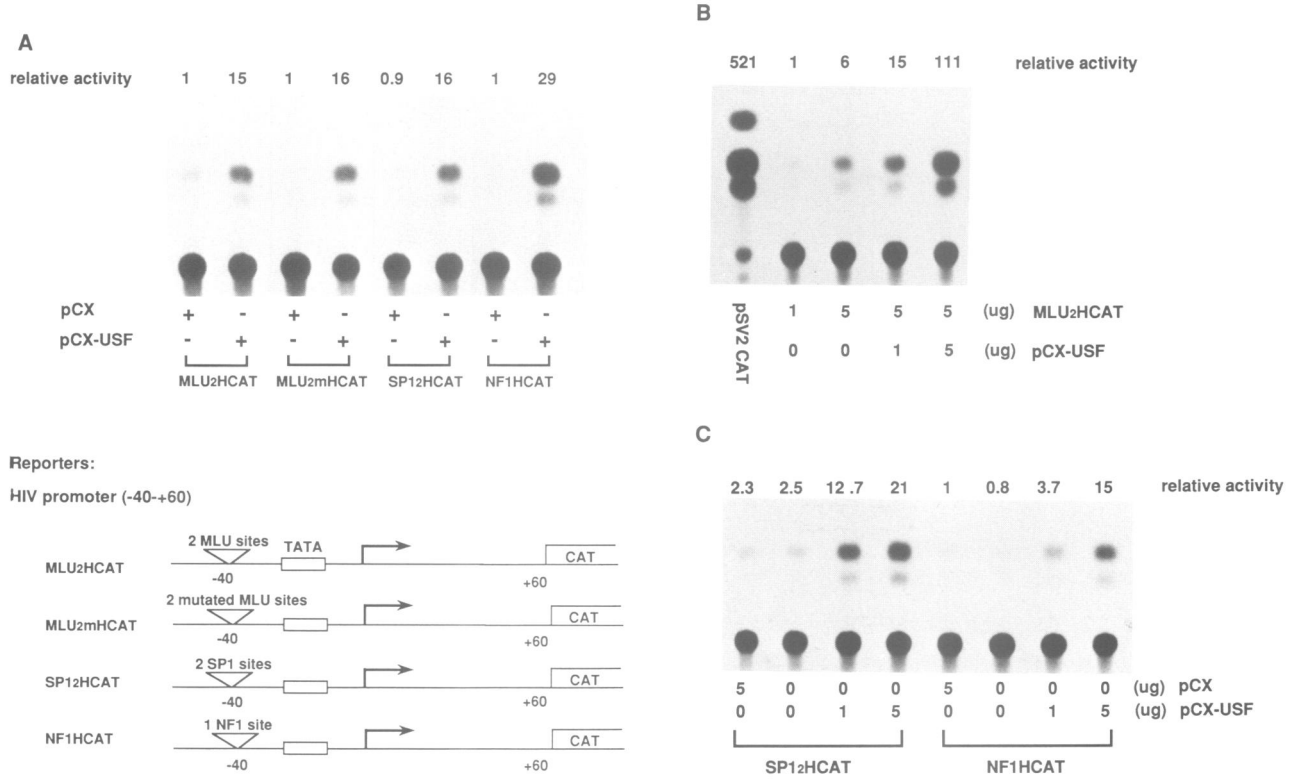


Fig. 2. *In vivo* activation of a TATA- and Inr-containing core promoter by ectopic USF[43]. (**A**) Stimulatory activity of USF[43] on the HIV-1 promoters with variable upstream elements. Five micrograms of the activator pCXUSF were co-transfected with 5 μg of either MLU_2HCAT, MLU_2mHCAT, $SP1_2HCAT$ or NF1HCAT reporter plasmids into HeLa cell monolayers. (**B**) USF[43] concentration-dependent stimulation of MLU_2HCAT transcription. (**C**) USF[43] concentration-dependent stimulation of $SP1_2HCAT$ and NF1HCAT transcription. One or 5 μg of each reporter plasmid was co-transfected with 0, 1 or 5 μg of pCXUSF into HeLa cell monolayers, as indicated pSV_2CAT was used as a positive reporter control for transfection. pCX was used as a negative control for the USF expression vector.

(also see Roy *et al.*, 1991). The first initiator site is located between positions −5 and +9, and the second initiator-like site is located between positions +29 and +42 (Figure 3D). We designated the former site as Inr 1 and the latter as Inr 2. As noted previously, two closely related sites are also found in very similar positions in the Ad-ML promoter (Roy *et al.*, 1991).

A gel retardation assay indicated that bacterially expressed and highly purified (95% homogeneity) recombinant USF[43] protein forms a specific complex with an oligonucleotide containing Ad-ML Inr 1. The protein−DNA complex was supershifted by an antiserum raised against full-length bacterially expressed USF (Figure 3A, lanes 2 and 3) and competed by unlabeled oligonucletoides containing either the ML-U (U) or the Ad-ML Inr site (I), but not by an oligonucleotide containing a mutated Inr 1 site (I_1R). Furthermore, a DNase I footprint analysis (Figure 3B) indicated that the same recombinant USF[43] protein bound to two distinct regions on the Ad-ML core promoter lacking the ML-U site. The first site (Inr 1) is located at position −3 to +20, which encompasses the previously described initiator site (Smale *et al.*, 1990). The second site (Inr 2) is located at positions +35 to +60 and bears considerable sequence similarity with Inr 1 (Figure 3D; Roy *et al.*, 1991). USF[43] also showed specific interactions with oligonucleotides encompassing either the Inr 1 or Inr 2 site in the HIV-1 core promoter in mobility shift assays (Figure 3C), as evidenced by competition of the corresponding DNA−protein complexes with either an unlabeled ML-U oligonucleotide (U) or a wild-type Ad-ML Inr 1 oligonucleotide (I), but not with a mutant Inr 1 oligonucleotide (I_1m).

USF[43] activation in vivo depends on the Inr 1 and Inr 2 sites on the HIV-1 and Ad-ML promoters

In order to assay the role of these two conserved DNA sequences in transcription, we generated a series of Inr 1 and Inr 2 mutations in both the HIV-1 and Ad-ML core promoters (Figures 4A and 5A). Mutated templates were compared to the wild-type templates both by *in vivo* transfection assays to assess responsiveness to USF[43] and by *in vitro* transcription assays to determine effects on both promoter strength and accuracy of initiation. The HIV-1 Inr mutants contained either 3 or 4 bp substitutions in Inr 1 or Inr 2, or double sets of mutations either in Inr 1 or in both Inr 1 and Inr 2. Each mutated promoter was inserted into the NF1HCAT reporter plasmid (Figure 4A). The NF1 site was chosen as the upstream activator element because NF1HCAT gave the lowest detectable basal level transcription and thus the strongest USF[43]-dependent transcription activity *in vivo* (see also Figure 2B). These constructs were then co-transfected with pCXUSF or pCX into HeLa cells and assayed for relative CAT activities. As shown in Figure 4A (summary) and B (primary data), the basal activity ranged from 65 to 85% of the level observed with the wild-type reporter. However, whereas USF[43] stimulated the wild-type HIV-1 core promoter 18-fold, relative to the level observed with the control plasmid pCX, most of the mutations in the Inr 1 site reduced the level of stimulation by USF[43]. Although the $NF1I_1a$ and $NF1I_1b$ mutations effected only

