

## Studies on transcriptional regulation of the mucosal T-cell integrin $\alpha E\beta 7$ (CD103)

PAUL W. ROBINSON, SALLY J. GREEN, CHRISTINE CARTER, JOHN COADWELL & PETER J. KILSHAW  
*The Babraham Institute, Babraham, Cambridge, UK*

### SUMMARY

Integrin  $\alpha E\beta 7$  is expressed almost exclusively by mucosal T cells and mucosal dendritic antigen-presenting cells (APCs) and is thought to be induced locally by transforming growth factor- $\beta$  (TGF- $\beta$ ). In mice, mRNA for the  $\alpha E$  subunit was found to be abundant in mucosal T cells but absent from other tissues. Exposure of a T-cell line to TGF- $\beta$  strongly up-regulated  $\alpha E$  mRNA levels within 30 min, and nuclear run-on experiments established that regulation occurred at the level of transcription. The organization of the human  $\alpha E$  gene and a very closely linked novel gene, ELG, was determined. The  $\alpha E$  promoter was tested in T cells and fibroblasts and functioned equally well in both cell types and did not confer TGF- $\beta$  responsiveness. Regions of the promoter providing enhancer activity and phorbol 12-myristate 13-acetate (PMA) responsiveness were identified by deletion studies. DNase I hypersensitivity analysis of 36 kb of the  $\alpha E$  gene revealed one hypersensitive site, found only in  $\alpha E^+$  cells, located near the transcription start points. These results show that, unlike the situation with other integrins, lineage specificity and cytokine responsiveness of  $\alpha E$  transcription are not conferred by the proximal promoter. Specificity may depend on distant control elements that have not yet been identified.

### INTRODUCTION

The integrin  $\alpha E\beta 7$  (CD103) was discovered in the late 1980s,<sup>1,2</sup> but many aspects of its function and regulation remain enigmatic.<sup>3</sup> In healthy humans or mice, expression is restricted almost exclusively to T cells and dendritic antigen-presenting cells (APC) in the mucosal immune system, and the integrin is prominent on intraepithelial lymphocytes in the gut.<sup>4–6</sup> In inflammatory diseases,  $\alpha E$  expression can be detected elsewhere. For example, T cells in the rheumatoid joint express  $\alpha E$  as do CD8 T cells infiltrating kidney tubular epithelium during acute rejection of renal allografts or associated with the ductal epithelium of salivary and lachrymal glands in Sjogren's syndrome.<sup>7–9</sup> Expression of the integrin is induced by transforming growth factor- $\beta$  (TGF- $\beta$ )<sup>4,5</sup> and this cytokine is prominent in all tissue microenvironments in which  $\alpha E^+$  cells are found. The only known ligand for  $\alpha E\beta 7$  is E-cadherin,<sup>10,11</sup> a homophilic cell adhesion molecule of the immunoglobulin superfamily that is expressed on the basolateral aspect of epithelial cells and in adherens junctions. Recent experiments with  $\alpha E$  null/null mice suggest that the integrin plays a role in

retention of T cells in or near the mucosal epithelia,<sup>12</sup> supposedly by interacting with E-cadherin.

The restricted distribution of  $\alpha E\beta 7$  is attributable largely to expression of the  $\alpha E$  subunit, because the beta chain, in association with  $\alpha 4$ , is widely distributed on T and B cells<sup>13,14</sup> and forms the principal mucosal homing receptor. In previous reports, Northern blot analysis of a range of human tissues showed  $\alpha E$  mRNA to be present in gut, lung, thymus and spleen<sup>15</sup> and, for mouse tissues, in the thymus and gut.<sup>16</sup> TGF- $\beta$  is the only cytokine known to induce  $\alpha E$  expression but it has not been clear in previous studies whether TGF- $\beta$  directly influences transcription of the  $\alpha E$  gene or modulates post-transcriptional events.

The organization of the  $\alpha E$  gene in mice, downstream from exon 2, has been described in connection with the preparation of a CD103 null/null mouse<sup>12</sup> but the upstream region encompassing the promoter has not been investigated. There have been no previous studies carried out on the human  $\alpha E$  gene or its promoter. Transcriptional regulation of most other integrin alpha subunit genes has been analysed using reporter constructs transiently transfected into appropriate cell lines. These experiments have generally been successful in revealing transcription control elements that determine cell specificity or responsiveness to cytokines or growth factors.<sup>17–20</sup> The rationale for studying transcriptional control of  $\alpha E$  embraces not only questions concerning its unique tissue specificity but also the possibility that *cis*-regulatory elements in the

Received 4 September 2000; revised 6 February 2001; accepted 12 February 2001.

Correspondence: Dr P. J. Kilshaw, The Babraham Institute, Babraham, Cambridge CB2 4AT, UK. E-mail: peter.kilshaw@bbsrc.ac.uk

promoter or elsewhere in the gene could have practical utility in directing expression of transgenes specifically to mucosal T cells in mice. Such a genetic tool would be of considerable value in manipulating the mucosal immune system.

The present report addresses preliminary questions concerning the regulation of  $\alpha E$  expression, specifically, whether  $\alpha E$  is transcriptionally regulated by TGF- $\beta$  and whether transcription control elements can be identified in the proximal promoter. We have determined the structure of the whole human  $\alpha E$  gene, including the upstream promoter region, and in doing so have identified a novel gene, ELG, which is very closely linked to  $\alpha E$ . Functional analysis of the human  $\alpha E$  promoter, and experiments on induction of  $\alpha E$  transcription by TGF- $\beta$ , were conducted using mouse T-cell lines because suitable TGF- $\beta$ -responsive human T-cell lines are not available. Similarly, the distribution of  $\alpha E$  mRNA *in vivo* was investigated using mouse tissues.

