The EMBO Journal vol.10 no.13 pp.4197-4207, 1991

Transcriptional termination between the closely linked human complement genes C2 and Factor B: common termination factor for C2 and c-*myc*?

Rebecca Ashfield, Pita Enriquez-Harris and Nick J.Proudfoot

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, UK

Communicated by N.J.Proudfoot

We have demonstrated, using a combination of nuclear run-off and poly(A) site competition assays, that transcriptional termination occurs between the closely spaced human complement genes, C2 and Factor B, soon after the C2 poly(A) site. A comparison of the C2 termination signal with a functionally similar sequence downstream of the human α2 globin gene reveals that both signals function in an orientation dependent manner, with subfragments of the whole signal displaying partial effects. In the case of the C2 termination sequence a protein binds within it, and is partially responsible for the termination effect. We further demonstrate that the same (or closely related) protein binds to the ME1a1 site in the murine c-*myc* promoter, which has been implicated in c-*myc* attenuation. We suggest that the termination/pause sequences positioned downstream of a gene's poly(A) site may constitute the general signals that elicit transcriptional termination in genes transcribed by RNA polymerase II.

Key words: c-*myc*/human complement genes/transcription termination

Introduction

Termination of transcription is defined as the release of the elongating RNA polymerase complex from the DNA template (Holmes *et al.*, 1983). In the case of eukaryotic genes transcribed by RNA polymerase II, the point at which this occurs does not correspond to the mature 3′ end of a mRNA molecule, which is instead generated by cleavage and polyadenylation of the nascent transcript (Proudfoot and Whitelaw, 1988). The mechanism of polII termination is not understood, and only in a few cases have sequences 3′ of polyadenylation sites been identified which are involved in this process. Specific termination signals have, however, been defined for polII transcribed snRNA and histone genes, which give rise to non-polyadenylated transcripts. The mechanism of termination appears unique to each class of polII gene: snRNA genes possess a consensus sequence ($GU_3A_3N_3AGA$) which gives rise to termination at, or close to, the mature 3′ end of the RNA (Kunkel and Pederson, 1985; Hernandez and Weiner, 1986; and Neuman de Vegvar *et al.*, 1986). In the case of the closely linked, repetitive sea urchin histone genes, termination is dependent on complex signals within and 3′ to the gene (at least in the case of the H2A gene). Furthermore this process occurs independently of the 3′ end processing mechanism (Birchenmeier *et al.*, 1984; Johnson *et al.*, 1986; Briggs *et al.*, 1989). However, for the mouse H2A gene, termination is dependent on mRNA 3′ end processing, possibly reflecting a less stringent requirement for termination 3′ to this gene (Chodchoy *et al.*, 1991). Termination of polII genes encoding polyadenylated mRNA has been shown in a number of cases to require an intact poly(A) signal, for example in human α2 globin, mouse β globin, adenovirus major late and polyoma virus late genes (Whitelaw and Proudfoot, 1986; Logan *et al.*, 1987; Connelly and Manley, 1988, 1989a; Lanoix and Acheson, 1988). This has given rise to models which postulate that after an RNA polymerase transcribes a poly(A) signal, it becomes destabilized, and liable to detach from the template. This is achieved either by release of an 'antitermination factor' from the polymerase (Logan *et al.*, 1987), or as a result of cleavage of the RNA at the poly(A) site (Connelly and Manley, 1988, 1989a; Proudfoot, 1989). This would expose the nascent RNA to the actions of an exonuclease and helicase (perhaps analogous to the prokaryotic rho factor; see Brennan *et al.*, 1987), resulting ultimately in the release of the polymerase from the template.

The position of termination has been characterized for several genes by nuclear run-off analysis, and is generally heterogeneous, occurring over hundreds of base pairs (Proudfoot, 1989; Citron *et al.*, 1984; Hagenbuchle *et al.*, 1984). This may be the case for many mRNA genes, where there is no need to have precise termination after the cleavage site. However, closely spaced genes are more likely to require closely coupled cleavage and termination; otherwise, the polymerase transcribing one gene could read into the promoter of a downstream gene, reducing initiation from that promoter. This is known as promoter occlusion or transcriptional interference, an example of which is seen in retroviruses, where transcription from the 5′ LTR prevents initiation at the 3′ LTR (Cullen *et al.*, 1984). In such cases, a termination or pause sequence downstream of a polyadenylation site could slow down or halt the polymerase, so enhancing its chances of being released from the template. This was directly demonstrated for tandem α globin gene constructs, which exhibit interference; this could be alleviated by insertion of a mouse β globin or sea urchin histone termination sequence between the genes (Proudfoot, 1986).

A termination sequence could take the form of a protein binding site, RNA secondary structure, unusual DNA structure, such as bent DNA, or a combination of these. Indeed, a CCAAT box binding protein has been shown to cause termination in the adenovirus major late promoter (Connelly and Manley, 1989a,b). This resembles the situation of polI genes: a specific protein, TTF-1, is required for termination of mouse rRNA genes (Grummt *et al.*, 1986; Henderson and Sollner-Webb, 1986; McStay and Reeder, 1986). Thus it appears possible that a transcription factor bound to a downstream promoter can terminate transcription from an upstream gene; since not all proteins are able to function in this way (Connelly and Manley, 1989a,b), the

mechanism is unlikely to be simply providing a steric block to polymerase elongation. Premature termination, or attenuation, has been well characterized in both the human and murine c-*myc* genes (Eick and Bornkamm, 1986; Bentley and Groudine, 1988; Wright and Bishop, 1989). Attenuation within the first exon of murine c-*myc* requires a 180 bp fragment of this exon, which includes a potential stem−loop followed by a T rich region. Attenuation is prevented by deletion of a promoter element, ME1a1 (Miller *et al.*, 1989), again implicating a protein in polII termination.

We have recently investigated the mechanism of transcriptional termination between the linked human $\alpha 2$ and $\alpha 1$ globin genes (separated by 3.5 kb). We identified a transcriptional pause site between these two genes which, in combination with a strong poly(A) signal, elicits transcriptional termination (Enriquez-Harris *et al.*, 1991). We wished to extend these observations to other polII genes to allow a comparison of their termination signals. This study investigates termination between the human complement genes C2 and Factor B, located in the MHC class III region. These are expressed simultaneously in the adult liver and are very closely spaced (421 bp from the C2 poly(A) site to the Factor B cap site) (Wu *et al.*, 1987). Factor B is expressed at several times the level of C2. This suggests that termination of C2 transcripts occurs within the first 400 nucleotides following the C2 poly(A) signal. We present data to show that termination does indeed occur, soon after the C2 poly(A) site. This evidence comes from two separate studies: nuclear run-off analysis, and a poly(A) site competition assay. This latter procedure was validated using the previous defined termination/pause sequence of the human $\alpha 2$ globin gene (Enriquez-Harris *et al.*, 1991). Using the poly(A) site competition assay, we define a 156 bp orientation specific termination sequence for the C2 gene, which binds a protein. This same protein is also shown to bind the ME1a1 promoter element of murine c-*myc*. It is postulated that this protein, in conjunction with some other property of this sequence, is responsible for termination of C2 transcription.

Results

Transcriptional analysis of the C2−Factor B intergenic region

Nuclei were isolated from human liver tissue and nascent RNA chains were elongated in the presence of [α-^{32}P]UTP. The conditions were such that *de novo* initiation and RNA processing should not occur; the assay is known as nuclear run-off analysis (Weber *et al.*, 1977; Bentley and Groudine, 1988). ^{32}P-labelled RNA was isolated and hybridized to excess single-stranded antisense DNA probes, derived from the 3′ end of C2, the 5′ end of Factor B, and the intergenic region. The precise location of these probes is shown in Figure 1A. The signal intensity from each probe is propor-