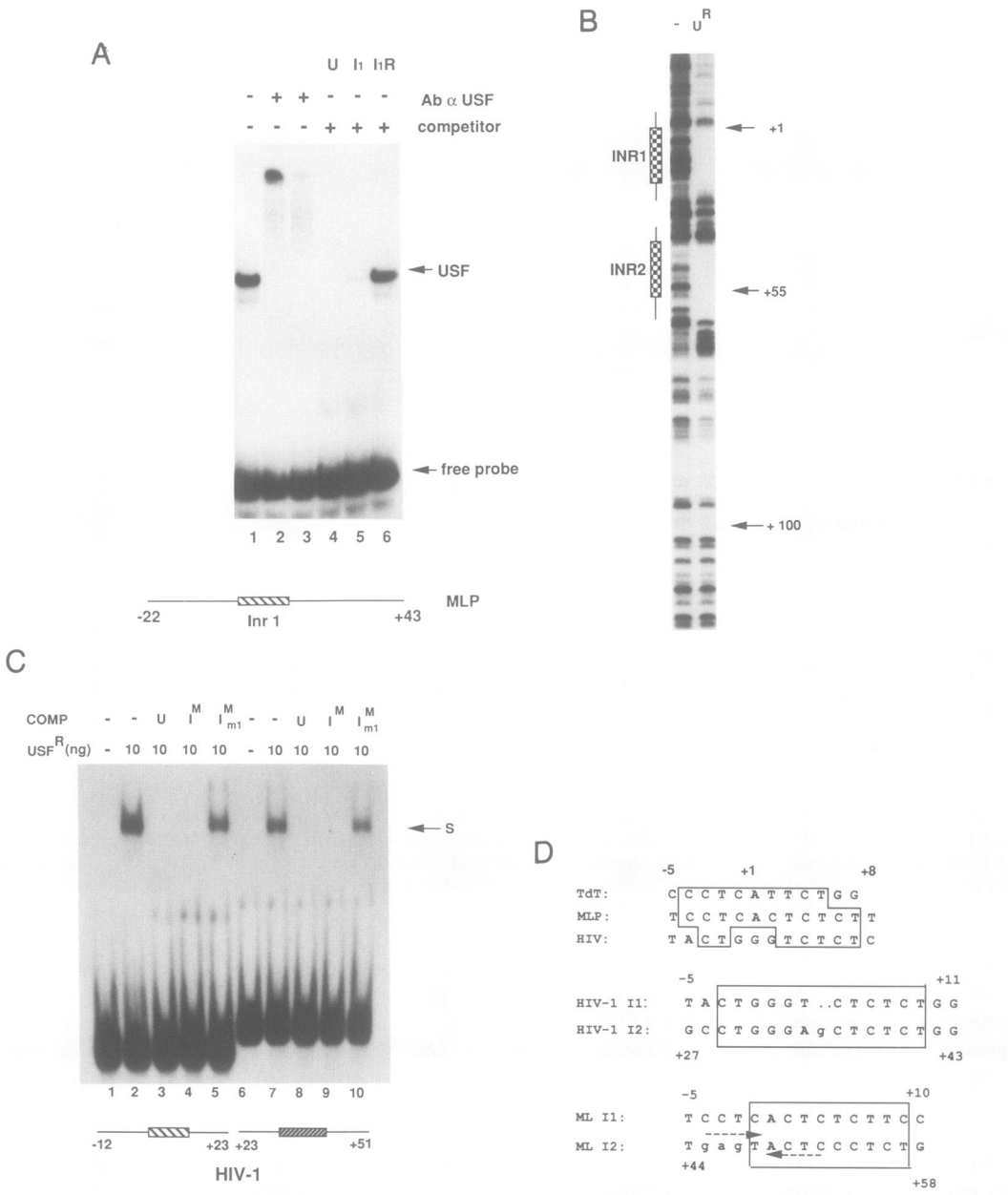


Fig. 3. Recombinant USF[43] binds to the Inr elements of the Ad-MLP and HIV-1 promoter. (**A**) Mobility shift assays of recombinant USF[43] binding at the Inr 1 sites of the Ad-ML promoter. Ten nanograms of highly purified (to 95%) bacterial recombinant USF[43] protein were incubated with 0.1−0.2 ng of radiolabeled probe corresponding to Inr 1 of Ad-ML promoter (−22/+43) at 30°C for 20 min. The DNA−protein complex was separated from the free probe by 5% non-denaturing polyacrylamide gels. Two antisera (1:100 dilution) against full-length bacterially expressed USF were added in the reaction in lanes 2 and 3. The DNA competitors were: U, the oligonucleotide corresponding to ML-U site; I, the wild-type oligonucleotide corresponding to the Inr 1 site of the Ad-ML core promoter; I_1R, the mutant oligonucleotide corresponding to the Inr 1 site of the Ad-ML core promoter. (**B**) DNase I footprint analysis of recombinant USF[43] binding at the Inr 1 and Inr 2 sites on the Ad-ML promoter. The assay procedure is described in Materials and methods. The wild-type Ad-ML promoter was incubated with no protein (−) or with highly purified (to 95%) recombinant USF[43] (U^R). (**C**) Mobility shift assays of recombinant USF[43] binding at the Inr 1 and Inr 2 sites on the HIV-1 promoter. Ten nanograms of highly purified (to 95%) bacterial recombinant USF[43] protein were incubated with 0.1−0.2 ng of radiolabeled probe corresponding to either Inr 1 or Inr 2 sites of the HIV-1 promoter, as indicated, at 30°C for 20 min. The DNA−protein complex was separated from the free probe by 4% non-denaturing polyacrylamide gels. The DNA competitors were the same as indicated in Figure 3A. (**D**) Sequence comparisons of Inr sites among the TdT, MLP and HIV-1 promoters.

minor reductions (~2-fold) in the level of stimulation by USF[43], a combination of these mutations ($NF1I_1c$) produced a more significant reduction (~4-fold). The $NF1I_1d$ and $NF1I_1e$ mutations dramatically reduced (4- to 7-fold) the level of stimulation by USF[43], whereas mutant $NF1I_1f$ had almost no effect. A limited set of mutations in Inr 2 ($NF1I_2$) also did not affect the level of stimulation by USF[43]. However, in combination with Inr 1 mutations, the Inr 2 mutation effected a several-fold reduction in the level of stimulation by USF[43], which in all cases but one ($NF1I_1d-I_2$) exceeded the inhibitory effects of either mutation alone. This is best illustrated by the results for the $NF1I_1a-I_2$, $NF1I_1b-I_2$, $NF1I_1c-I_2$ and $NF1I_1e-I_2$ reporters. A more modest reduction in USF[43] response was seen with $NF1I_1f-I_2$, while essentially no effect was observed with either $NF1I_1f$ or $NF1I_2$ alone. The modest