## MATERIALS AND METHODS

### *Cell lines and culture conditions*

The mouse T-cell lymphoma, TK1, and the mouse T-cell hybridoma, MTC-1,<sup>4</sup> were maintained in RPMI-1640 containing 10% fetal calf serum (FCS). When induction or maintenance of  $\alpha E\beta 7$  expression was required, the medium was supplemented with recombinant human (rh)TGF- $\beta 2$  (5 ng/ml). In some experiments MTC-1 cells were treated with phorbol 12-myristate 13-acetate (PMA) (15 ng/ml) and ionomycin (400 ng/ml) 1 hr before transfection with promoter reporter constructs. The human T-cell lymphoma Molt16, which secretes TGF- $\beta$  constitutively and expresses  $\alpha E\beta 7$ , was cultured in Dulbecco's modified Eagle's minimal essential medium (DMEM) + 10% FCS. Mouse L-cell fibroblasts and human MRC-5 fibroblasts were also cultured in this medium.

### *Northern blotting*

Tissues and gut intraepithelial lymphocytes were prepared from 6-week-old BALB/c specific pathogen-free (SPF) mice. Total cellular RNA was extracted from these samples or from cultured cells by standard procedures involving homogenization in Trizol (Gibco, BRL, Life Technologies Ltd, Paisley, UK), treatment with chloroform and precipitation with isopropanol. RNA was resolved on a 1% agarose formaldehyde gel and transferred onto a nitrocellulose membrane. Blots were hybridized with denatured probe at 42° overnight in 5× Denhardt's solution, 5× sodium chloride/sodium citrate (SSC), 50% (w/v) formamide, 0.15% sodium dodecyl sulphate (SDS) and 100 µg/ml of salmon-sperm DNA.<sup>21</sup> Membranes were washed twice at 50° in 1× SSC+0.1% SDS and autoradiographed. The following probes were generated by restriction digestion of plasmids or by the polymerase chain reaction (PCR): an 800-bp fragment from the 3' end of mouse  $\alpha E$  cDNA and a 250-bp fragment from the 5' end of glyceraldehyde phosphate dehydrogenase (GAPDH) cDNA. Probes were labelled using the Prime-It II, Random Primer Labelling Kit (Stratagene Europe, Amsterdam, the Netherlands).

### *Nuclear run-on*

TK1 cells ( $5 \times 10^7$ ) were cultured overnight with or without TGF- $\beta$ , washed in 10 mM Tris/HCl, pH 7.4, containing 3 mM CaCl<sub>2</sub> and 2 mM MgCl<sub>2</sub>, and then lysed in the same buffer

containing 1% Nonidet P-40 (NP-40). After homogenization with a Dounce homogeniser, nuclei were recovered by centrifugation (500 g at 4°) and resuspended in 200 µl of storage buffer [50 mM Tris/HCl (pH 8.3), 40% (w/v) glycerol, 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA] and stored temporarily at -80°. To generate <sup>32</sup>P-labelled unprocessed mRNA transcripts, nuclei were thawed and incubated for 30 min at 30° with 200 µl of reaction buffer: 10 mM Tris/HCl (pH 8.0), 5 mM MgCl<sub>2</sub>, 0.3 M KCl, 10 µl each of 100 mM ATP, 100 mM CTP and 100 mM GTP, and 10 µl of 10 mCi/ml [ $\alpha$ -<sup>32</sup>P]UTP. Extraction, purification and hybridization of labelled RNA to cDNA adsorbed to nitrocellulose followed a standard protocol for nuclear run-on analysis.<sup>21</sup> The membrane for hybridization was prepared as follows: maxipreps of pcDNA3 containing mouse  $\alpha E$  cDNA were prepared using the Qiagen Maxiprep Kit. One-microgram samples in 200 µl of 6× SSC were denatured at 100°, chilled on ice, diluted with 300 µl of 20× SSC and then adsorbed to a nitrocellulose membrane using a dot-blot manifold. Control cDNA for  $\beta$ -actin was prepared similarly. The membrane was air-dried and the cDNA immobilized with a UV-Stratalinker.

### *Determination of the transcription start sites*

The transcription start sites were identified by RNA Ligase-Mediated Rapid Amplification of cDNA Ends (RLM-RACE; Promega UK Ltd, Southampton, UK), according to the supplier's protocol. This technique depends on the selection of full-length and capped mRNA for analysis. Briefly, mRNA was prepared from Molt16 cells by adsorption to oligo (dT) cellulose (FastTrack 2.0; Invitrogen, Groningen, the Netherlands). It was then treated with calf intestinal phosphatase to dephosphorylate mRNA with incomplete, uncapped, 5' ends, followed by treatment with Tobacco Acid Pyrophosphatase (TAP) to remove the cap from full-length mRNA, leaving an intact 5' phosphate. A synthetic RNA adaptor was then added using T4 RNA ligase. The product was reverse transcribed using random primers and the cDNA subjected to nested PCR using primers to the 5' adaptor and outer and inner  $\alpha E$ -specific reverse primers: outer, 5'-TGCATATCAAACACCGTGGTGGC (323 bp downstream of the ATG initiation codon); inner, 5'-TGGGGACATGCTCTACAGGATGGC (260 bp from the initiation codon). In the  $\alpha E$  gene, the primers were located in exon 4 (outer primer) and in the region spanning exons 3 and 4 (inner primer). The RACE products were cloned into pT-Adv (Clontech Labs, Basingstoke, Hampshire, UK) and sequenced. A negative control was performed in which treatment with TAP was omitted. In addition, conventional reverse transcription (RT)-PCR was performed on mRNA using a forward primer (5'-CCTCCTGGCTGAGGGGAAGCTG) encompassing a putative initiator sequence. This primer was used in combination with the inner reverse primer given above.