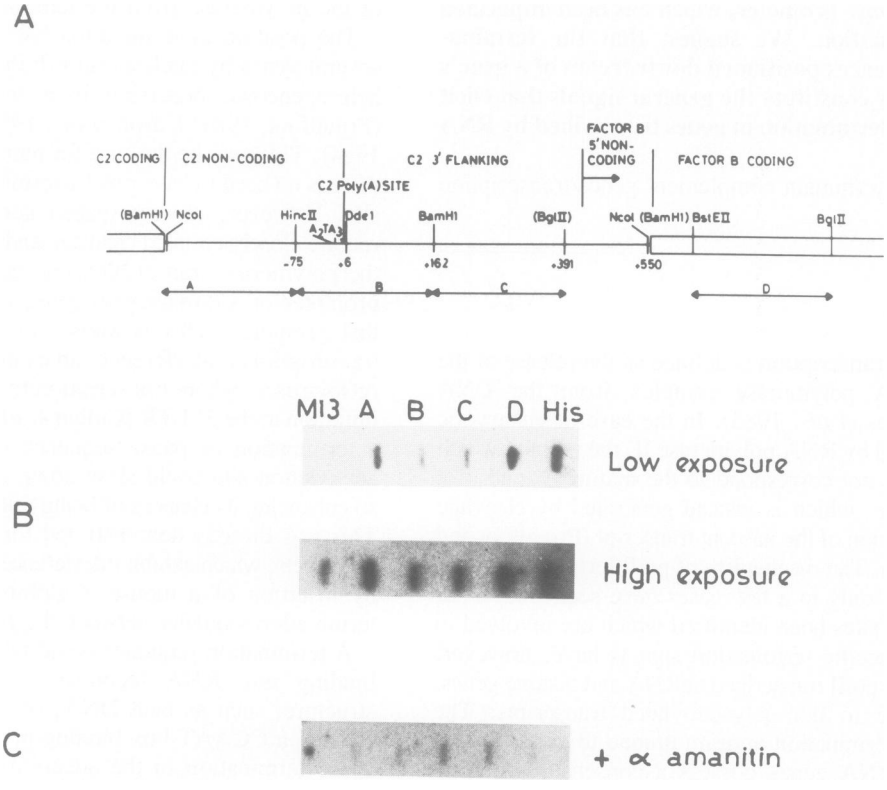

Fig. 1. A. Structure of the C2-Factor B intergenic region, showing nuclear run-off probes A−D. Numbering of sites is with respect to the C2 poly(A) site. The *Dde*I site is not unique. Restriction sites given in brackets are only present after the *Nco*I fragment has been subcloned to create the *Bam*HI sites (in pMLC2.B) and Bal31 deleted to give the *Bgl*II site (in clone pMLC2.BΔ24); see Materials and methods. The open boxes are C2 and Factor B coding sequences, and the solid line is non-coding and intergenic DNA. **B** and **C**. Nuclear run off analysis of human liver nuclei, using probes A−D and His. Probe M13 is M13mp19 DNA, and detects background hybridization. Probe His is mouse histone H4 DNA, and controls for the level of polII transcription. Probes A and D gave signals above the M13 background signal in a ratio of 1:2 respectively (once the background M13 signal is subtracted, and the number of U residues in each probe is taken into account). C. is an experiment in which nuclei were preincubated with 3.8 ng/μl α-amanitin. Under such conditions polII but not polI or polIII transcription is inhibited.

tional to the number of RNA polymerases present over that region of DNA in the liver nuclei. The only other factor that could affect the signal is the number of U residues in each probe; this is taken into account when calculating the relative signals. Thus, if transcription terminates, RNA polymerases will detach from the DNA, and the signal from a probe 5′ to the point of termination will be greater than the signal from a probe 3′ to this point.

Figure 1B shows two exposures of the nuclear run-off signals obtained from liver nuclei using the probes A−D as indicated in Figure 1A. Although these signals are relatively low, it is clear that both probes A and D give signals well above the M13 background. Probe D gives a 2-fold greater signal than probe A as it detects transcripts from the 5′ end of factor B, which is expressed at higher levels than C2 (Wu *et al.*, 1987). The fact that the signal from probe A is above background while those of probes B and C are not, demonstrates that a significant degree of transcriptional termination is occurring, probably within 100 bp of the C2 poly(A) site (since probe B ends 162 bp 3′ of the poly(A) site). The probe His signal in Figure 1B is for histone H4 transcripts. These transcripts as for C2-B (probes A−D) are clearly sensitive to α amanitin inhibition (used at levels specific for polII transcription) as shown in Figure 1C, confirming that the signals detected on probes A and D derived from polII transcription.

C2 poly(A) signals

The nuclear run-off analysis described in the previous section indicates that transcriptional termination occurs immediately downstream of the C2 poly(A) site. We therefore carried out experiments to delineate the C2 poly(A) site so that it could be distinguished from adjacent termination signals. In particular, an experiment was performed to determine how much sequence 3′ to the A_2TA_3 poly(A) signal is required for poly(A) site function. Fragments of DNA containing the C2 A_2TA_3 and 12 bp, 46 bp and 82 bp of 3′ sequence (to generate C2Δ1, C2Δ2 and C2Δ3 respectively), were cloned

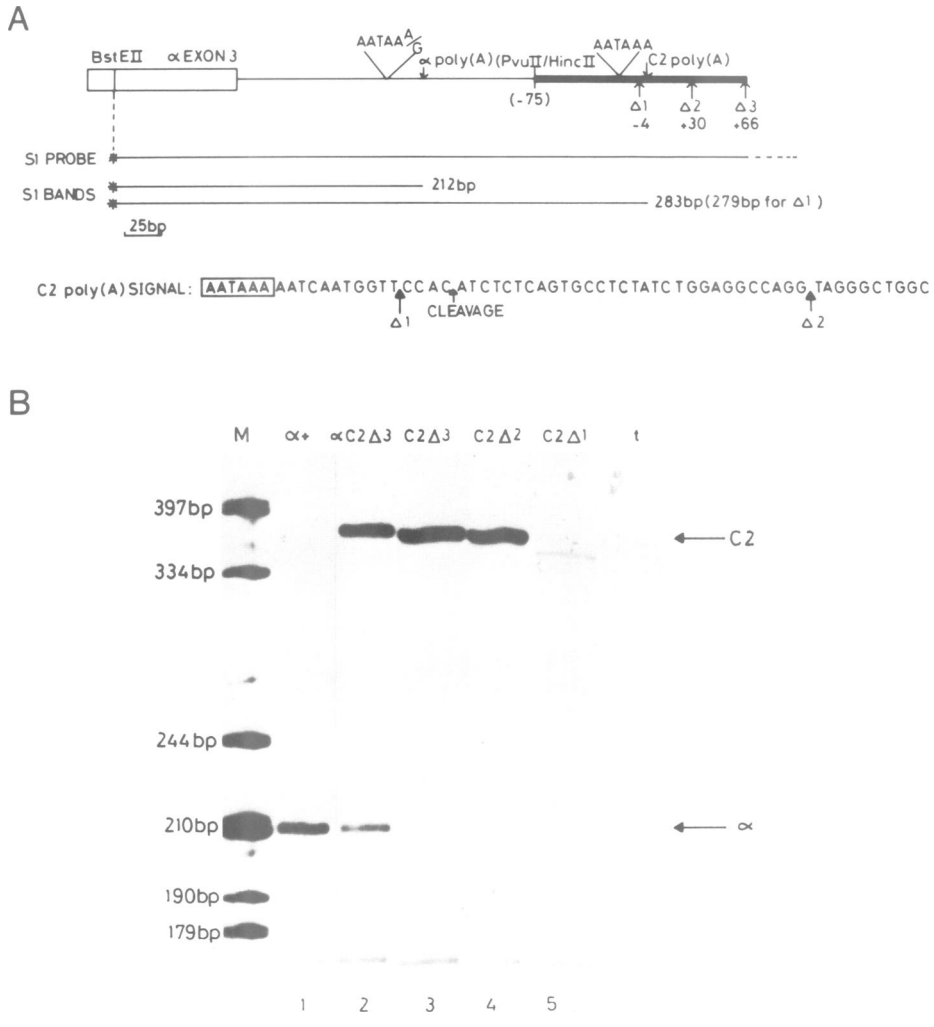

Fig. 2. A. Structure of αC2Δ3 and C2Δ1-3. The open box is the 3rd exon of the human α2 globin gene, and the solid box is a 70−140 bp segment of C2 3′ DNA, containing the C2 poly(A) signal, and sequence up to the Bal31 end points of Δ1-3. The α globin poly(A) signal is mutant in clones C2Δ1-3 (AATAAG), and wild type is αC2Δ3 (AATAAA). The S1 probe was generated by cleavage at the BstEII site in exon 3 of α globin, and expected bands for use of the α and C2 poly(A) signals are shown. The sequence of the C2 poly(A) signal is given, with the A_2TA_3 motif boxed, and the positions of poly(A) cleavage and Bal31 end points Δ1 and Δ2 indicated. **B.** S1 analysis of HeLa RNA, transfected with αC2Δ3 and C2Δ1-3. 'C2' and 'α' denote bands from usage of the C2 and α2 poly(A) signals, respectively, and 't' is a tRNA control lane. α^+ (lane 1) is RNA from pα2W3′PS, a plasmid containing the wild type α2 globin gene (only gives an α poly(A) band). The band at ~170 bp is a co-transfection control (rabbit β globin RNA), to measure for transfection efficiency.

into the 3′ flanking sequence of the human α globin gene, with a mutant, non-functional poly(A) signal (Whitelaw and Proudfoot, 1986). A fourth clone, αC2Δ3, was identical to C2Δ3, except that the α globin gene had a functional poly(A) signal. This clone gives a measure of the relative strengths of the C2 and α globin poly(A) signals. Figure 2A shows maps of these clones, while Figure 2B presents the S1 analysis of HeLa cells transiently transfected with them.