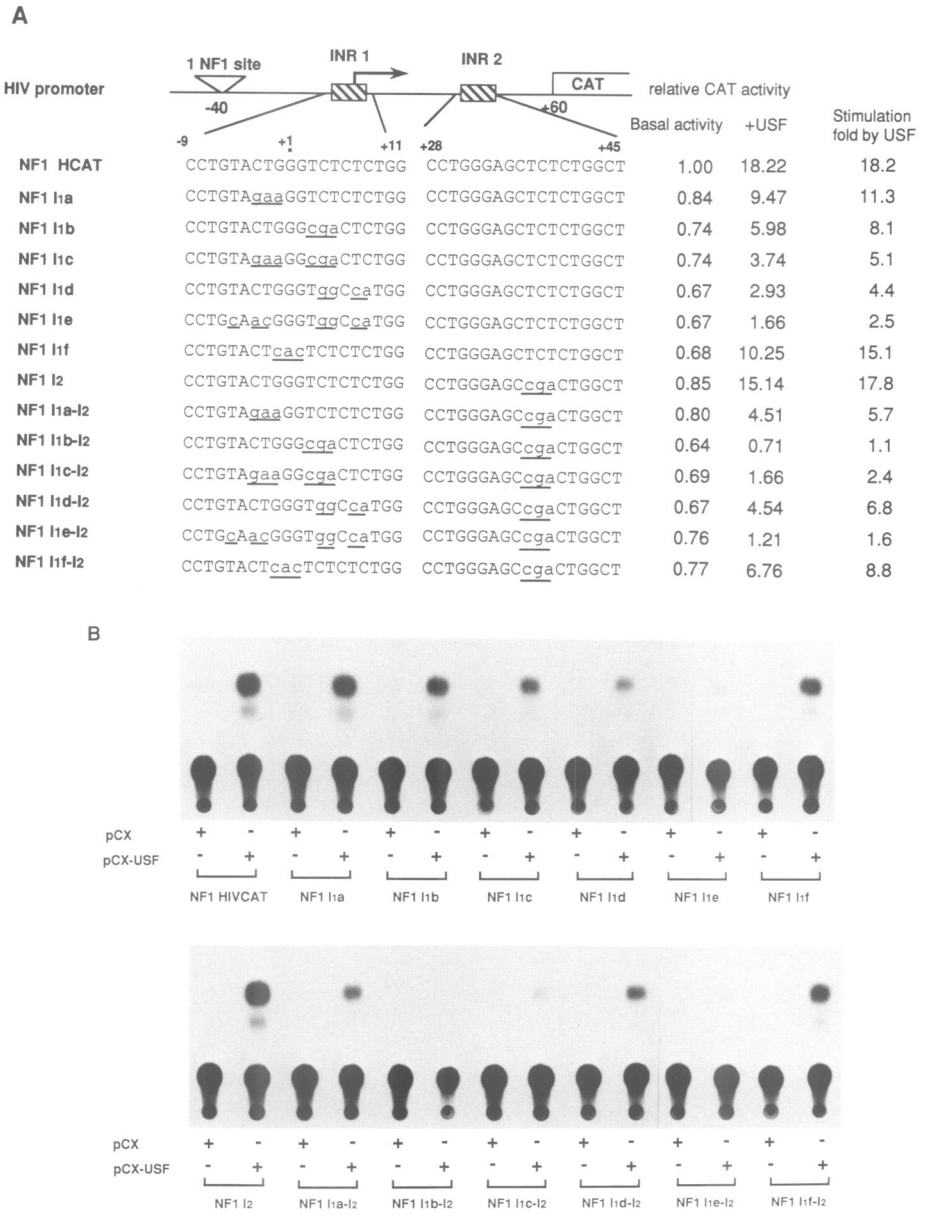


Fig. 4. Effect of Inr mutations on USF[43]-dependent transcription from the HIV-1 promoter *in vivo*. (**A**) Diagram of the mutations at either Inr 1 or Inr 2, or at both sites on the HIV-1 core promoter and summary of the stimulatory effects of USF[43]. (**B**) CAT assays. Five micrograms of pCX or pCXUSF were co-transfected with 5 μg of the wild-type or mutant HIV-1 promoters (as indicated) into HeLa cell monolayers. The relative CAT activities are indicated on the top.

positive effect of the double mutation in $NF1I_1d-I_2$, relative to the activity of $NF1_1d$ or $NF1I_2$, is unexplained. Overall, these results indicate: (i) that the downstream sequence TCTCTC (+3 to +8) in Inr 1 is more important than the sequence just upstream of the initiation site; (ii) that sequences in the Inr 2 are less important (when analyzed independently) than sequences in Inr 1, but that they may act synergistically with the Inr 1 sequence; and (iii) that the initiation site sequence GGG (−1 to +2) in the HIV promoter is not critical for basal or USF[43]-dependent activity. The latter may reflect conversion of the unusual HIV-1 Inr 1 site (containing three G residues at the initiation site) to a site that more closely matches the consensus for a TdT-like initiator (Smale *et al.*, 1989).

Since the Ad-ML promoter contains similar Inr 1 and Inr 2 elements, we next analyzed the Ad-ML core promoter (missing the upstream high-affinity USF site) for USF[43]-dependent activation. A series of test plasmids were constructed which contained sequences from −45 to +65, with either wild-type or altered sequences in the Inr elements, fused to the CAT reporter gene (mutated sequences are shown in Figure 5A). Co-transfection of pCXUSF and the wild-type Ad-ML core promoter into HeLa cells resulted in an 11-fold stimulation of activity by USF[43]. Although the Inr mutations had virtually no negative effects on basal activity, the mutated templates did show reduced levels of USF[43]-dependent transcription. The exception was MLI_1L (mutations in the pyrimidine tract just upstream of the initiation site), which showed a significant (unexplained) increase in both basal and USF-dependent activity. On the other hand, the I_1R mutations (in the downstream pyrimidine cluster of Inr 1) significantly reduced stimulation

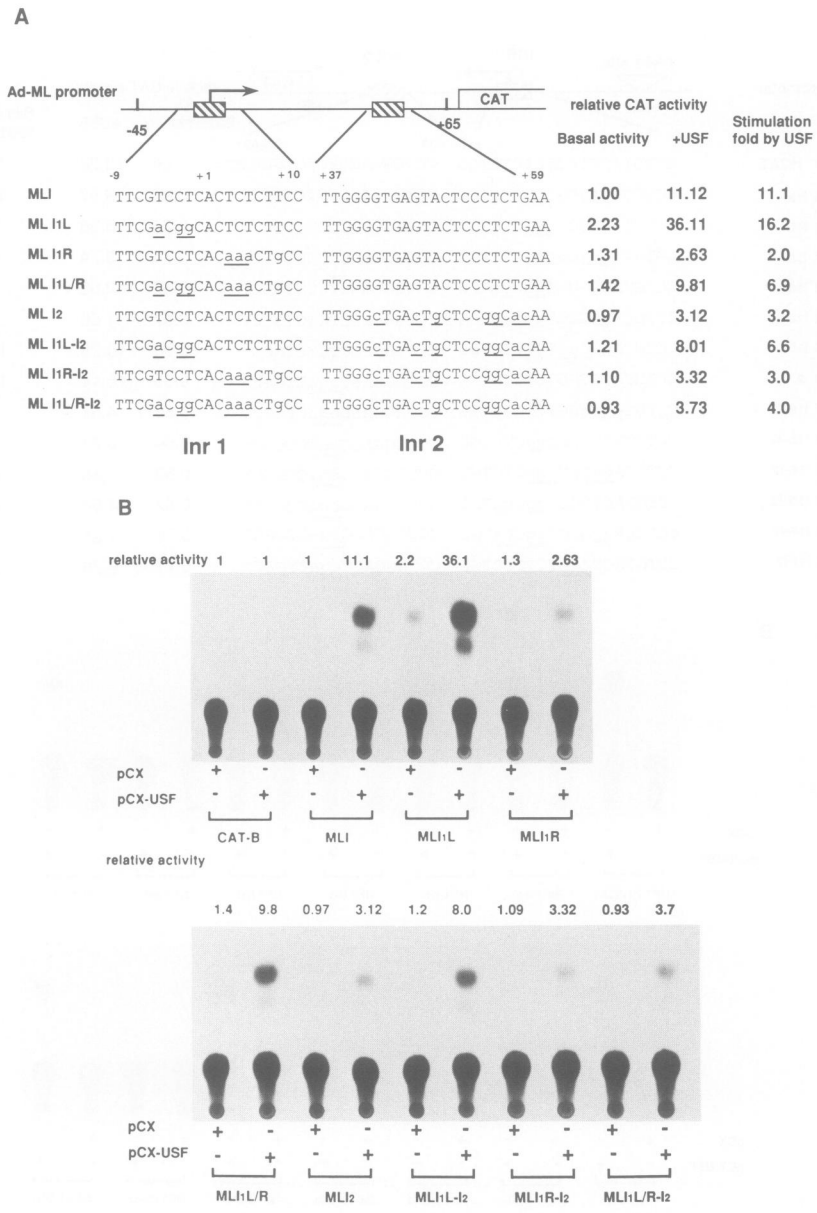


Fig. 5. Effect of Inr mutations on USF[43]-dependent transcription from the Ad-ML core promoter *in vivo*. (**A**) Diagram of the mutations at Inr 1 or Inr 2 sites, or at both sites on the Ad-ML core promoter and summary of the stimulatory effect of USF[43]. (**B**) CAT assays. Five micrograms of pCX or pCXUSF were co-transfected with 5 μg of the wild-type or mutant Ad-ML promoters (as indicated) into HeLa cell monolayers. The relative CAT activities are indicated on the top.

by USF[43] both in the wild-type background (>5-fold) and in the MLI_1L background. Similarly, I_2 mutations (throughout Inr 2) markedly reduced (3- to 4-fold) the level of stimulation in the wild-type and I_1L backgrounds, but had less effect in the I_1L/R background. These results further support the notion that the Inr 1 and Inr 2 sites mediate the USF[43] stimulatory effects when the ML-U site is absent in the HIV-1 or Ad-ML promoters.