### *Promoter reporter assays*

Genomic DNA was prepared from buffy coat human blood leucocytes using the Genomic G2 Kit (Helena Biosciences UK, Sunderland, UK). The region of the  $\alpha E$  gene immediately adjacent to the translation start codon and extending to 4 kb upstream was prepared by PCR using *Pfu* polymerase. Deletion mutants were prepared from the full-length product by using a series of forward primers and a

common reverse primer. Forward and reverse primers included a *Kpn*I site and a *Hind*III site, respectively: -4 kb forward, 5'-TACGGTACCAACATGGTGAACCCCGTCTCT-3'; -1909 bp forward, 5'-GGCGGTACCTGGGTTTGTGGGAAAGCA-3'; -1510 bp forward, 5'-GGCGGTACCGAGGAGGGTCACGAGC-3'; -1096 bp forward, 5'-GGCGGTACCAGACACTCATGTTAGAGAGC-3'; -516 bp forward, 5'-GGCGGTACCGGCTGGGGAGGCACTTCCTA-3'; -446 bp forward, 5'-GGCGGTACCTGGGCTGGAAGCTGGGCTGA-3'; -176 bp forward, 5'-GGCGGTACCTGTGTACGAGCCGGAGTGTC-3' and reverse, 5'-GAGAAGCTTCCTTGCTGGAGCAGAGGCG-3'.

PCR products were cloned into firefly luciferase reporter vectors pGL3-Basic and pGL3-Enhancer (Promega), according to standard procedures. The following control constructs were used: pGL3-Control containing the SV40 promoter/enhancer linked to firefly luciferase; pRL-CMV, a Renilla luciferase construct containing the cytomegalovirus (CMV) promoter used to normalize transfection efficiency; and (CAGA)<sub>12</sub>MLP-Luc containing the initiator sequence of an adenoviral minimal promoter plus a classical TGF- $\beta$  response element linked to firefly luciferase (kindly supplied by Dr J. M. Gauthier, Glaxo Wellcome, 91951 Les Ulis Cedex, France).<sup>22</sup> Promoter constructs (45  $\mu$ g) were cotransfected with pRL-CMV (5  $\mu$ g) into cell lines ( $4 \times 10^6$  cells) by electroporation (Gene Pulser; BioRad, Hemel Hempstead, Herts, UK). Twenty-four hours later, cell lysates were assayed sequentially for Firefly luciferase and Renilla luciferase using Promega reagents and a luminometer. Results are expressed as relative luciferase activity, representing the ratio Firefly: Renilla signals.

#### DNase hypersensitivity analysis

Buffers and procedures for preparation of nuclei and subsequent DNase I digestion have been described in detail previously.<sup>21</sup> Briefly, nuclei were prepared from  $4 \times 10^7$  Molt16 T cells or MRC5 fibroblasts by lysis in 0.4% NP-40. Aliquots (200  $\mu$ l) of nuclei suspension were treated with 0–20 U/ml of DNase I (DNase Grade 1; Boehringer Mannheim, Roche Diagnostics, Lewes, East Sussex, UK) for 10 min at 25° followed by addition of 200  $\mu$ l of stop solution and overnight incubation at 50° in proteinase K (200  $\mu$ g/ml). DNA was purified and incubated overnight with *Cl*aI or with *E*spI plus *M*luI. Probes (459 bp to 744 bp) were generated by PCR and labelled using a Prime It Random Primer Labelling Kit (Stratagene). The  $\alpha$ E gene was analysed in the region -6.1 kb upstream of the translation start codon to +8.7 kb downstream, encompassing exon 1 and the first half of intron 1. The segments of DNA and the primers used to generate probes specific for the 5' or 3' ends of the fragments are defined as follows, with numbering in accordance with the EMBL database entry AC002993: 37235–45145 (5' upstream and exon 1), forward primer 5'-GTCGGTGAGGAAAGTGAGAGC-3', reverse primer 5'-CAGAGGTGATGGTGCAATTTGAG-3'; 45046–52118 (5' part of intron 1), forward primer 5'-CCATTACCCCAACAAGAACG-3', reverse primer 5'-TCTTCACATCTCTGCCTCGGCCT-3'; Electrophoresis and hybridization were carried out as described above for Northern blotting.

## RESULTS

### Tissue distribution of $\alpha$ E mRNA

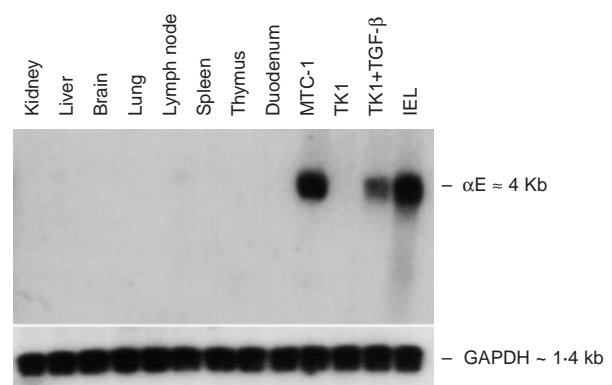
Northern blot analysis was conducted on a range of BALB/c mouse tissues (Fig. 1). The  $\alpha$ E<sup>+</sup> mouse T-cell hybridoma, MTC-1 (cultured in the presence of TGF- $\beta$ ), was included as a positive control, and the mouse T-cell lymphoma, TK1, was tested before and after overnight treatment with TGF- $\beta$ . Intraepithelial lymphocytes (IEL) prepared from small intestine gave a strong  $\alpha$ E signal at  $\approx$  4 kb, as did MTC-1 cells and TK1 cells after treatment with TGF- $\beta$ . All tissue samples, including duodenum, were negative. Failure to detect  $\alpha$ E transcripts in duodenum was a result of the low number of IELs present in the gut of 3-month-old barrier-reared mice.