Lanes 3 and 4 (Figure 2B) show that C2Δ3 and C2Δ2 give a strong C2 poly(A) band. C2Δ1, however, gives a much weaker band (lane 5), indicating that sequences necessary for the function of the poly(A) signal have been removed (the band is slightly lower than for C2Δ3, as it is a 'mismatch band', see Figure 2A). These constructs demonstrate that a functional C2 poly(A) signal requires, at most, 46 bp of sequence 3′ to the A_2TA_3 motif. It is interesting that clone αC2Δ3 (lane 2), gives a much stronger C2 poly(A) band than α poly(A) band, despite the fact that the poly(A) signal is over 100 bp downstream of that of α globin. The C2 poly(A) signal would therefore appear to be a strong RNA processing signal (more than α globin), even though it has no GT or T rich region 3′ to the A_2TA_3 (see Figure 2A). This is in contrast to our previous finding that poly(A) signals require such a region (Gil and Proudfoot, 1987; Levitt *et al.*, 1989). The C2 poly(A) signal appears to belong to a new class of efficient poly(A) signals without a conventional downstream sequence.

A protein binds the C2-Factor B intergenic region

If there is a sequence downstream of the C2 poly(A) site important in termination, a possible mechanism is that a protein binds, and acts as a block to polymerase elongation. This may be a specific termination factor, or may have a dual role as a transcription factor for Factor B. Indeed, the Factor B enhancer residues in a 450 bp DNA fragment, the 5′ end of which lies 50 bp upstream of the C2 poly(A) signal; the enhancer region contains various consensus sequences for DNA binding proteins (Wu *et al.*, 1987). To determine if the region of DNA implicated in C2 termination does bind a protein, *in vitro* footprint analysis was performed on a 180 bp sequence containing the 156 bp *Dde*I−*Bam*HI fragment lying 3′ to the C2 poly(A) site (see Figure 1A). The DNA was end-labelled with [α-^{32}P]nucleotides, incubated with either rat liver or HeLa cell extract, and cleaved with DNase I. Controls were treated identically, but contained no extract. The cleavage products, after varying lengths of incubation with DNase, were run on an acrylamide gel alongside a sequence ladder of the fragment. Figure 3A shows that a region of DNA is strongly protected from digestion in the presence of rat liver extract; HeLa cell extract gave an identical footprint (data not shown). The figure also shows the sequence of the protected DNA, and its position relative to the C2 poly(A) signal. In addition to the footprint, there is a DNase hypersensitive site 33 bp downstream of the 3′ end of the footprint (at +97).

Figure 3B shows a gel retardation experiment, in which a 31 bp double-stranded DNA oligonucleotide containing the 27 bp footprinted sequence was labelled with ^{32}P and incubated with HeLa cell extract. Lane 1 shows that the oligonucleotide binds a protein from HeLa extract to give one major retarded complex. Lanes 2−5 contain the same amount of labelled probe and extract, with gradually increasing amounts (10−200 times that of probe) of unlabelled, homologous oligonucleotide, which outcompetes

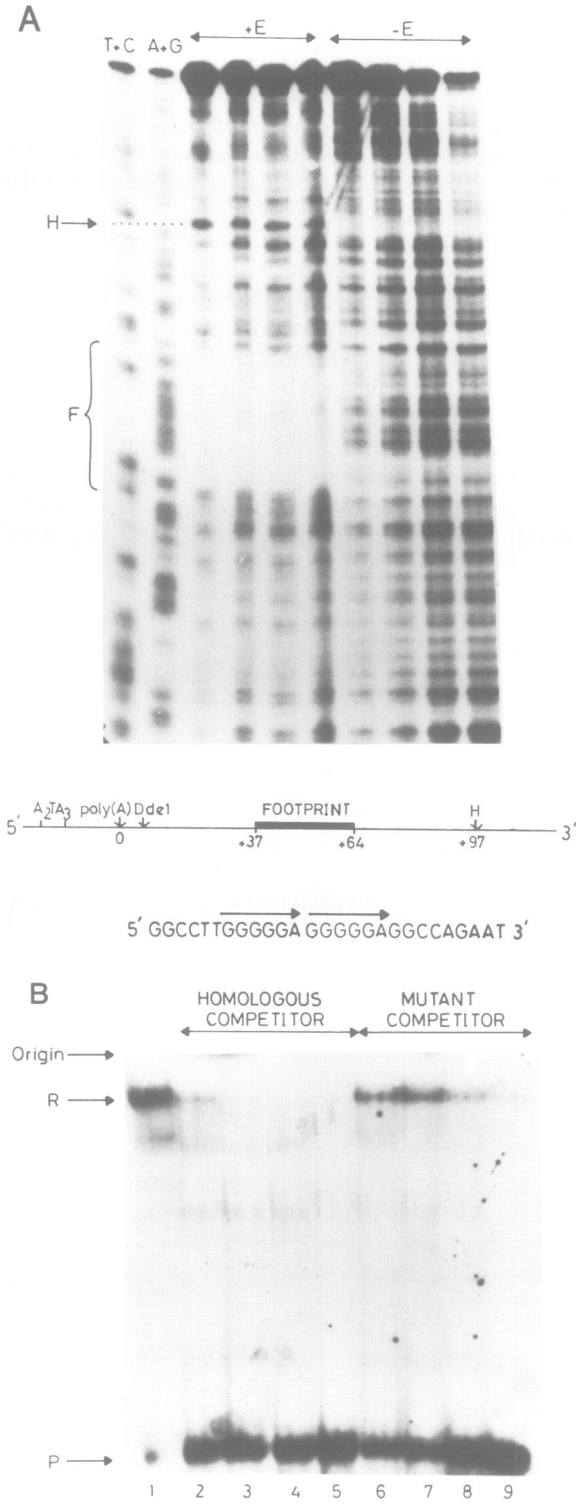

Fig. 3. A. Footprint analysis of DNA 3′ to the C2 poly(A) signal. 'F' denotes the region of protected DNA, and 'H' a hypersensitive site present only when protein is bound. +E and −E indicate presence and absence respectively of rat liver extract. Each set of four lanes was treated with DNase for 1, 2, 5 and 10 mins (from left to right). Lanes T+C and A+G are sequence ladders, with DNA cleaved at the bases indicated. The position of the footprint (solid block) and hypersensitive site relative to the C2 poly(A) site is shown. The sequence of the protected DNA is given, with a 6 bp direct repeat indicated by arrows. **B.** Gel retardation of a 31 bp oligonucleotide homologous to footprinted DNA. 'R' denotes a retarded protein−DNA complex, and 'P' free DNA probe. Lanes 2−5 contain 10, 50, 100 and 200 ng respectively of unlabelled, wild type oligonucleotide, and lanes 6−9 the same quantities of mutant oligonucleotide.

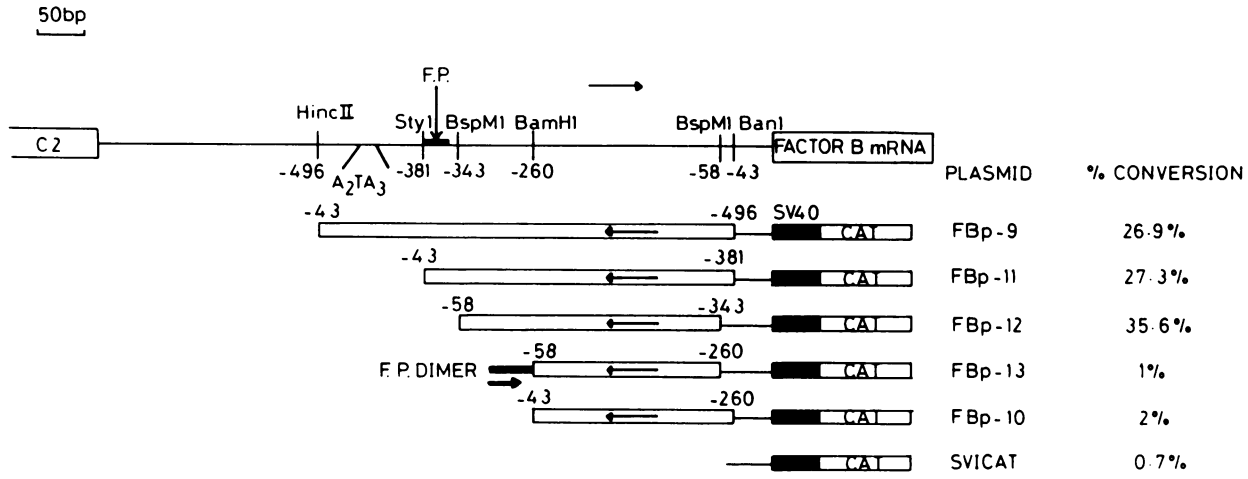

Fig. 4. The diagram illustrates restriction fragments of the Factor B upstream region cloned in reverse orientation into the *Nde*I site of pSV1CAT (Gorman *et al.*, 1982). This plasmid requires enhancer activity in order to express CAT from the SV40 early promoter. Relative CAT activities are given for these constructs transfected into HepG2 cells; assays were done in duplicate, using volumes of extract containing equal amounts of β-galactosidase activities from a cotransfected β-Gal expressing plasmid. Restriction sites are numbered relative to the Factor B mRNA cap site. F.P. denotes the footprint described in Figure 3A, and F.P. dimer a dimerized 31 bp oligonucleotide containing this sequence. The solid box is the SV40 early promoter, and open boxes are C2, Factor B and CAT mRNA sequences as indicated. Arrows indicate the orientation of DNA fragments with respect to the C2–Factor B intergenic region.

the probe for the protein present; thus no probe is retarded. Lanes 6–9 contain gradually increasing amounts of unlabelled, mutant oligonucleotide (5 Gs mutated to Ts: see Materials and methods), which competes with the probe for the protein to a much lesser extent. This demonstrates that the protein–DNA interaction is specific for the footprinted DNA sequence.