Mapping the transcription start site of USF[43]-induced Ad-ML core promoter in vivo

The above studies show an *in vivo* activation by USF[43] that requires intact Inr elements. In order to show that this activation occurs at the level of transcription and from the normal +1 initiation site, and that the Inr mutations which block the activation are not simply eliminating all functional interactions of the general transcriptional machinery, primer extension assays were employed to map the RNA start site on wild-type and mutated Ad-ML promoter templates *in vivo* in the presence and absence of ectopic USF[43]. In order to obtain detectable levels of transcripts in the absence of USF[43] and with mutated templates, an SV40 enhancer was placed downstream of the CAT reporter gene and the derived plasmids (MLIE and MLI_1R-I_2E) were transfected into a cell line (T22) reported to have a high transfection efficiency (Figure 6). Under these conditions, as determined by primer extension analyses, USF[43] expression effected a 7-fold activation of the wild-type template (lane 3 versus lane 2), but only a 1.5- to 2-fold activation of an Inr 1/Inr 2 mutated template (lane 5 versus lane 4). Importantly, in each case the major transcription initiation site was at the normal +1 site within Inr 1. Interestingly, however, the wild-type and

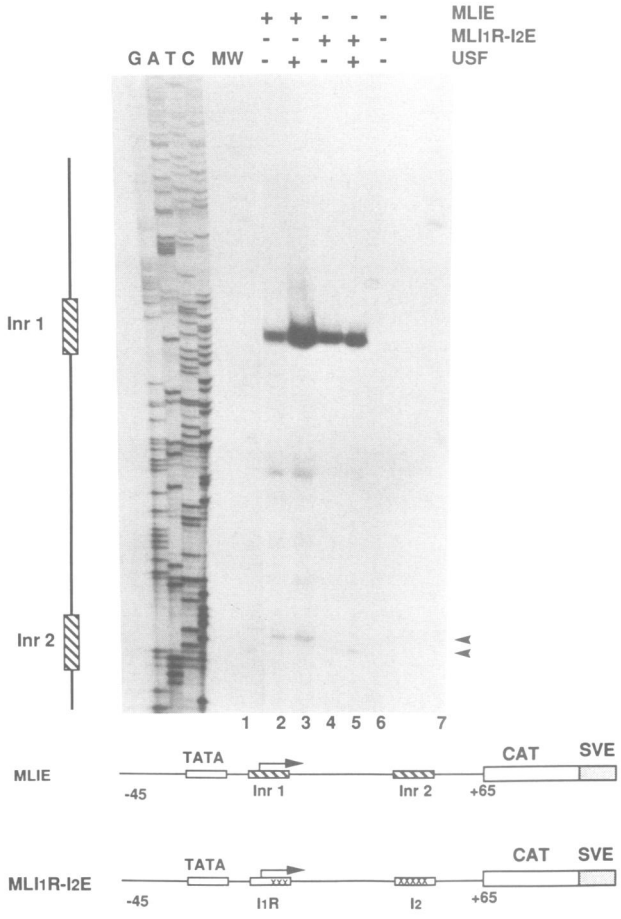


Fig. 6. Mapping the transcription start site of USF[43]-induced Ad-MLP *in vivo*. Six micrograms of mRNA from co-transfected T-22 cells were isolated and assayed by primer extension as described in Materials and methods. The wild-type template of the Ad-ML promoter was sequenced by the dideoxynucleotide method with the same CAT primer used in the primer extension, and shown on lane 'G' 'A' 'T' 'C'. ^{32}P end-labeled pBR322/*Hae*III fragments were used as molecular weight markers and are shown on **lanes 1** and **7**. Signals of mRNA samples from mock transfection (**lane 6**), co-transfection of pCX with wild-type (MLIE, **lane 2**) or mutant template (MLI_1R-I_2E, **lane 4**), co-transfection of pCXUSF with wild-type (MLIE, **lane 3**) or mutant template (MLI_2R-I_2E, **lane 5**) are shown as indicated.

mutated templates appeared to generate, in addition to the major +1 initiated transcript, several minor transcripts; one of these (arrow in Figure 6) mapped to the Inr 2 region (+55 for wild-type and +57 for the mutated template). Primer extension analyses clearly showed that the USF[43]-dependent activation of the MLI template (without an enhancer) in transfected HeLa cells also resulted from an enhanced level (9-fold) of +1 initiated transcripts, but the low level of transcripts from cells transfected with the corresponding mutant (MLI_1R-I_2) template or with the wild-type (MLI) template in the absence of USF[43] precluded firm conclusions about the start site in these cells (data not shown). The *in vivo* mRNA mapping studies provide evidence indicating: (i) that USF[43] activation via the Inr elements occurs at the transcription level from initiation events at the major +1 start site; (ii) that Inr mutations do not affect the accuracy of transcription initiation (including recognition by the basal transcriptional machinery) on these TATA-containing promoters; (iii) that the Inr 2 element can serve as a weak initiation start site in the absence of an intact Inr 1.

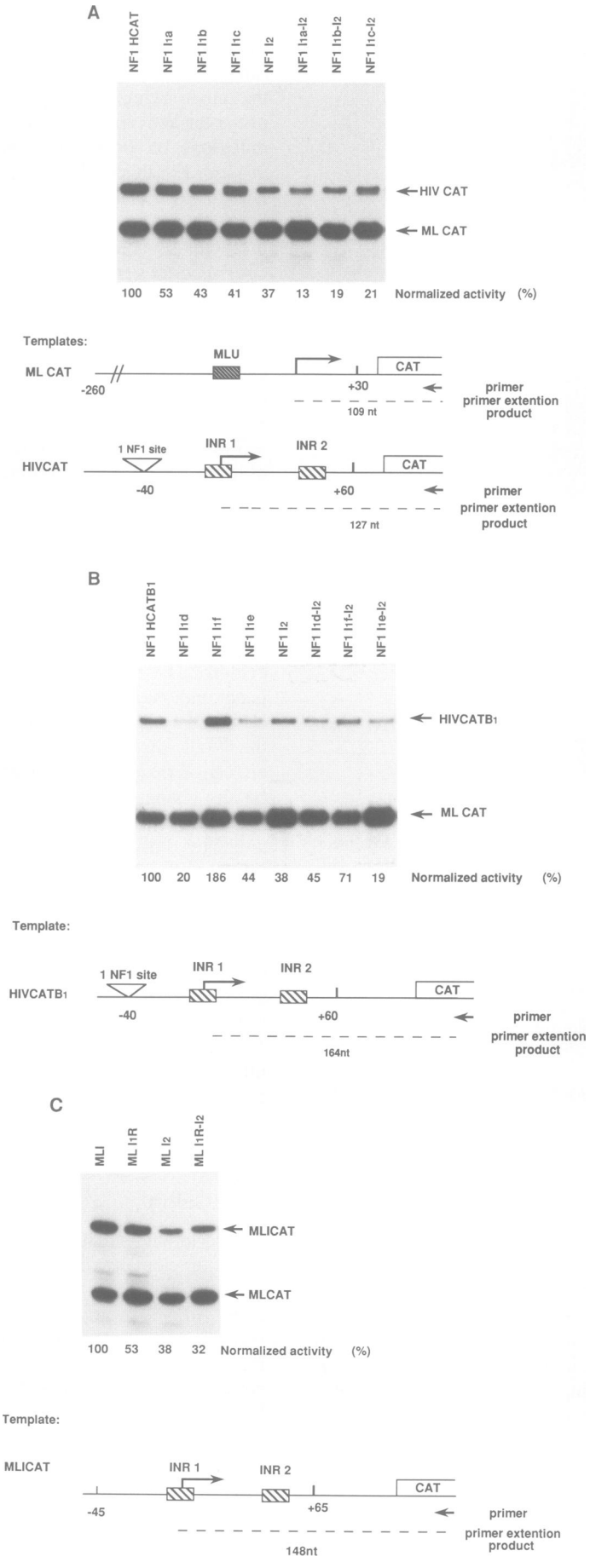
The Inr 1 and Inr 2 sites are important for in vitro transcription activity of the HIV-1 and Ad-ML promoters