### TGF- $\beta$ induces $\alpha$ E transcription

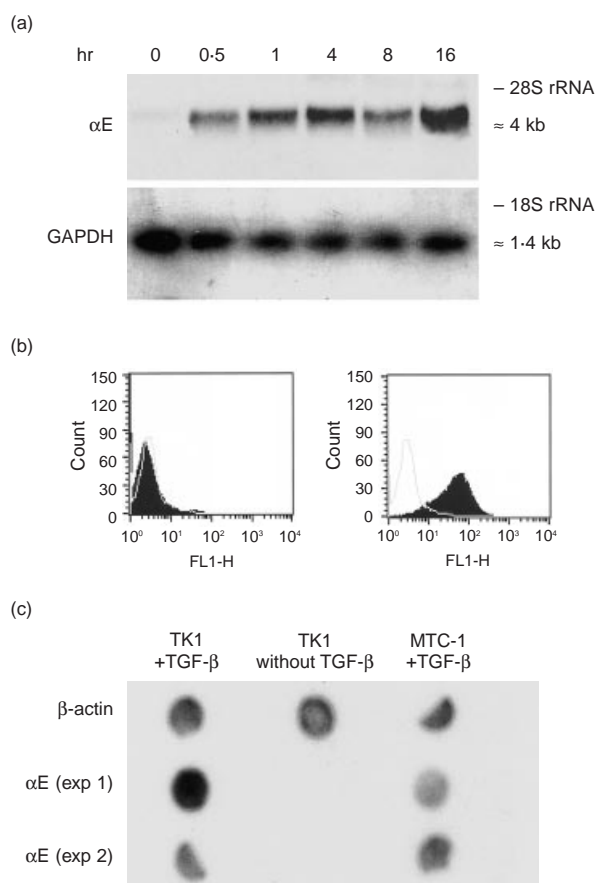
Mouse TK1 T lymphoma cells were cultured for 16 hr with TGF- $\beta$ , and RNA was sampled at specific time intervals within this period. Northern blotting showed that a strong  $\alpha$ E mRNA signal was detected after 30 min and was maximal at 16 hr (Fig. 2a). Cell-surface expression of  $\alpha$ E $\beta$ 7 at the beginning and end of this incubation period is shown (Fig. 2b). Nuclear run-on experiments were performed to determine whether regulation by TGF- $\beta$  occurred at the level of transcription. Newly synthesized, unprocessed, nuclear RNA isolated from TK1 cells, before and after stimulation with TGF- $\beta$ , was tested for nascent  $\alpha$ E transcripts. Figure 2(c) shows that such transcripts were present after, but not before, exposure of the cells to TGF- $\beta$ , clearly demonstrating regulation of  $\alpha$ E transcription. Nuclear RNA from  $\alpha$ E<sup>+</sup> MTC-1 cells cultured continuously in the presence of TGF- $\beta$  provided a positive control, and detection of  $\beta$ -actin mRNA in all samples validated the technique.

### Organization of the $\alpha$ E gene

The human  $\alpha$ E gene was identified in the EMBL database entry AC002993, which consists of a series of contiguous lengths of DNA from chromosome 17 with no previously identified



**Figure 1.** Northern blotting for  $\alpha$ E mRNA in a range of mouse tissues. RNA extracted from BALB/c mouse tissues, from small intestinal intraepithelial lymphocytes (IEL) and from the T-cell lines MTC-1 and TK1, was tested by Northern blotting for  $\alpha$ E mRNA. Blotting for glyceraldehyde phosphate dehydrogenase (GAPDH) verified equal loading.



**Figure 2.** Stimulation of transcription and expression of  $\alpha E$  by transforming growth factor- $\beta$  (TGF- $\beta$ ). (a) TK1 cells were treated for 16 hr with TGF- $\beta$  and RNA was extracted at specific time-points and tested for  $\alpha E$  transcripts. Probing for glyceraldehyde phosphate dehydrogenase (GAPDH) verified equal loading. (b) Cell-surface expression of  $\alpha E\beta 7$  by TK1 cells before (left panel) or after (right panel) treatment for 16 hr with TGF- $\beta$ ; black profiles show staining with monoclonal antibody (mAb) M290 (specific for  $\alpha E$ ), unfilled profiles show the negative control with isotyped-matched irrelevant mAb. (c) Nuclear run-on analysis to test for newly synthesized, nascent,  $\alpha E$  transcripts in TK1 cells before and after treatment with TGF- $\beta$ .  $^{32}P$ -labelled nuclear RNA was extracted from TK1 cells before or after overnight stimulation with TGF- $\beta$  and from  $\alpha E^+$  MTC-1 cells cultured in the presence of TGF- $\beta$ . The RNA was tested for hybridization to  $\alpha E$  cDNA. Detection of nascent mRNA for  $\beta$ -actin in all samples validated the technique.

affiliation to known genes. The entry was detected using BLAST searches with the published human  $\alpha E$  cDNA sequence.<sup>15</sup> The structure of the  $\alpha E$  gene was deduced from the unordered contigs using LocalFASTA and reiterative searching; intron/exon boundaries were confirmed by known properties of splice junctions. Alpha E contains 31 exons spanning at least 85 kb (Fig. 3a). There were three gaps in continuity of sequence in the database entry. A gap in intron 1 was analysed by PCR and found to be 10 bp. Two gaps at the 3' end have not yet been analysed. The alpha E protein contains a unique small acidic region, the X-domain, in which there is a post-translational cleavage site that has no clear functional role.<sup>23</sup> It is encoded by exon 6. Immediately downstream, the I-domain is encoded by four exons, 7–10. In keeping with