These two *in vitro* assays have shown that a protein binds to the C2 3′ flanking DNA, 37–64 bp 3′ to the C2 poly(A) site. The sequence of the footprinted region is shown in Figure 3A: it contains a [G_5A] direct repeat. As described below, this sequence is homologous to the ME1a1 site in the murine c-*myc* promoter, and, to a lesser extent, to the ME1a2 site (Asselin *et al.*, 1989) (see Figure 7). The ME1a1 sequence is required for c-*myc* expression, so it is possible that the transcription factor to which it binds also participates in Factor B expression.

C2 termination region DNA binding protein is not associated with the Factor B enhancer

It has previously been reported by Wu *et al.* (1987) that the Factor B gene is activated by a liver specific enhancer positioned in the C2-Factor B intergenic region and overlapping with the C2 poly(A) signal. We therefore wished to determine whether or not the DNA binding protein we described in the previous section is part of the Factor B enhancer. The fact that this protein is detected in HeLa cell nuclear extracts (see previous section) indicates that it is not liver specific and is therefore unlikely to be responsible for this liver specific enhancer activity. Even so, ubiquitous transcription factors often act in association with tissue specific ones to activate tissue specific enhancers or promoters.

Figure 4 shows a set of plasmid constructs based on pSV1CAT (Gorman *et al.*, 1982) which has the CAT gene driven by the SV40 early promoter. As shown by Wu *et al.* (1987), the insertion of the Factor B enhancer in reverse orientation upstream of the SV40 promoter (as in plasmid FBp-9) confers high level liver specific expression. In contrast FBp-10, which has a truncated enhancer fragment, gave only low levels of expression. We constructed the additional plasmids FBp-11, 12 and 13 to test the role of the protein binding site described above in the factor B enhancer. Plasmids FBp-11 and 12 delete the 5′ side of the enhancer to just before and after the footprint region respectively. As indicated FBp-9, 10 and 11 all gave nearly equal, high levels of CAT activity when transfected into the liver cell line HepG2. This result indicates that the DNA binding protein is not part of the factor B enhancer. By way of confirmation of this result, a dimer of the footprint oligonucleotide (Figure 3) was inserted into FBp-10 to generate FBp-13. However, this construct was still inactive, as indicated by the low CAT activity value obtained for FBp-10 and 13 (comparable to pSV1CAT). From these results we can conclude that the DNA binding protein footprinting to the termination region of the C2 gene is unlikely to be directly associated with the factor B enhancer. Such observations make it more plausible that this protein may have a direct role in C2 termination.

A 156 bp fragment of C2 3′ flanking DNA behaves as a termination sequence: the protein binding site is partially responsible

We have devised an assay for identification of putative termination signals. The assay is based on poly(A) signal competition where a strong poly(A) site is positioned immediately downstream of a weaker poly(A) site. Normally the strong poly(A) site will be predominantly utilized. However, insertion of transcriptional pause or termination sequences between the two poly(A) sites will favour greater usage of the weaker upstream poly(A) site. The strong, synthetic poly(A) site or SPA (Levitt *et al.*, 1989) was cloned 45 bp downstream of the weaker α globin poly(A) signal in a plasmid containing the human α globin gene; a unique site between the two poly(A) signals allows insertion of possible termination sequences. Relative usage of the signals is determined by S1 analysis of cytoplasmic RNA from transiently transfected HeLa cells (see Figure 5A). In this parent clone, SPA^+, the α poly(A) signal is used in only 8% of mRNA (Figure 5B, lane 2). Any fragment of DNA inserted between the poly(A) signals will alter this figure. A 'neutral' sequence will merely increase the spacing

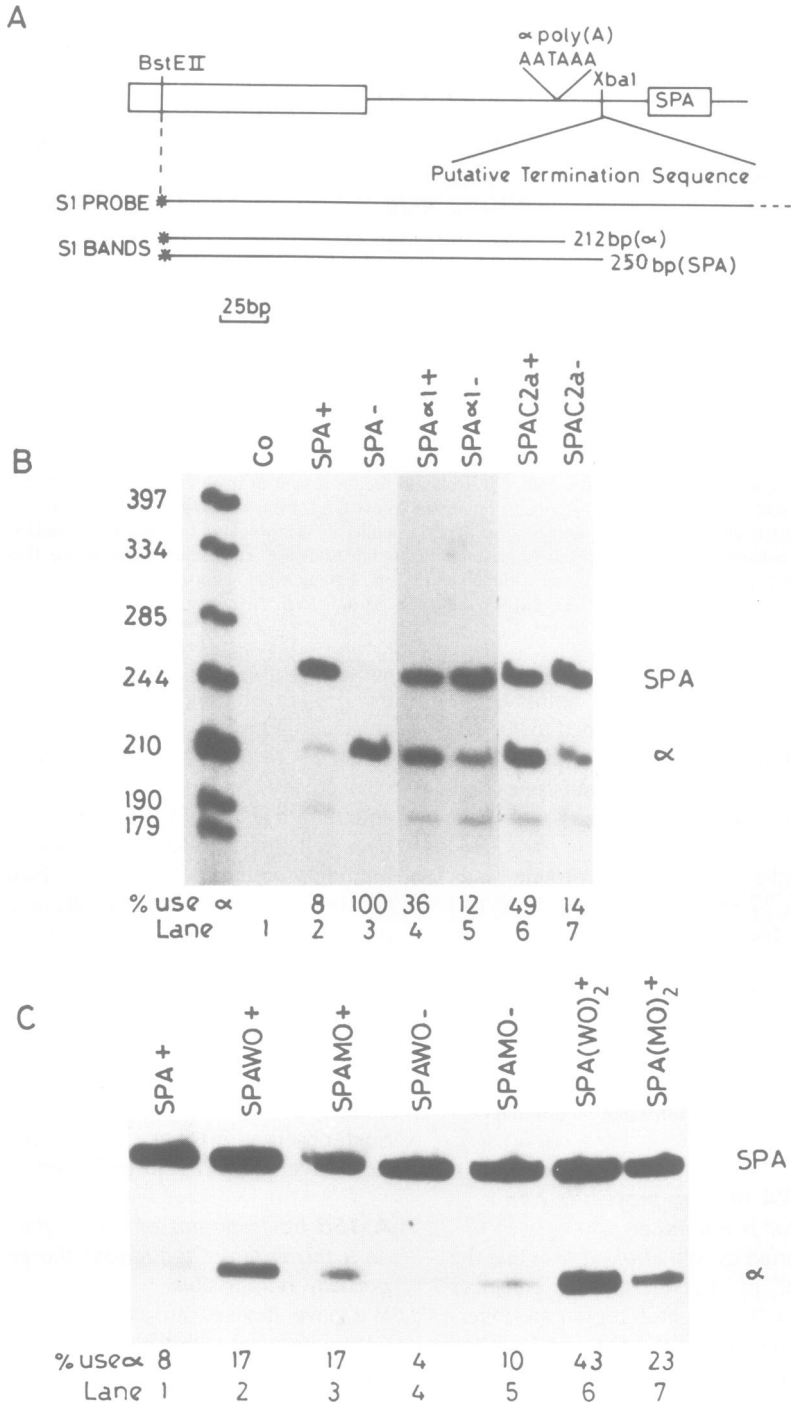

Fig. 5. A. Diagram of SPA$^+$, showing relative positions of α poly(A) and SPA signals. The open box is exon 3 of human α2 globin; S1 probes are generated by cleavage at the *Bst*EII site, and the bands expected from use of the α poly(A) and SPA signals are shown. Putative termination or pause sequences can be inserted into the *Xba*I site between the two poly(A) signals. **B**. S1 analysis of SPAα, and SPAC2a clones. 'SPA' and 'α' denote use of the SPA and α poly(A) signal respectively. No co-transfection control is required, as it is the ratio of the two bands which is measured. The figures beneath the lanes represent the percentage use of the α poly(A) signal, i.e. the proportion of the signal in the lower, α, band. 'SPA$^-$' (lane 3) is RNA from a plasmid containing the SPA in the antisense orientation and so only gives an α band. This shows that SPA bands are from use of the SPA, not read-through transcripts which are polyadenylated at some other site. Co is a tRNA control lane. **C**. S1 analysis of SPAC2 oligonucleotide clones as in Figure 5B.