Since the *in vivo* transcription assays clearly indicated that the initiator elements, especially Inr 1, are crucial for core promoter activation by USF[43], we tested the effects of Inr mutations on promoter strength and initiation accuracy *in vitro*. For this analysis, we used unfractionated HeLa cell nuclear extracts and various HIV-1 templates with mutations in the Inr 1 or Inr 2 sites. MLCAT (containing Ad-ML promoter sequences from −260 to +30 connected to the CAT reporter gene) was used as the internal control template. As shown in Figure 7A and B, mutations either in Inr 1 or Inr 2 or in both sites of the HIV-1 promoter clearly decreased *in vitro* transcription activity. Compared to the transcriptional activity of the wild-type HIV-1 promoter, mutants $NF1I_1a$ (−3 to −1), $NF1I_1b$ (+3 to +5), $NF1I_1c$ (−3 to −1 and +3 to +5), $NF1I_1d$ (+4 to +8), $NF1I_1e$ (−2 to −5 and +3 to +8) and $NF1I_2$ (+37 and +29) had only 20−60% of the wild-type transcriptional activities. The combined Inr 1 and Inr 2 mutants $NF1I_1a-I_2$, $NF1I_1b-I_2$, $NF1I_1c-I_2$ and $NF1I_1e-I_2$ showed even more severe effects with activities that were <20% of the control (wild-type template) value. Although the $NF1I_1f-I_2$ template showed a comparatively high activity (71%) relative to the wild-type template, an effect of the I_2 mutation on the abnormally high activity of the $NF1I_1f$ template (186% of the wild-type) was still evident. The relatively high activity of the double mutant is not understood, but mirrors the abnormally high activity observed *in vivo*. Except for the latter case, the effects of the double mutants were multiplicative. Despite the drastic effects of many of the mutations, there did not appear to be any changes in the transcription start sites.

For the Ad-ML promoter, we chose as representative templates those which showed the most dramatic effects on USF[43]-dependent *in vivo* transcription assays. As shown in Figure 7C, the Inr 1 mutant MLI_1R, the Inr 2 mutant MLI_2 and the double mutant MLI_2R-I_2 (Figure 5A) had ~30−50% of the wild-type activity. Again no different transcription initiation sites were observed in these assays, in agreement with the result of the *in vivo* assays (Figure 6). Thus, these mutational studies have suggested that the Inr 1 and Inr 2 sites on both HIV-1 and Ad-ML promoters are important *cis*-elements for achieving the full range of transcriptional activity *in vitro*. These results also agree with the observations from *in vivo* assays that intact Inr 1 and Inr 2 sequences are required for full promoter activity.

Discussion

Transcription initiation is regulated not only by core promoter elements and associated factors, but also by distal elements and associated factors. An example of the latter is the Ad-ML promoter ML-U site (E box motif) which is recognized by the cellular factor USF. The USF/ML-U site interactions are of general relevance since the 43 kDa component of USF (USF[43]) is a member of the helix−loop−helix family of regulatory proteins which also recognize E-box motifs. A direct role for USF[43] in stimulating promoter function through the ML-U site, in the apparent absence of the 44 kDa USF component, was provided by earlier *in vitro* studies with recombinant USF[43] and USF-depleted extracts (Pognonec and Roeder, 1991). In the present analyses, *in vivo* transfection assays clearly


A
NF1 HCAT
NF1 I1a
NF1 I1b
NF1 I1c
NF1 I2
NF1 I1a-I2
NF1 I1b-I2
NF1 I1c-I2
HIV CAT
ML CAT
100 53 43 41 37 13 19 21 Normalized activity (%)
Templates:
ML CAT
MLU
CAT
-260
+30
109 nt
primer
primer extention product
HIVCAT
1 NF1 site
INR 1
INR 2
CAT
-40
+60
127 nt
primer
primer extention product
B
NF1 HCATB1
NF1 I1d
NF1 I1f
NF1 I1e
NF1 I2
NF1 I1d-I2
NF1 I1f-I2
NF1 I1e-I2
HIVCATB1
ML CAT
100 20 186 44 38 45 71 19 Normalized activity (%)
Template:
HIVCATB1
1 NF1 site
INR 1
INR 2
CAT
-40
+60
164nt
primer
primer extention product
C
MLI
ML I1R
ML I2
ML I1R-I2
MLICAT
MLCAT
100 53 38 32 Normalized activity (%)
Template:
MLICAT
INR 1
INR 2
CAT
-45
+65
148nt
primer
primer extention product

showed that ectopic expression of USF^{43} stimulated transcription of reporter plasmids (U_2E1bCAT and U_4-E1bCAT) containing ML-U sites, but lacking known initiator elements. USF^{43}-dependent activation of these promoters was dependent on ML-U sites and showed a USF^{43} dosage response (data not shown). These results demonstrate, for the first time, that the cloned USF^{43} is biologically active *in vivo*, similar to the natural USF, and indicate its ability to function through the ML-U site.

Quite unexpectedly, USF^{43} also stimulated HIV-1 core promoter transcription regardless of the presence of an ML-U site. The implication of this result is that an alternative core promoter sequence, distinct from the ML-U (E-box) site, is responsible for this stimulatory activity. In this report, we showed by gel shift and DNase I footprint analysis that USF^{43} binds to two regions on the Ad-ML core promoter: one covering a proximal region (−3 to +20) that contains a previously recognized initiator element (Inr 1) and a second (+35 to +60) containing a region (designated Inr 2) that is similar in sequence to Inr 1. Analogous regions with significant sequence similarities are also apparent in the HIV-1 promoter (Figure 3D) and oligonucleotides containing these two regions formed specific complexes with USF^{43} in electrophoretic mobility shift assays, just as they did with oligonucleotides containing Ad-ML Inr elements. These observations are consistent with our previous findings that USF^{43} can interact with the Inr element in the Ad-ML promoter (Roy *et al.*, 1991). Further confirmation of functional interactions (leading to transcription enhancement) was provided by *in vivo* analyses which showed that the activation by ectopic USF^{43}, of promoters lacking an ML-U site, required intact initiator elements on both HIV-1 and Ad-ML promoters. The activation by USF^{43} through Inr elements was shown to occur at the transcription level and via initiation at the normal +1 site. Furthermore, using a chimeric DNA template (pE1bCAT) which lacks the Inr 1 and Inr 2 sites and contains only a TATA element, no stimulation by ectopic USF^{43} was observed. Therefore, activation by USF^{43} of the HIV-1 and Ad-ML core promoters is correlated with the presence of initiator elements.