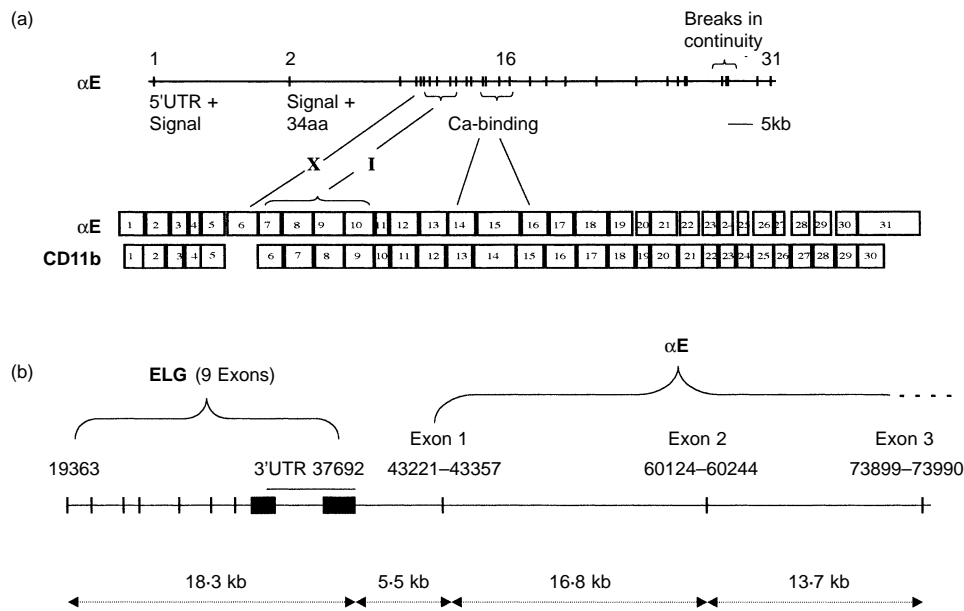
general principles governing the insertion of 'functional modules' into genes, both these regions have phase 1 intron/exon boundaries.<sup>24</sup> Figure 3a shows an alignment of the  $\alpha E$  gene with functionally homologous exons in CD11b, the most closely related integrin (15). Intron/exon boundaries were compared and shown to be closely similar, showing displacement of six amino acids or fewer. The seven repeats of the  $\beta$  propeller of  $\alpha E$  are encoded by exons 2–17, the  $Ca^{2+}$ -binding regions being in three consecutive exons, 14–16.

We have identified a new gene closely linked to  $\alpha E$  (Fig. 3b). The genomic sequence 5.5 kb upstream of  $\alpha E$  matched four clusters of overlapping expressed sequence tags. The only possible open reading frame was identified and the amino acid sequence of a novel protein was deduced (acc. no.: AJ277841). The new gene has been named ELG. It contains nine exons and has an open reading frame of 1023 bp and a 3' untranslated region (UTR) of 3 kb. mRNA transcripts of ELG have been detected in a wide range of cell lines. Using multiple sequence alignment analysis, ELG was shown to have some characteristics in common with nucleosome assembly proteins, including possession of a nuclear localization signal and a highly acidic region. Taken together, these characteristics suggest that the new protein may be a nuclear chaperone.<sup>25</sup> Because ELG is very closely linked to  $\alpha E$ , and that we have analysed most of the 5.5 kb intervening region for promoter activity, the transcriptional regulation of this new gene was examined in parallel with  $\alpha E$ . ELG was shown to be regulated independently of  $\alpha E$  (data not included).

In keeping with most other integrin  $\alpha$ -chain genes, the  $\alpha E$  gene does not contain a TATA box. Analysis of full-length, capped,  $\alpha E$  mRNA by 5' RLM-RACE, identified two transcription start sites at 51 bp and 44 bp, respectively, upstream of the initial methionine codon (Fig. 4a, 4b). A region  $\approx 40$  bp further upstream contains a consensus initiator motif within close proximity of an SP1 site (Fig. 4c). It is probable that transcription starts here also, but we were not able to detect an RLM-RACE product of this length. Nevertheless, a longer mRNA encompassing the putative initiator was detected by conventional RT-PCR (Fig. 4a). RLM-RACE may have failed to reveal this longer species because it is present in low abundance or, alternatively, technical reasons may have prevented ligation of this mRNA to the RLM-RACE adaptor. It is notable that the sequence immediately downstream of the putative initiator was unusually GC-rich and may have displayed secondary structure, which would militate against its detection by RLM-RACE.

#### Activity of the $\alpha E$ promoter

Luciferase reporter constructs were prepared containing 1.9 kb of the  $\alpha E$  promoter extending 5' upstream from the initial methionine codon in exon 1 towards the ELG gene (see Fig. 3b). Two reporter vectors were used: pGL-B (a basic version) and pGL-Enh (which contains the SV40 enhancer incorporated downstream of luciferase). Promoter constructs were tested in TK1 cells (a T-cell line) in the presence or absence of TGF- $\beta$  (Fig. 5a). A construct containing a synthetic TGF- $\beta$  response element (CAGA)<sub>12</sub>, linked to an adenoviral minimal promoter, was included as a positive control. Reporter constructs were also tested in mouse fibroblast L-cells or in the T-cell hybridoma MTC-1 (Fig. 5b). The promoter showed



**Figure 3.** The  $\alpha E$  gene. (a) Organization of intron/exon boundaries and alignment of exons with those of the CD11b gene. Exons coding for the X-domain, the I-domain and the calcium-binding region are identified. (b) The proximal region of the human  $\alpha E$  gene and the closely linked novel gene ELG. Numbers refer to nucleotide positions in the EMBL database entry accession no.: AC002993. Exons are marked as vertical bars or blocks.