between the signals, resulting in a small increase in use of the α signal. A sequence that causes termination or pausing the polymerases will have a more drastic effect: termination causes fewer SPA signals to be transcribed, so reducing usage of this signal, whereas a sequence that pauses polymerases increases use of the α poly(A) signal, since it prolongs the period in which only this signal is present in the transcript. It is unlikely that a fragment altering relative use of the signals could do so by favouring recognition of one signal, for example by providing the α poly(A) signal with a more efficient downstream region. This is because the downstream sequences of poly(A) sites have been

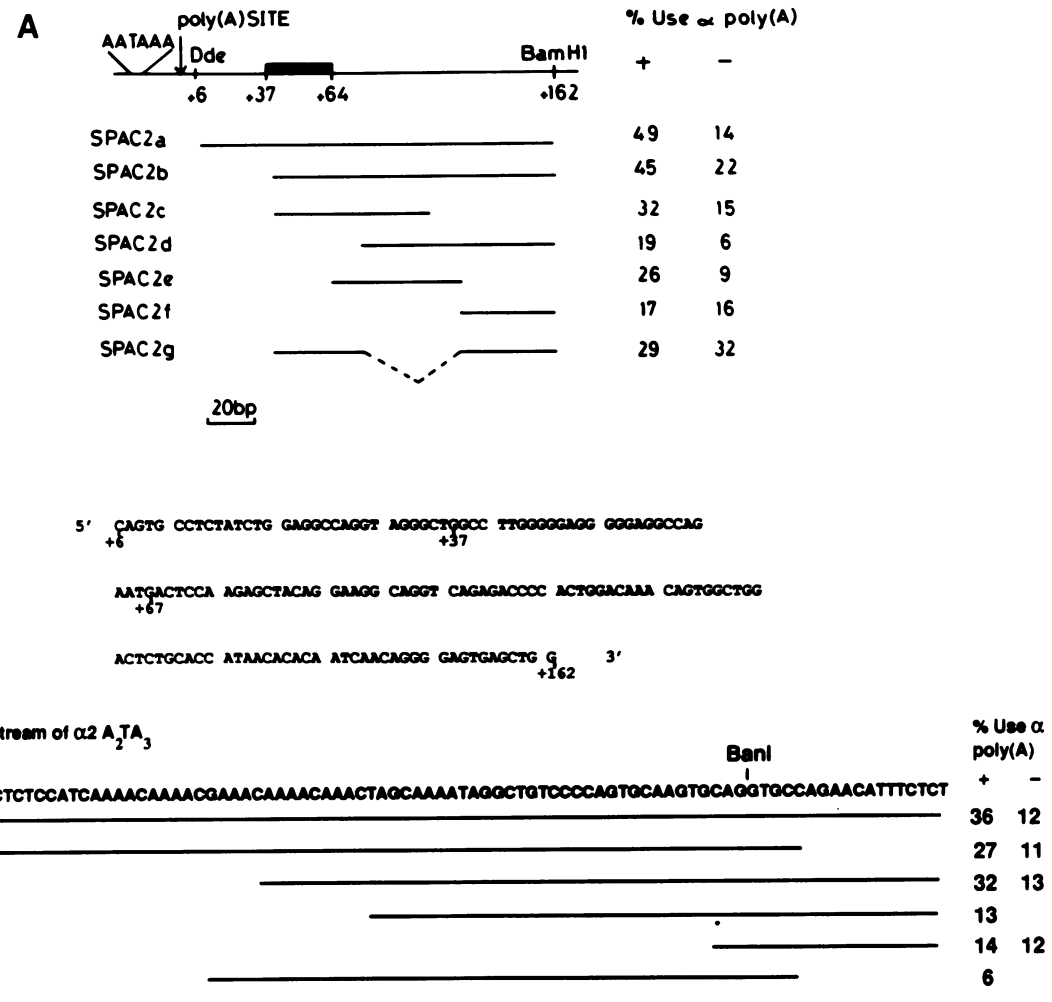

Fig. 6. Analysis of subfragments of the α pause site and C2 termination sequences of the poly(A) site competition assay. **A**. Diagram showing the 156 bp *Dde*I−*Bam*HI C2 fragment in relation to the C2 poly(A) site. The solid block is the region of footprinted DNA. The solid lines represent subfragments cloned into SPA^+, positioned beneath their exact location in the C2 3′ flanking region. Constructs $SPAC2g^{+/-}$ contain two separate fragments cloned adjacent to each other. The figures give the percentage use of the α poly(A) signal on S1 analysis of the clones, in sense (+) and antisense (−) orientations, as in Figure 5B (S1 analysis not shown for these clones). The sequence of the 156 bp fragment is shown beneath the subfragments. **B**. Diagram showing the locations of five subfragments (α2−6) of the α globin pause site (α1) present in clones SPAα1−6. The figures give the percentage use of the α poly(A) signal for each clone. The sequence of the 92 bp pause site is given, with its position downstream of the α poly(A) signal indicated.

demonstrated to show position dependence and only function immediately following the 3′ cleavage site (McDevitt *et al.*, 1987; Gil and Proudfoot, 1987). Sequence added at the cloning site is too far 3′ to the α poly(A) site to have any significant influence on the α poly(A) site signals.

To validate the use of SPA^+ as an assay system for termination signals, we inserted the previously defined α globin gene pause site (Enriquez-Harris *et al.*, 1991) between the two competing poly(A) signals in SPA^+ to make SPAα1. As indicated in Figure 5B (lanes 4 and 5), the α pause site sequence activates the upstream α poly(A) site in an orientation specific manner. In a similar but more pronounced effect, the 156 bp *Dde*I−*Bam*HI fragment 3′ of the C2 poly(A) signal (SPAC2a) also gave an orientation specific activation of the upstream α poly(A) site. Thus, $SPAC2a^+$ (lane 6) enhances the relative use of the α signal so that it is used in ~50% of transcripts, whereas the reverse orientation (lane 7) only has a slight effect, probably entirely due to spacing. This result is consistent with the C2 156 bp fragment causing termination downstream of the α poly(A) site, in an orientation specific manner.

We next wished to determine whether or not the protein binding site is responsible for the C2 termination effect. The 31 bp wild type and mutant oligonucleotides described previously (wo and mo respectively) were therefore cloned into SPA^+, as monomers (both orientations) and dimers (+ orientation only, e.g. $SPA(wo)_2^+$). Figure 5C shows the S1 analysis of these clones. The wild type monomer gives a small but orientation specific termination or pausing effect, while the mutant gives a similar but less orientation dependent effect. The difference between wild type and mutant dimers is much more pronounced with the wild type dimer giving an effect nearing that of the 156 bp C2 fragment. As shown in Figure 3B, the mutant oligonucleotide still displays some capacity to bind the protein since at high concentrations it does compete with the wild type oligonucleotide for protein binding. This weaker binding effect of the mutant oligonucleotide may account for the greater difference seen between the dimers of the two oligonucleotides. Therefore coupled with the gel retardation data shown in Figure 3B, these data suggest that the effects seen with the oligonucleotides in this assay are due to a

protein which binds more strongly to the wild type sequence, especially as a dimer. However, the effect of inserting a single oligonucleotide downstream of the α poly(A) signal is considerably smaller than that of the 156 bp fragment containing this sequence. We therefore conclude that this fragment contains other sequences contributing to the termination or pausing effect.

Dissection of the α globin and C2 pause site/termination sequences

Six different subfragments of the 156 bp C2 terminator fragment were inserted, in both orientations, into SPA^+, and the resulting clones analysed as in Figure 5. The positions of these sequences within the 156 bp fragment, and the percentage use of the α poly(A) signal for each clone, is shown in Figure 6A. All of the fragments give an effect, mostly orientation specific, but none as great as that for $SPAC2a^+$. The clone $SPAC2b^+$ (+40 to +162) is almost as active as $SPAC2a^+$, suggesting that sequence 3′, not 5′, to the footprint is important. Furthermore SPAC2b lacks any of the C2 poly(A) site downstream sequences defined in Figure 2. The possible activation of the α poly(A) site by insertion of C2 poly(A) site downstream sequences can therefore be discounted. Thus the putative C2 termination sequence resides in a 156 bp fragment of DNA from +6 to +162 past the poly(A) site. This region contains the protein binding site, but this is only partially responsible for the termination effect.

Figure 6B shows a similar dissection of the human α2 globin gene's pause site, which causes transcriptional termination when placed downstream of a strong poly(A) signal (Enriquez-Harris *et al.*, 1991). This 92 bp sequence, which does not appear to bind a protein (data not shown), and has no sequence homology with the C2 termination sequence. Five subfragments of the α pause site, generated by Bal31 deletion, were inserted into SPA^+ in sense and (for three fragments only) antisense orientations. These clones, and the resultant percentage usage of the α signal, are shown in Figure 6B. None of these sequences gives the entire effect seen with the 92 bp fragment (Figure 5B), so it must be concluded that most, if not all, of this sequence is required. (Clone SPAα6, however, shows very little activity and is the only clone which lacks the 13 bp sequence at the 3′ boundary of the 92 bp fragment.) The α2 globin pause site, therefore, resembles the C2 termination sequence in that a specific sequence of DNA, much longer than a protein binding site, is responsible for pausing/terminating RNA polymerases, in an orientation specific manner.