These observations imply that USF^{43} can stimulate transcription initiation through two entirely different regions in a sequence-specific manner: (i) USF^{43} preferentially binds to the ML-U site (CACGTG) and can activate the basic transcriptional machinery when the Inr 1 and Inr 2 sites are absent (Figure 1); (ii) USF^{43} can interact functionally with Inr 1 and Inr 2 sites (consensus YYAYYTCYY), even in the absence of ML-U sites, to stimulate basal transcription (Figure 2). This provides an intriguing example of a single transcription factor that can interact functionally with different promoter elements and stimulate transcription in distinct ways under different circumstances.

Exactly how USF^{43} interacts with and stimulates transcription through the Inr 1 and Inr 2 sites in the cell is still unclear. USF^{43} can bind directly to the Inr 1 and Inr 2 sites *in vitro*, and mutations in these sites in the HIV-1 and Ad-ML promoters show parallel effects on USF^{43} binding and promoter activity (data not shown). However, the binding affinity for the Inr 1 and Inr 2 sites is much lower (~20-fold) than that for the ML-U site (Roy *et al.*, 1991). Therefore, it is not clear whether transcriptional stimulation by USF^{43} *in vivo* reflects direct and independent binding of USF^{43} to the Inr 1 and Inr 2 sites, cooperative interactions with another factor(s) at these sites or even an indirect effect. The existence of another factor(s) is implied by several findings. The first is the observation that the USF^{43} cDNA concentration-dependent stimulation of activity *in vivo* is followed by an inhibition of activity at high cDNA levels (H.Du and R.G.Roeder, unpublished observation), suggesting that excessive amounts of ectopically expressed USF^{43} may titrate (squelch) other factors important for transcription initiation. The second is based on *in vitro* transcription studies with USF-depleted nuclear extracts. These experiments show a several-fold reduction of Inr-dependent transcription that is only partially restored by recombinant USF^{43}, suggesting that USF depletion may remove an additional factor(s) involved in the USF^{43} effect (H.Du and R.G.Roeder, unpublished observations).

One factor that might be involved in USF^{43} function is the 44 kDa polypeptide which is associated with USF^{43} in natural USF (Sawadogo *et al.*, 1988) and apparently immunoprecipitated with USF^{43} (Pognonec and Roeder, 1991). Another good candidate for a factor interacting with USF^{43} and with Inr elements *in vivo* is the recently identified TFII-I. TFII-I is immunologically related to USF^{43} and binds specifically to the Inr 1 and Inr 2 sites on both Ad-ML and HIV-1 promoters, as well as to the ML-U site (Roy *et al.*, 1991). Furthermore, *in vitro* transcription assays with purified factors have shown that TFII-I stimulates core promoter activity in a fashion dependent on binding to the Inr sequence, while DNA binding assays have shown highly cooperative interactions between USF^{43} and TFII-I at both ML-U and Inr sites (Roy *et al.*, 1991). These results suggest that TFII-I could be involved in modulation of the USF^{43} binding and transcriptional activities at either ML-U or Inr sites when present individually or together on the same template. The observed cooperativity between TFII-I and USF^{43} at Inr sites might also explain the apparent lack of ectopic USF^{43} function through ML-U sites on the HIV promoter, assuming that such interactions maximally stimulate the general transcription machinery. It is also possible that the ML-U site mutations studied here (Figure 2A), while sufficient to reduce USF binding *in vitro*, nevertheless allow USF interactions (perhaps cooperative) *in vivo*. The structural and functional relationships between USF^{43} and TFII-I pose interesting and important questions for further study.

The protein sequence of USF^{43} revealed that it is a member of a family of proteins which contain adjacent basic (BR), helix−loop−helix (HLH) and, in some cases, leucine repeat (LR) regions (Gregor *et al.*, 1990). Recent studies have revealed >20 members of this family, which includes proto-oncogenes, enhancer-binding proteins and cell

Fig. 7. Effect of Inr mutations on transcription from HIV-1 and Ad-ML core promoters *in vitro*. Mutations are indicated in Figures 4A and 5A. Each reaction contained 250ng (**A**) or 100 ng (**B**) of wild-type or mutant HIV-1 template, or 100 ng (**C**) of wild-type or mutant Ad-MLP template and 100 ng (A, B) or 50 ng (C) of MLCAT internal control template. The transcribed RNAs were analyzed as described in Materials and methods. The lengths of the primer extension products are indicated in the diagrams. After gel electrophoresis, each specific band was excised and radioactivity was quantitated by liquid scintillation counting.

determination and differentiation factors (Blackwood and Eisenman, 1991; reviewed in Visvader and Begley, 1991; Weintraub *et al.*, 1991). An interesting feature of HLH proteins is their capacity to form either homodimeric or heterodimeric complexes with certain other members of the family, thereby modulating their DNA-binding capacity and/or their functional specificity (Murre *et al.*, 1989; Benezra *et al.*, 1990; Blackwood and Eisenman, 1991; Weintraub *et al.*, 1991). Therefore, it is not unlikely that other HLH proteins might interact with USF[43] to regulate either positively or negatively both DNA binding and *trans*-activation activity through either the ML-U site or Inr elements.

As indicated, USF[43] can stimulate transcription through two different elements: a typical upstream element (ML-U) and more recently described Inr elements on the HIV-1 and Ad-ML promoters. The HIV-1 core promoter also contains other control elements in addition to Inr elements. Overlapping Inr 1 (−5 to +9) are three sites (−4 to +1, +7 to +11 and +17 to +22) which bind a cellular factor (designated LBP-1) that, at high concentrations, negatively regulates the HIV-1 promoter transcription activity both *in vivo* and *in vitro* (Kato *et al.*, 1991). Overlapping Inr 2 is the TAR element (+19 to +42) which mediates (via the derived RNA) activation by the viral-coded TAT protein (Jakobovits *et al.*, 1988; Berkhout *et al.*, 1989; Garcia *et al.*, 1989). These observations raise the possibility that Inr 1- and Inr 2-bound factors such as USF[43] and/or TFII-I may modulate the function of other elements. In addition, it has been shown that there is an E-box sequence that binds USF at the HIV-1 promoter upstream region (−167 to −175) (Sawadogo and Roeder, 1985b), but whether USF[43] or another protein interacts with the E-box to regulate the HIV promoter activity remains to be determined. Thus, the HIV-1 promoter provides an interesting model system to study functional interactions of both positive factors (USF[43] and TFII-I) and negative factors (LBP-1) on the transcription initiation region (Inr 1, Inr 2 and LBP-1 sites).

Both the HIV-1 and the Ad-ML promoters have two initiator elements, one located at the transcription initiation region from −3 to +8, and another located at a downstream region extending from +38 to +55 (Figure 3D). Other promoters, including the mouse TdT gene and rabbit β-globin gene, also contain a second Inr element downstream of the +1 site (A.L.Roy and R.G.Roeder, unpublished observations). Our mutation analyses of the HIV-1 and Ad-ML promoters indicate that Inr 2 is an important functional element for transcription initiation. This is further confirmed by the observation that a mRNA transcript actually starts at the Inr 2 site. It is unlikely that this transcript is caused by reverse transcriptase pausing, since the transcription start sites at Inr 2 were constantly observed in both HeLa (data not shown) and T-22 (Figure 6) cell lines for the Ad-ML promoter, as well as for the HIV-1 promoter (data not shown). Even though the level of transcripts starting at Inr 2 is much less than that starting at Inr 1, it could be an important compensatory system. So far, gel shift and footprint analyses have identified two *trans*-acting factors, USF[43] (this study) and TFII-I (Roy *et al.*, 1991), that bind to the Inr 2 site. Thus, we believe that the involvement of the Inr 2 site in conjunction with the Inr 1 site in transcription initiation is of potential importance for certain promoter activation mechanisms.