neither TGF- $\beta$  responsiveness nor T-cell specificity, functioning equally well in T cells or fibroblasts regardless of the presence of TGF- $\beta$ . In contrast, the (CAGA)<sub>12</sub> construct responded strongly to TGF- $\beta$ . The presence of the SV40 enhancer in the reporter vector doubled the activity of the  $\alpha E$  promoter. To gain information about the strength of the  $\alpha E$  promoter relative to that of a known strong viral promoter, we tested the SV40 early promoter in MTC-1 cells in parallel with  $\alpha E$ . The viral promoter gave a relative luciferase value of  $4.0 \pm 0.3$ ,  $\approx 10$ -fold higher than the basic  $\alpha E$  construct. The  $\alpha E$  promoter construct was extended to 4 kb upstream of the open reading frame and a series of deletion mutants was prepared and tested in MTC-1 cells, which were cultured in the presence of TGF- $\beta$  and were strongly  $\alpha E\beta 7^+$  (Fig. 5c). The experiments were also performed in MTC-1 cells that had been treated with PMA and ionomycin immediately before transfection (Fig. 5d). The region  $-176$  bp to  $-446$  bp displayed enhancer activity, and the region  $-516$  bp to  $-1096$  bp was required for responsiveness to PMA/ionomycin. Inspection of the nucleotide sequence in these two regions showed potentially significant transcription-factor binding motifs, including GATA-1 sites in the putative enhancer region and a series of AP1 sites in the PMA-responsive region, but these have not yet been analysed for functional activity. Motifs for transcription factors of particular significance to T cells,<sup>26</sup> e.g. GATA3, TCF1, LEF1 and Sox4, were not present in the promoter. The 4-kb promoter construct was tested also in fibroblasts (L-cells) and gave values similar to those obtained with the 1.9-kb promoter, as shown in Fig. 5(b).

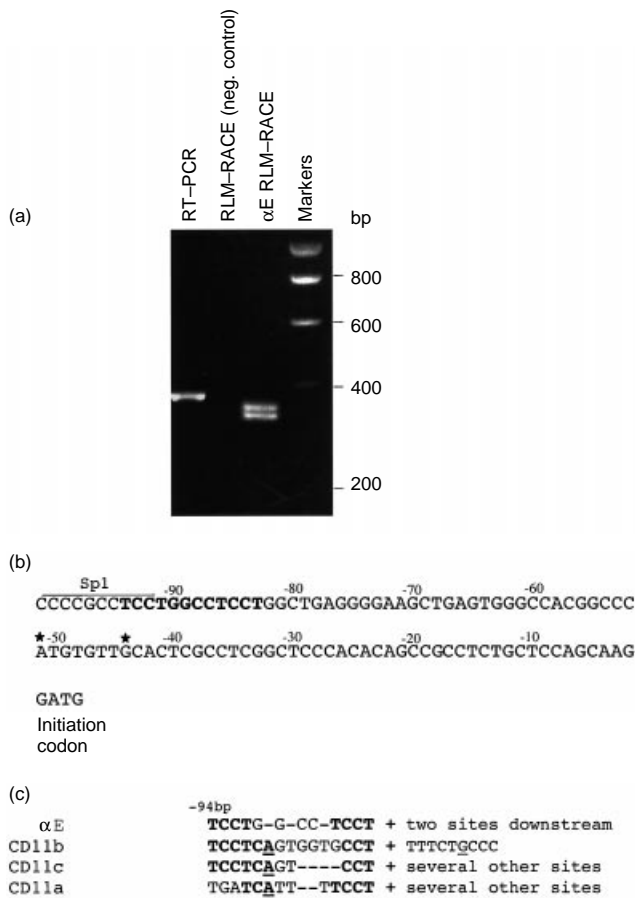
#### DNase 1 hypersensitivity

Because response elements conferring TGF- $\beta$  responsiveness or T-cell specificity were not detected in promoter constructs, we

sought evidence of more distant control elements by testing for hypersensitivity to DNase 1. Human Molt16 cells, which were actively transcribing  $\alpha E$ , were compared with human MRC5 fibroblasts, which cannot transcribe  $\alpha E$ . Figure 6 shows results from the 5' upstream region (left panel) and the first half of intron 1 (right panel). A hypersensitive region is clearly identified,  $\approx 1.9$  kb 5' of the *Esp*I restriction site. This corresponds to a position near the ATG translation start codon and probably reflects the position of the transcription start sites. Analyses of the remainder of intron 1, as well as exon 2 and intron 2, gave negative results closely similar to those in the right panel of Fig. 6, and are not shown. Therefore, no additional DNase 1 hypersensitive sites were detected over the entire 36.5 kb analysed. Detection of the hypersensitivity site near the first methionine codon was used as a positive control in all experiments. It is striking in Fig. 6 that the susceptibility of the  $\alpha E$  gene to degradation by increasing concentrations of DNase was identical in Molt16 and MRC5 cells. This suggests that the chromatin organization in this region was similar in both cell types.