The C2 protein binding site and the murine c-myc ME1a1 site bind an identical, or closely related, protein

Following the demonstration that the C2 binding site is partially responsible for the termination effect seen in Figure 5, and the observation that it bears homology to a c-*myc* promoter element, the two sites were compared in a gel retardation assay. Figure 7B shows the sequences of the two oligonucleotide probes used in the assay, a 24 bp sequence containing the c-*myc* ME1a1 site and the 31 bp C2 sequence used previously (see Figure 3B). These oligonucleotides were labelled with ^{32}P and incubated with HeLa cell extract and varying amounts of unlabelled oligonucleotide competitors: a non-specific 31 bp sequence, the 31 bp C2 sequence, and the 24 bp c-*myc* sequence.

Figure 7A shows the result of one such gel retardation experiment. Lanes 1−6 show that both C2 and c-*myc* probes bind a protein specifically (the probes cannot be outcompeted by a 500-fold excess of an unrelated sequence). The retarded complex appears identical in size for both probes, suggesting that they are binding the same protein. Lanes 7−9 and 10−12 contain gradually increasing amounts of unlabelled C2 oligonucleotide, with C2 and c-*myc* probes respectively. This competes with both probes to the same extent. Lanes 13−15 and 16−18 are similar experiments, but substituting unlabelled C2 with c-*myc* oligonucleotide. Again, this competes with both probes, but to a slightly less extent than the C2 competitor. These competition experiments show that the C2 binding site recognizes a protein that is very similar, if not identical, to that binding the murine ME1a1 site. The fact that the C2 site competes more efficiently than ME1a1 suggests that this sequence binds the protein with greater affinity (although this may be caused by differences in probe length).

Finally, Figure 7B compares the nucleotide sequences of the C2 and c-*myc* protein binding sites. As indicated, both sequences possess the identical $PuPuG_3AG_4PuPuG_2$ sequence element. Significantly the asterisked nucleotide positions in the c-*myc* sequence have been shown to be required for protein binding by methylation interference experiments (Asselin *et al.*, 1989). Most of these critical nucleotides fall within the common sequence motif between C2 and c-*myc*.

Discussion

We have shown that termination of the C2 gene occurs immediately after the polyadenylation site. Furthermore, a 156 bp fragment from this region behaves as a terminator in an indirect assay system. Within this fragment is a binding site for a protein, identical or closely related to that binding the c-*myc* ME1a1 motif, which is partially responsible for this termination effect. As described in the Introduction, there are two models for transcriptional termination of mRNA genes. One postulates that an antitermination factor leaves the polymerase as it passes a poly(A) signal, so destabilizing it (Logan *et al.*, 1987). The other states that cleavage at the poly(A) signal results in degradation of the nascent, uncapped transcript, leading to termination (Proudfoot, 1989). Both models predict that closely spaced genes require, in addition to a poly(A) signal, a pause site to slow down elongating polymerases (such as that characterized for α2 globin), or a 'termination sequence', to halt or further destabilize polymerases. Such signals would ensure that termination follows swiftly after cleavage and polyadenylation. They may also be important for genes where termination controls differential splicing or poly(A) site usage, for example in the immunoglobulin heavy chain constant region (Mather *et al.*, 1984). The C2 termination sequence defined here is consistent with either model for termination. We suggest that a strong poly(A) signal coupled with a termination sequence or pause site constitutes a signal for termination. Such a signal was created when the α pause site was placed 3′ of the SPA (Enriquez-Harris *et al.*, 1991), and a similar signal appears to be present 3′ of C2, where a termination sequence lies downstream of the strong C2 poly(A) site. Other examples of termination sequences have been found, both *in vitro* (Kerppola and Kane, 1988; Resnekov and Aloni, 1989; Baek *et al.*, 1986; Dendrick

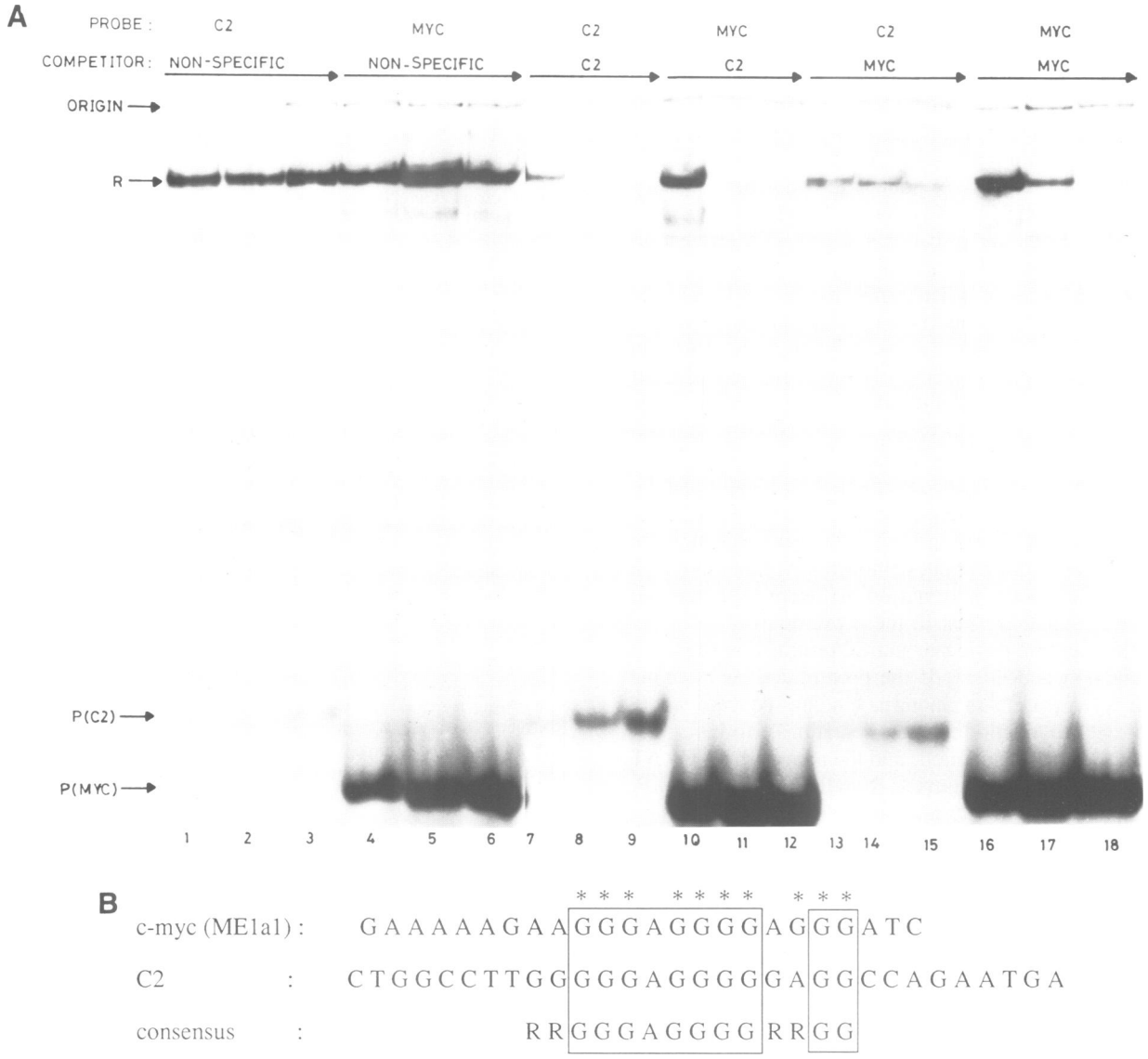

Fig. 7. A. Gel retardation of oligonucleotide homologous to the C2 3′ footprint (C2) and to the c-*myc* ME1a1 promoter footprint (myc). P(C2) and P (MYC) indicate unbound probes (C2 and *myc* respectively), and R the retarded DNA−protein complex. 1 ng of each probe was competed with 10 100 and 50 ng of unlabelled non-specific, C2 and *myc* oligonucleotides, as shown above the retardation (increasing amounts of competitor towards the head of each arrow). **B**. Sequence comparison of the C2 and c-*myc* ME1a1 oligonucleotides. G residues in the *myc* sequence found selectively in the unbound form when methylated are indicated by asterisks (Asselin *et al.*, 1989).

et al., 1987) and *in vivo* (Frayne and Kellems, 1986; Logan *et al.*, 1987; Connelly and Manley, 1989a). The sequences of the C2 termination sequence and α pause site are given in Figure 4C and D respectively.