Materials and methods

Plasmid construction

All plasmids were constructed and manipulated using standard techniques (Sambrook *et al.*, 1989). All DNA sequences and orientations were confirmed by dideoxy chain termination using Sequenase (United States Biochemicals). The plasmids U_2E1bCAT and U_4E1bCAT were generated by the insertion of two or four copies of the synthetic ML-U site at the *Sal*I site of pE1bCAT (a gift from J.W.Lillie and M.R.Green). The plasmids MLU_2-HCAT, $SP1_2$HCAT and NF1HCAT contain the USF, SP1 and NF1 binding site oligonucleotides at the −40 site of the HIV-1 promoter of the pHCAT plasmid, respectively (H.Kato and R.G.Roeder, unpublished data). The plasmid MLU_2mHCAT is the same as plasmid MLU_2HCAT except that the triple point mutations were introduced into the ML-U sites (CACGTG to CCCGAT).

Templates with mutations at the Inr 1 and Inr 2 sites of the HIV-1 promoter were made by ligation of synthetic oligonucleotides containing different regions (−40 to −13, −12 to +19, and +20 to +60) of the HIV-1 promoter with either the *XhoI*/*Bgl*II fragment of pHCAT ($NF1I_1a$, $NF1I_b$ and $NF1I_c$), the *XhoI*/*Hin*dIII fragment of pHCAT ($NF1I_2$, $NF1I_1a-I_2$, $NF1I_1b-I_2$ and $NF1I_1c-I_2$), the *XhoI*/*Bgl*II fragment of pHCATB1 ($NF1I_1d-I_2$, $NF1I_1e-I_2$ and $NF1I_1f-I_2$), or the *XhoI*/*Hin*dIII fragment of pHCATB1 ($NF1I_2$, $NF1I_1d-I_2$, $NF1I_1e-I_2$ and $NF1I_1f-I_2$). The mutation sites are illustrated in Figure 5A. The templates with mutations at the Inr 1 and Inr 2 site of the Ad-ML core promoter were made by ligation of synthetic oligonucleotides containing different regions (−45 to −16, −15 to +23, and +24 to +65) of the Ad-ML promoter with the *Sal*I/*Xba*I fragment of pCAT (MLI, MLI_1L, MLI_1R, MLI_1L/R, MLI_2, MLI_1-I_2, MLI_1R-I_2 and $MLI_1L/R-I_2$) or with the *Sal*I/*Xba*I fragment of pCATE (MLIE, MLI_1R-I_2E). The mutation sites are illustrated in Figure 5A. The oligonucleotides for both strands of the HIV-1 and Ad-ML promoter were synthesized by the protein sequence facility at The Rockefeller University, as shown below.

ML-U:
t c g a GGCCACGTGACCGGGTGTTCCTGGCCACGTGACCG
CCGGTGCACTGGCCCACAAGGACCGGTGCACTGGC a g c t
ML-Um:
t c g a GGCCCCGATACCGGGTGTTCCTGGCCCCGATACCG
CCGGGGCTATGGCCCACAAGGACCGGGGCTATGGC a g c t
NF_1 and HIV-1 TATA element:
t g c a GTTGGCACTGTGCCAACTCAGATGCTGCATATAAGCAGCTGCTTT
CAACCGTGACACGGTTGAGTCTACGACGTATATTCGTCGACGAAAAACGG
HIV-1 Inr 1 (−12 to +19):
TTGCCTGTACTGGGTCTCTCTGGTTAGACCA
ACATGACCCAGAGAGACCAATCTGGT c t a g
HIV-1 Inr 2: (+20 to 60):
GATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCA
ACTCGGACCCTCGAGAGACCGATTGATCCCTTGGGTTCGA
Ad-MLP TATA element: (−45 to −16):
t c g a CTCCTGAAGGGGGGCTATAAAAGGGGGTGGG
GAGGACTTCCCCCCGATATTTTCCCCCACCCCCGCG
Ad-ML Inr 1 (−15 to +23):
GGCGCGTTCGTCCTCACTCTCTTCCGCATCGCTGTCTG
CAAGCAGGAGTGAGAGAAGGCGTAGCGACAGACGCTCCC
Ad-ML Inr 2 (+24 to +65):
CGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAAAAGCGGT
GGTCGACAACCCCACTCATGAGGGAGACTTTTCGCCA g a t c

The mutant oligonucleotides are shown in Figures 4A and 5A.

The pCXUSF expression vector was constructed by insertion of the full-length USF cDNA (Gregor *et al.*, 1990) into the pCX expression parent vector at the *Hin*dIII and *Xho*I sites. The pCX vector contains the human cytomegalovirus enhancer/promoter, the β-globin gene splicing signal and the polyadenylation signal (Boshart *et al.*, 1985).

In vivo transfection, CAT assays and mRNA mapping

HeLa cells were seeded to 7×10^5 cells/10 cm dish and grown overnight in DME medium containing 2.5% bovine calf serum and 2.5% fetal bovine serum. On the following day, cells at 60−70% confluence were transfected with plasmid DNA by the calcium phosphate co-precipitation procedure as described previously (Gerster *et al.*, 1990). To each 10 cm culture dish was added reporter plasmid (usually 5 μg, otherwise as indicated in the figure legends), activator plasmid (amounts indicated in the figure legends), pXGH5 (5 μg) as an internal control and pBluescript as carrier (to bring the total DNA to 20 μg). The levels of human growth hormone expressed from pXGH5 (Selden, 1986) were detected by immunoradiometric assays (Nichols Institute) and used to determine the relative transfection efficiencies and to adjust the amounts of the cellular extracts for CAT assays (Gorman *et al.*, 1982). Each transfection was repeated at least 2−4 times with different plasmid preparations. After autoradiography of the separated forms of acetylated chloramphenicol forms, spots were excised and quantitated by liquid scintillation counting.

For mRNA mapping studies, T-22 cells (293 cells transformed with SV40 large T-antigen) were seeded at 1×10^6 cells/10 mm dish and grown

overnight in DME containing 5% fetal bovine serum. Cells were transfected the next day by the calcium phosphate method (5 h exposure to DNA) and then re-fed with complete medium. Total RNA was extracted at 48 h post-transfection by the guanidine–HCl method and poly(A)$^+$ mRNAs were isolated following the procedure of 'Micro-fast track' mRNA isolation (Invitrogen). For each set, 6 μg of mRNA were analyzed by primer extension as described previously (Lillie *et al.*, 1986). The synthetic oligonucleotide primer (5'CAACGGTGGTATATCCAGTG3'), complementary to the coding region of the CAT gene, was 5' end-labeled with [γ-^{32}P]ATP (ICN) by T4 polynucleotide kinase (Boehringer Mannheim). A dideoxynucleotide sequencing reaction was performed on the wild-type template with the same CAT primer and the products were loaded on a same 6% polyacrylamide urea gel sequencing along with the primer extension products. The radioactive signals were visualized by autoradiography.

Mobility shift assays

Mobility shift assays with purified recombinant USF[43] (Pognonec and Roeder, 1991) were performed as follows: the protein was incubated for 20 min at 30°C with 0.1–0.2 ng of radiolabeled probe in a buffer containing 20 mM Tris–HCl (pH 7.9), 10% glycerol, 10 mM EDTA, 5 mM dithiothreitol (DTT), 80 mM KCl, 200 ng of poly(dAdT):poly(dAdT) as non-specific DNA and bovine serum albumin to bring the total protein concentration to 500 ng, with or without specific competitors as indicated. The reaction products were analyzed subsequently by electrophoresis in 4% non-denaturing polyacrylamide gels for 3 h at 120 V. The radioactive signals were visualized by autoradiography.