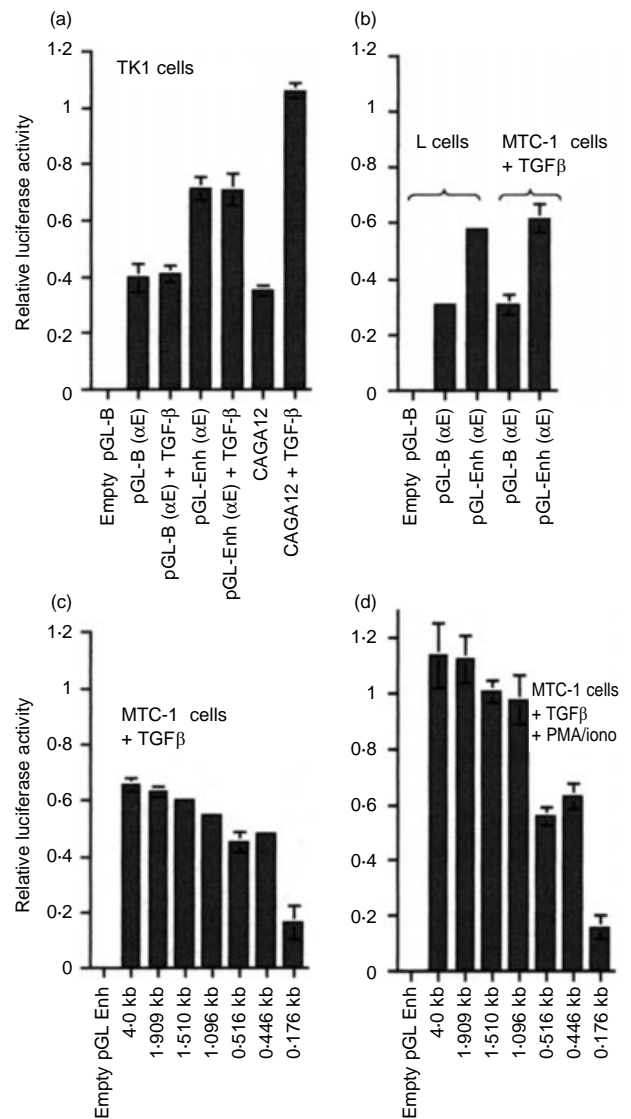
#### DISCUSSION

Our objectives were to affirm that transcription of the  $\alpha E$  integrin gene is normally restricted to mucosal T cells and to gain insight into the regulation of transcription. In BALB/c mice,  $\alpha E$  mRNA was detected only in IELs. The data extend and largely agree with previous results in human and mouse for the tissue distribution of  $\alpha E$  mRNA,<sup>15,16</sup> taking into account differences in the sensitivity of the detection systems used. The distribution of  $\alpha E$  mRNA therefore reflects the tissue distribution of the  $\alpha E\beta 7$  heterodimer.<sup>1,2</sup> Results from nuclear run-on and Northern blotting show unequivocally that TGF- $\beta$  acts within 30 min to induce  $\alpha E$  transcription. Most other



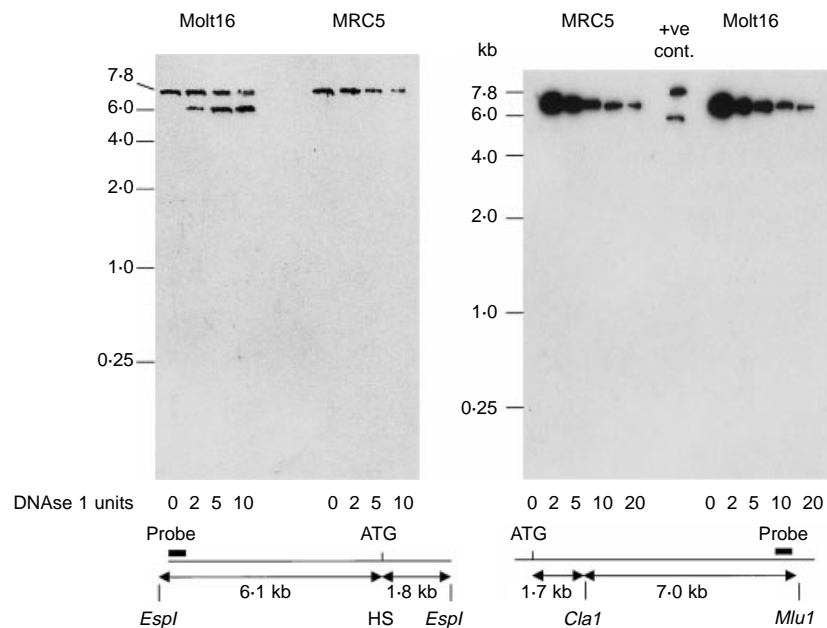
**Figure 4.** Determination of  $\alpha E$  transcription start sites. (a) RLM-RACE was performed on mRNA prepared from Molt16 cells. The track next to the markers shows two closely spaced products, of  $\approx 330$  bp in size, which define two transcription start points (note that the 5' adaptor adds 38 bp to the 5' end of the mRNA). The track on the left shows a longer product of 360 bp, which was generated by conventional RT-PCR using a forward primer spanning a consensus transcription-initiator sequence located  $\approx 45$  bp upstream of the other start sites. This product reflects the existence of a longer, probably minor, species of mRNA and an additional start point; it could not be detected by RML-RACE. (b) The  $\alpha E$  gene immediately 5' of the ATG methionine codon. A consensus initiator sequence is identified in bold type, closely adjacent to an Sp1 site. The two positively identified transcription start sites are identified with stars. (c) Comparison of initiation sequences among related integrin  $\alpha$ -chain genes. Underlined bases indicate the start points.

TGF- $\beta$ -responsive genes are also regulated at the level of transcription. Nevertheless, in a small number of studies, TGF- $\beta$  has been shown to elevate steady-state levels of mRNA by increasing message stability.<sup>27-29</sup> The mechanism depends on recognition of nucleotide motifs in the 3'-UTR of the mRNA by cytoplasmic transacting factors. Recognition sequences are usually repeated or palindromic, and are U-rich. There are no such motifs in the 3'-UTR of  $\alpha E$  mRNA and no conserved motifs common to the human and mouse 3'-UTR. Because we could not detect nascent or mature  $\alpha E$  mRNA in unstimulated TK1 cells, it was considered unlikely that regulation of message stability was a significant factor in the TGF- $\beta$  responsiveness we observed.



**Figure 5.** Alpha E promoter analysis. (a) The 1.9-kb  $\alpha E$  promoter was incorporated into a basic luciferase reporter construct (pGL-B) or into a similar vector containing the SV40 enhancer (pGL-Enh) and transiently transfected into TK1 cells, which were cultured for 24 hr in the presence or absence of TGF- $\beta$ . A positive control containing a TGF- $\beta$  response element (CAGA)12 was tested in parallel. (b) Similar constructs were tested in L-cells in the absence of TGF- $\beta$  or in MTC-1 cells in the presence of TGF- $\beta$ . (c) deletions prepared from a 4-kb  $\alpha E$  promoter construct incorporated into the pGL-Enh vector were tested in MTC-1 cells cultured in the presence of TGF- $\beta$ . (d) The deletion series was tested in MTC-1 cells that had been stimulated, in addition, with phorbol 12-myristate 13-acetate (PMA) and ionomycin.