It is not known how termination sequences and pause sites cause an RNA polymerase to slow down or stop transcribing. For C2 termination, it appears that a DNA binding protein is important. The protein also binds to the ME1a1 site of murine c-*myc*, which is located between the two promoters, P1 and P2 (Asselin *et al.*, 1989). Deletion of this site inhibits P2 expression, causes an increase in P1 expression and also prevents attenuation in exon 1 (Miller *et al.*, 1989). The protein binding to ME1a1 may be participating directly in attenuation, or, if attenuation only occurs when P2 is used, it may act by controlling the relative usage of P1 and P2, causing termination of P1 transcripts within the P2 promoter. This is highly reminiscent of the C2−Factor B system, where the identical (or closely related) protein is required to terminate transcription from an upstream gene, preventing readthrough into the downstream promoter. The protein is present in a variety of cell types (Asselin *et al.*, 1989), and may participate in termination of a range of genes. Sequences homologous to ME1a1 are also found in the attenuation regions of murine genes c-*myb* and c-*fos* (S.Wright, unpublished data).

The protein binding region was found to be insufficient to give the maximal termination effect, although it does contribute to it. Some other property of the 156 bp fragment appears to act in concert with the protein to promote termination. Others have suggested that RNA secondary structure is involved in termination, for example in the chicken β^H globin gene (Pribyl and Martinson, 1988) and the SV40 late gene premature termination sequence (Hay *et al.*, 1982). However, neither the C2 terminator nor α globin pause site are predicted to have any significant structure in the RNA. It is probable, then, that some property of the DNA is important. One possibility is that the DNA is 'bent'. It is interesting that protein binding caused a

hypersensitive site to appear, 33 bp 3′ of the footprint. This raises the possibility that protein binding alters the physical structure of the DNA. Certain sequences are known to be prone to forming a bent conformation, but only do so when a protein binds, for example the *lac* promoter of *E.coli*, which bends when it binds CAP (Wu and Crothers, 1984). It is difficult to predict the tendency to bend from primary sequence, but bent DNA is typically AT rich on one side of the helix. Both the α pause site and C2 termination sequence are A rich (40% and 29% respectively). Intriguingly, the sequences of other terminators are AT rich: for example, those of histone H2A (A rich) (Johnson *et al.*, 1986; Briggs *et al.*, 1989), SV40 (AT rich, with a bend in the DNA) (Hsieh and Griffith, 1988) and rho independent prokaryotic terminators (hairpin loop, then T rich) (Platt, 1987). Certainly for procaryotes, this functions to weaken the interaction between the nascent RNA and the DNA template. The SV40 late gene premature terminator and the c-*myc* attenuator possess sequences which could form stable RNA stem−loops, followed by T rich regions, which are both required for termination, demonstrating that more than one structural motif is required to form a functional termination signal, at least in these examples.

Both the α globin pause site and C2 termination sequence are orientation dependent. If the protein acts as a simple block to polymerase elongation, its binding sequence would also be expected to work in both orientations. However, not all proteins cause termination, and the adenovirus CCAAT box, which terminates transcripts from the adenovirus MLP, only functions in one orientation (Connelly and Manley, 1989a,b). Other terminators are also orientation dependent, for example those of human histone H3.3 and human c-*myc* genes (Kerppola and Kane, 1990). These terminators were, in addition, shown to contain bent DNA, although this was insufficient to cause termination. The protein binding site involved in C2 termination is asymmetric, and only worked in the sense orientation in the poly(A) signal competition assay. We therefore postulate that C2 termination requires an orientation specific interaction between RNA polymerase II, an asymmetric protein, and some other, as yet undefined, property of the DNA sequence.

Materials and methods

Plasmid constructions

The *Hin*dIII fragment of 'Cos10', containing the 3′ end of C2, 5′ end of Factor B and the intergenic region, was cloned into pAT153 to give pH1.3Ba (Wu *et al.*, 1987). The 890 bp *Nco*I fragment of this was cloned into the *Bam*HI site of plasmid pML(C_2AT) (Sawadogo and Roeder, 1985), which contains the adenovirus MLP and a G-less cassette cloned into pU13, to give pMLC2.B. Ligations were of blunt-ended fragments, filled in with the Klenow fragment of DNA polymerase.

NRO probes. Two *Bam*HI−*Hin*cII fragments of pMLC2.B (A,B), a *Bam*HI−*Bgl*II fragment of pMLC2.BΔ24 (from Bal deletion of pMLC2.B; see below, and Figure 1A) (C), a *Bst*EII−*Bgl*II fragment of pH1.3BA (D), and the *Pvu*II−*Afl*II fragment of the mouse histone H4 gene (His) were filled in and blunt-end ligated into M13mp9 cut with *Hin*cII. The antisense orientation was screened for by restriction fragment analysis.

αC2poly(A) clones. pMLC2.B was linearized at the *Xba*I site (3′ to the Factor B sequence), and Bal31 exonuclease used to delete in both directions. Endpoints were blunt-ended and ligated with *Bgl*II linkers to give pMLC2.BΔ clones. *Hin*cII−*Bgl*II fragments of pMLC2.B1Δ1, Δ2 and Δ3 (with 12, 46 and 82 bp of sequence 3′ of A_2TA_3) were blunt-end ligated into the *Pvu*II site of α2W3′PS or α2M3′PS (Whitelaw and Proudfoot, 1986).

CAT clones. The 336 bp *Sty*I−*Ban*I and 285 bp *Bsp*MI fragments from pMLC2.B were inserted in the reverse orientation into the *Nde*I site of pSV1CAT (Gorman *et al.*, 1982) to generate FBp-11 and FBp-12 respectively. The 31 bp wild type oligonucleotide (see gel retardation method for sequence) was cloned as a dimer in sense orientation into the *Bsp*MI site of FBp-10 (Wu *et al.*, 1987), to generate FBp-13. FBp-9 and FBp-10 are gifts from R.D.Campbell.

Poly(A) competition clones. SPA^+ (originally α2W3′PSΔ3SPA) was created by insertion of an *Eco*RI−*Hin*dIII fragment containing the SPA in the *Sma*I site of a pUC119 polylinker into the *Bgl*II linker of Δ3C (Levitt *et al.*, 1989), downstream of the intact α2 globin gene. SPA^- has the SPA inserted in the opposite (antisense) orientation. Oligonucleotide or fragments containing portions of the C2 termination region or the α2 pause site were blunt-end ligated into the *Xba*I site of SPA^+ (see Figure 4 for details).

Transfection and RNA analysis

HeLa, 293 and HepG2 cells were transfected using calcium phosphate precipitates, left on the cells (~50% confluent) for 8−10 h (or 4 h for 293 cells, followed by a 10% glycerol shock). 20 μg of test plasmid was used per 15 cm plate. Cytoplasmic RNA was isolated after 36−48 h; for details of RNA preparation and S1 analysis, see Johnson *et al.*, 1986. All plasmids were cotransfected with 7 μg of RβSVpBR328, expressing SV40 T antigen, to promote replication (Grosveld *et al.*, 1982). S1 probe/RNA hybrid digestion products were subjected to electrophoresis through 4% polyacrylamide 7.3 M urea gels, and autoradiography performed at −70°C with intensifying screens.

β-Gal and CAT assays

HepG2 cells were transfected using the calcium phosphate precipitation method. Cells were harvested 48 h after removal of the precipitate. Cellular lysates were prepared by three rounds of freeze−thawing. Lysates were assayed for β-galactosidase activity as described by Herbomel *et al.* (1984). CAT assays were performed as described by Gorman *et al.* (1982). CAT assays were quantified by cutting chloramphenicol and three acetyl chloramphenicol spots from the TLC plate and counting them in a liquid scintillation counter.

NRO analysis

Human liver nuclei were prepared according to the method of Graves *et al.* (1986). In essence, about 10 cm^3 of liver was homogenized (using a polytron homogenizer), in a buffer containing 0.25 M sucrose. The homogenate was filtered through nylon, made up to 0.5% NP40, and rehomogenized. The homogenate was mixed with an equal volume of buffer containing 2.6 M sucrose, and layered on top of a third buffer, containing 2.1 M sucrose, in a centrifuge tube. This was spun in a swing-out rotor (SW28) at 26 000 r.p.m. for 45 min at 0°C to pellet the nuclei. Nuclei were incubated at 37°C for 5 min in the presence of 44 mM Tris pH 8.3, 6.3 mM $MgCl_2$, 0.08 mM EDTA, 173 mM KCl, 0.29 mM rATP, rGTP, rCTP, 200 μCi [α-^{32}P]rUTP and 30% glycerol. Nuclei treated with 3.8 ng/μl α-amanitin were pre-incubated for 15 min at 37°C. Elongation was terminated by the addition of ~200 U of RNase-free DNase (Boehringer Mannheim), and incubation at 37°C for 45 min. Deproteinization was achieved by addition of 25 μg of proteinase K, and 275 μl of 0.2 M Tris pH 7.5, 0.3 NaCl, 2% SDS, 25 mM EDTA, followed by incubation for 30 min at 37°C. RNA was phenol−chloroform extracted, ethanol precipitated and resuspended in 50 μl TE (10 mM Tris, pH 7.5, 1 mM EDTA) and then 50 μl hybridization solution (4× SSC, 10× Denhardt's, 200 μg/ml sonicated herring DNA, 0.2% SDS, 50% formamide, 10% dextran sulphate). The RNA was heated at 85°C for 5 min, then hybridized to M13 single-stranded probes immobilized on nylon filters. Filters were washed in 0.5 × SSC, 0.1% SDS at 65°C for 30 min, and autoradiography performed as above. For further details, see Enriquez-Harris *et al.* (1991) and Whitelaw and Proudfoot (1986).