DNase I footprint assays

Footprint reactions were conducted under essentially the same conditions as described for mobility shift assays, except that the DNA probe (−53 to +192) was prepared by 3' end labeling with [^{32}P]ATP and Klenow enzyme. After the initial incubation, 1 μl of DNase I (15 μg/ml) was added and the reactions were further incubated for an additional 1 min at room temperature. The reactions were stopped and phenol extracted. Digested products were processed as before (Nakajima *et al.*, 1988) and analyzed in 8% sequencing gels containing urea. The radioactive signals were visualized by autoradiography.

In vitro transcription

Transcription reactions were performed as previously described (Pierani *et al.*, 1990) except that 250 ng of the supercoiled template DNA and 10 μl (200 μg of total protein) of HeLa cell nuclear extract were used. A total of 100 ng of pMLCAT (containing the Ad-ML promoter region from −260 to +30 fused to the CAT gene) was also added as an internal control. The transcribed RNAs were analyzed by primer extension as described previously (Lillie *et al.*, 1986). The synthetic oligonucleotide primer (5'CAACGG-TGGTATATCCAGTG3'), complementary to the coding region of the CAT gene, was 5' end-labeled with [γ-^{32}P]ATP (ICN) by T4 polynucleotide kinase (Boehringer Mannheim). The products of transcripts (127 nt for HCAT, 164 nt for HCATB1, 106 nt for MLCAT and 148 nt for MLICAT) were analyzed by electrophoresis on 6% polyacrylamide gels with 50% urea and autoradiography.

Acknowledgements

We thank H.Kato for providing MLU$_2$HCAT, SP1$_2$HCAT and NF1HCAT templates, P.Pognonec for providing purified recombinant USF[43], L.Q.Rodriguez for excellent technical assistance, V.Kozler and L.K.Beyah for tissue culture assistance, P.Gregor, R.Kovelmen, E.Martinez, S.Malik and E.Sinn for critical readings of the manuscript, and members of the laboratory for discussion, support and encouragement. This work was supported by Public Health Service grants CA42567 and AI27397 to R.G.R. and by a grant from the Pew Charitable Trusts to the Rockefeller University. H.D. is supported by postdoctoral training grant CA09673 from the National Institutes of Health. A.L.R. is a Damon Runyon–Walter Winchell Cancer Research Fund postdoctoral fellow.

References

Ayer,D.E. and Dynan,W.S. (1988) *Mol. Cell. Biol.*, **8**, 2021–2033.
Benezra,R., Davis,R.L., Lockshorn,D., Turner,D.L. and Weintraub,H. (1990) *Cell*, **61**, 49–59.
Berkhout,B., Silbverman,R.H. and Jeang,K. (1989) *Cell*, **59**, 273–282.
Blackwood,E.M. and Eisenman,R.N. (1991) *Science*, **251**, 1211–1217.
Boshart,M., Weber,F., Jahn,G., Dorsch-Hasler,K. Fleckenstein,B. and Schaffnen,W. (1985) *Cell*, **41**, 521–530.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.
Bungert,J., Kober,I., During,F. and Seifart,K.H. (1992) *J. Mol. Biol.*, **223**, 885–898.
Carthew,R.W., Chodosh,L.A. and Sharp,P.A. (1985) *Cell*, **43**, 439–448.
Carthew,R.W., Chodosh,L.A. and Sharp,P.A. (1987) *Genes Dev.*, **1**, 973–980.
Chodosh,L.A., Carthew,R.W., Morgan,J.G., Crabtree,G.H. and Sharp,P.A. (1987) *Science*, **238**, 684–688.
Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kedinger,C. and Chambon,P. (1980) *Science*, **209**, 1405–1414.
Dierks,P., Ooyen,A., Cochran,M.D., Dobkin,C., Reiser,J. and Weissmann,C. (1983) *Cell*, **32**, 695–706.
Garcia,J.A., Harrich,D., Soultanakis,E., Wu,F., Mitsuyasu,R. and Gaynor,R.B. (1980) *EMBO J.*, **8**, 765–778.
Gerster,T., Balmaceda,C. and Roeder,R.G. (1990) *EMBO J.*, **9**, 1635–1643.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Gregor,P.D., Sawadogo,M. and Roeder,R.G. (1990) *Genes Dev.*, **4**, 1730–1740.
Grosschedl,R. and Birnstiel,M.L. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1432–1436.
Hall,R.K. and Tayor,W.L. (1989) *Mol. Cell. Biol.*, **9**, 5003–5011.
Hariharan,N. and Perry,R.P. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1526–1530.
Hen,R., Sasson-Corsi,P., Coden,J., Gaub,M.P. and Chambon,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7132–7136.
Jakobovits,A., Smith,D.H., Jakobovits,E.B. and Capon,D.J. (1988) *Mol. Cell. Biol.*, **8**, 2555–2561.
Kato,H., Horikoshi,M. and Roeder,R.G. (1991) *Science*, **251**, 1476–1479.
Lillie,J.W., Green,M. and Green,M.R. (1986) *Cell*, **40**, 1043–1051.
Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambon,P. (1985) *EMBO J.*, **4**, 3563–3570.
Murre,C., McCaw,P.S., Vassin,H., Caudy,M., Jan,L.Y., Jan,Y.N., Cabrera,C.V., Buskin,J.N., Haushchka,S.D., Lasser,A.B., Weintraub,H. and Baltimore,D. (1989) *Cell*, **58**, 537–544.
Nakajima,N., Horkoshi,M. and Roeder,R.G. (1988) *Mol. Cell. Biol.*, **8**, 4028–4040.
Peritz,L.N., Fodor,E.J.B., Silversides,D.W., Cattini,P.A., Baxter,J.D. and Eberhardt,N.L. (1988) *J. Biol. Chem.*, **263**, 5005–5007.
Pierani,A., Heguy,A., Fujii,H. and Roeder,R.G. (1990) *Mol. Cell. Biol.*, **10**, 6204–6215.
Pognonec,P. and Roeder,R.G. (1991) *Mol. Cell. Biol.*, **11**, 5125–5136.
Roeder,R.G. (1991) *Trends Biochem. Sci.*, **16**, 402–411.
Roy,A.L., Meisterernst,M., Pognonec,P. and Roeder,R.G. (1991) *Nature*, **354**, 245–248.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual.* Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sawadogo,M. (1988) *J. Biol. Chem.*, **263**, 11994–12001.
Sawadogo,M. and Roeder,R.G. (1985a) *Cell*, **43**, 165–175.
Sawadogo,M. and Roeder,R.G. (1985b) *Cancer Cells 4/DNA Tumor Viruses*. Cold Spring Harbor, New York, pp. 147–154.
Sawadogo,M., Van Dyke,M.W., Gregor,P.D. and Roeder,R.G. (1988) *J. Biol. Chem.*, **263**, 11985–11993.
Scotto,K.W., Kaulen,H. and Roeder,R.G. (1989) *Genes Dev.*, **3**, 651–662.
Selden,R.F., Burle-Howie,K., Rowe,M.E., Goodman,H.M. and Moore,D.D. (1986) *Mol. Cell. Biol.*, **6**, 3173–3179.
Seto,E., Shi,Y. and Shenk,T. (1991) *Nature*, **345**, 241–245.
Sirito,M., Walker,S., Lin,Q., Kozlowski,M.T., Klein,W.H. and Sawadogo,M. (1992) *Gene Expr.*, **2**, 231–240.
Smale,S.T. and Baltimore,D. (1989) *Cell*, **57**, 103–113.
Smale,S.T., Schmidt,M.C., Berk,A.J. and Baltimore,D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4509–4513.
Visvader,J. and Begley,C.G. (1991) *Trends Biochem. Sci.*, **16**, 330–333.
Weintraub,H., Davis,R., Tapscott,S., Thayer,M., Krause,M., Benezra,R., Blackwell,T.K., Turner,D., Rupp,R., Hollenberg,R., Zhuang,Y. and Lassar,A. (1991) *Science*, **251**, 761–766.

Received on September 16, 1992; revised on November 6, 1992