DNase footprinting and gel retardation assays

The footprinting probe was obtained by filling in (using Klenow and radiolabelled nucleotides) the *Bgl*II site of the 180 bp *Pst*I−*Bgl*II fragment of $SPAC2a^-$, which contains the 156 bp *Dde*I−*Bam*HI fragment of pMLC2.B, followed by purification from a 4% polyacrylamide gel. 150−250 counts of probe (measured on a Geiger monitor) was incubated with and without 6 μl rat liver extract or 4 μl HeLa cell extract per time point, in 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA, 5% glycerol and 10 μg of poly(dIdC) (Pharmacia). Extracts were prepared according to Dignam *et al.* (1983). After 15 min at room temperature, 2U of RQ1 DNase were added, and aliquots removed after 1−10 min. After phenol−chloroform extraction and ethanol precipitation, samples were resuspended in a formamide dye mix, and equal numbers of counts loaded onto a 6%

acrylamide 7.3 M urea gel. Sequencing (Maxam and Gilbert, 1980) was carried out on the 180 bp probe, and run alongside the DNase I products.

Gel retardation was performed using 31 bp oligonucleotides homologous to the footprinted region (5'-CTGGCCTTGGGGGAGGGGGAGGCCA-GAATGA-3', wild type, and 5'-CTGGCCTTGGGGGATTTTTAGGC-CAGAATGA-3', mutant). The 24 bp c-*myc* oligonucleotide is shown in Figure 5. These were annealed to their complementary strands, to generate competitor DNA, or (one strand end-labelled with [^{32}P] using polynucleotide kinase, and gel purified), to generate probe. 1 μl of HeLa cell extract was incubated for 5 min with 2 μg of poly(dIdC) in the glycerol buffer described previously, with or without competitor DNA (from 10–500 ng). One nanogram of probe (wild type oligonucleotide) was added, and after a further 20 min, the products were run on a 6% acrylamide gel at 350 V. Autoradiography as before.

Acknowledgements

We gratefully acknowledge Duncan Campbell for encouraging us to investigate the C2 termination process and for providing us with the C2-factor B DNA clones. We thank the Wellcome Trust for financial support (programme grant 032773).

References

Asselin,C., Nepveu,A. and Marcu,K.B. (1989) *Oncogene*, **4**, 549–558.
Baek,K.-H., Sato,K., Ito,R. and Agarwal,K. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7623–7627.
Bentley,D.L. and Groudine,M. (1988) *Cell*, **53**, 245–256.
Birchenmeier,C., Schumperli,D., Sconzo,G. and Birnstiel,M.L. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1057–1061.
Brennan,C.A., Dombroski,A.J. and Platt,T. (1987) *Cell*, **48**, 945–952.
Briggs,D., Jackson,D., Whitelaw,E. and Proudfoot,N.J. (1989) *Nucleic Acids Res.*, **17**, 8061–8071.
Chodchoy,N., Pandey,N.B. and Marzluff,W.M. (1991) *Mol. Cell. Biol.*, in press.
Citron,B., Falck-Pedersen,E., Salditt-Georgieff,M. and Darnell,J.E.Jr (1984) *Nucleic Acids Res.*, **12**, 8723–8731.
Connelly,S. and Manley,J. (1988) *Genes Dev.*, **2**, 440–452.
Connelly,S. and Manley,J. (1989a) *Cell*, **57**, 561–571.
Connelly,S. and Manley,J. (1989b) *J. Mol. Cell. Biol.*, **9**, 5254–5259.
Cullen,B.R., Lemedico,P.T. and Ju,G. (1984) *Nature*, **307**, 241–245.
Dendrick,R.L., Kane,C.M. and Chamberlin,M.J. (1987) *J. Biol. Chem.*, **262**, 9098–9108.
Dignam,J.D., Lebowitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
Eick,D. and Bornkamm,G. (1986) *Nucleic Acids Res.*, **14**, 8331–8346.
Enriquez-Harris,P., Levitt,N., Briggs,D. and Proudfoot,N.J. (1981) *EMBO J.*, **10**, 1833–1842.
Frayne,E.G. and Kellems,R.E. (1986) *Nucleic Acids Res.*, **14**, 4113–4125.
Gil,A. and Proudfoot,N.J. (1987) *Cell*, **49**, 399–406.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Graves,B.J., Johnson,P.F. and McKnight,S.L. (1986) *Cell*, **44**, 565–576.
Grosveld,G.G., de Boer,E., Shewmaker,C.K. and Flavell,R.A. (1982) *Nature*, **295**, 120–126.
Grummt,I., Kuhn,A., Bartsch,I. and Rosenbauer,H. (1986) *Cell*, **49**, 901–911.
Hagenbuchle,O., Wellauer,P.K., Cribbs,D.L. and Schibler,U. (1984) *Cell*, **38**, 737–744.
Hay,N., Skolnik-David,H. and Aloni,Y. (1982) *Cell*, **29**, 183–193.
Henderson,S. and Sollner-Webb,B. (1986) *Cell*, **47**, 891–900.
Herbomel,P., Bourachot,B. and Yanic,M. (1984) *Cell*, **39**, 653–662.
Hernandez,N. and Weiner,A.M. (1986) *Cell*, **47**, 249–258.
Holmes,W.M., Platt,T. and Rosenburg,M. (1983) *Cell*, **32**, 1029–1032.
Hsieh,G.-H. and Griffiths,J.D. (1988) *Cell*, **52**, 535–544.
Johnson,M., Norman,C., Reeve,M.A., Scully,J. and Proudfoot,N.J. (1986) *Mol. Cell. Biol.*, **6**, 4008–4018.
Kerppola,T.K. and Kane,C.M. (1988) *Mol. Cell. Biol.*, **8**, 4389–4394.
Kerppola,T.K. and Kane,C.M. (1990) *Biochemistry*, **29**, 269–278.
Kunkel,G.R. and Pederson,T. (1985) *Mol. Cell. Biol.*, **5**, 2332–2340.
Lanoix,J. and Acheson,N.H. (1988) *EMBO J.*, **7**, 2515–2522.
Levitt,N., Briggs,D., Gil,A. and Proudfoot,N.J. (1989) *Genes Dev.*, **3**, 1019–1025.
Logan,J., Falck-Pedersen,E., Darnell,J.E.Jr and Shenk,T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8306–8310.
McDevitt,M.A., Hart,R.P., Wong,W.W. and Nevins,J.R. (1986) *EMBO J.*, **5**, 2907–2913.
McStay,B. and Reeder,R.H. (1986) *Cell*, **47**, 913–920.
Mather,E.L., Nelson,K.H., Haimovich,J. and Perry,R.P. (1984) *Cell*, **36**, 329–338.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
Miller,H., Asselin,C., Dufort,D., Yang,J.-Q., Gupta,K., Marcu,K.B. and Nepveu,A. (1989) *Mol. Cell. Biol.*, **9**, 5340–5349.
Neuman de Vegvar,H.E., Lund,E. and Dahlberg,J.E. (1986) *Cell*, **47**, 259–266.
Platt,T. (1986) *Annu. Rev. Biochem.*, **55**, 339–372.
Pribyl,T.M. and Martinson,H.G. (1988) *Mol. Cell. Biol.*, **8**, 5369–5377.
Proudfoot,N.J. (1986) *Nature*, **322**, 562–565.
Proudfoot,N.J. (1989) *Trends Biochem. Sci.*, **14**, 105–110.
Proudfoot,N.J. and Whitelaw,E. (1988) In Hames,B.D. and Glover,D.M. (eds), *Transcription and Splicing*. IRL Press, Oxford, pp. 97–129.
Resnekov,O. and Aloni,Y. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 12–16.
Sawadogo,M. and Roeder,R.G. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4394–4398.
Weber,J., Jelinek,W. and Darnell,J.E.Jr (1977) *Cell*, **10**, 611–616.
Whitelaw,E. and Proudfoot,N.J. (1986) *EMBO J.*, **5**, 2915–2922.
Wright,S. and Bishop,J.M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 505–509.
Wu,H.-M. and Crothers,D.M. (1984) *Nature*, **308**, 509–513.
Wu,L., Morley,B.J. and Campbell,R.D. (1987) *Cell*, **48**, 331–342.

Received on August 9, 1991