The human promoter constructs were tested in mouse T-cell lines. This was necessary because we have not yet identified a convenient human T-cell line that synthesizes  $\alpha E$  in response to stimulation by an exogenous source of TGF- $\beta$ . It should be emphasized that the tissue distribution of  $\alpha E$  *in vivo*, and its induction in peripheral T cells *in vitro* by TGF- $\beta$ , is essentially identical in human and mouse. Our  $\alpha E$  promoter constructs showed neither T-cell specificity nor TGF- $\beta$  responsiveness in transient transfection assays. Furthermore, TK1 cells that had



**Figure 6.** DNase 1 hypersensitivity analysis of the proximal region of the human  $\alpha$ E gene. DNA from Molt16 cells that were transcribing  $\alpha$ E was compared with DNA from MRC5 fibroblasts. The left panel shows the region flanking the initial methionine codon and the right panel extends the analysis through half of intron 1. HS indicates the position of a hypersensitive site. +ve cont., positive control.

been stably transfected with the 1.9-kb promoter linked to GFP expressed the reporter in a TGF- $\beta$ -independent manner (data not shown). It is very unlikely that the use of mouse cell lines, as opposed to human cells, had any bearing on our findings. An extensive body of literature validates the use of heterologous mouse/human systems for promoter studies, and human integrin promoters are known to confer tissue-specific expression in transgenic mice.<sup>30–32</sup> The absence of cell lineage specificity and TGF- $\beta$  responsiveness in our experiments was unexpected in view of results with other integrin subunit promoters, many of which have revealed cell lineage-specific regulatory elements in transient transfection assays. For example, a 1-kb construct of the human  $\alpha$ 2 promoter conferred both cell type and differentiation-dependent expression during megakaryocytic differentiation.<sup>20</sup> Studies on the  $\alpha$ 6 integrin, which is expressed mainly in epithelial and endothelial cells, showed that the promoter was functional in keratinocytes and breast carcinoma cells, but not in a leukaemia cell line, again demonstrating lineage specificity.<sup>19</sup> Similarly, the proximal promoter for the CD11c integrin subunit was functional only in LFA-1<sup>+</sup> cells.<sup>18</sup>

DNase 1 hypersensitivity has been used extensively to identify control elements, particularly enhancers and locus control regions, within transcriptionally active DNA.<sup>33</sup> In the present study we compared the  $\alpha$ E gene in Molt16 cells that were actively transcribing  $\alpha$ E mRNA with that in MRC5 fibroblasts which cannot transcribe  $\alpha$ E. Surprisingly, only one hypersensitive site was detected in Molt16 cells, in the proximal part of the promoter. Apart from this site, the  $\alpha$ E gene in Molt16 and MRC5 showed equal susceptibility to DNase 1 digestion, suggesting that the chromatin may have been in an open configuration in both cell types and that transcriptional

regulation of  $\alpha$ E may not involve chromatin remodelling. DNase 1 sensitivity of  $\alpha$ E and the closely linked gene ELG has not yet been examined in non-dividing cells and it is possible that the open configuration is related to cell division or activation.

We have been unable to reveal a TGF- $\beta$  response element in the promoter region of the  $\alpha$ E gene or a DNase 1 hypersensitive site that might contain one. Because our promoter constructs were active in all circumstances, we suggest that TGF- $\beta$  responsiveness may be conferred by relief from transcriptional repression caused by a control element some distance away from the region analysed. De-repression by TGF- $\beta$  receptor signalling has been reported in other circumstances.<sup>34,35</sup> Testing larger constructs in transgenic mice will be required to reveal such an element and to explain the bias of  $\alpha$ E transcription towards CD8<sup>+</sup> cells.<sup>36</sup> The stimulatory effect of PMA/ionomycin on our reporter constructs, and the presence of several AP1 sites in the responsive region of DNA, is consistent with the observation that T-cell activation is important in initiating  $\alpha$ E transcription.<sup>36,37</sup>

The very rapid increase of  $\alpha$ E mRNA in TK1 cells following exposure to TGF- $\beta$  suggests that transcription is a direct and immediate consequence of TGF- $\beta$  signalling. Recently, the principal components of the TGF- $\beta$  signalling pathway have been identified.<sup>38–40</sup> Serine phosphorylation of Smad 2 or Smad 3 by the TGF- $\beta$  receptor is a key event in the initial stages of the pathway. We recently confirmed that induction of  $\alpha$ E expression in TK1 cells is dependent on serine/threonine phosphorylation, but not tyrosine phosphorylation, and showed also that Smad 3 is not required for expression of  $\alpha$ E *in vivo* (P. J. Kilshaw, unpublished). The significance of Smad 2 in this context awaits investigation. The possibility that

stimuli other than TGF- $\beta$  could induce  $\alpha E$  synthesis cannot be entirely ruled out but, apart from one report suggesting that ligation of  $\beta 1$  integrins induces  $\alpha E$  expression,<sup>41</sup> there is no evidence for other mechanisms.

It is probable that transcriptional regulation of the  $\alpha E$  gene will prove to be considerably more complex than that of other integrin genes. Nevertheless, our observation that the distribution of  $\alpha E$  mRNA is exquisitely tissue specific and mainly, if not solely, regulated at the transcriptional level, offers promise that the  $\alpha E$  gene could be used as a vehicle for targeting expression of transgenes to the mucosal immune system by homologous recombination.

#### ACKNOWLEDGMENTS

This work was supported by the Biotechnology and Biological Sciences Research Council, UK. We thank Dr P. Evans for his help in determining the transcription start sites.

#### REFERENCES

- Cerf-Bensussan N, Jarry A, Brousse N, Lisowska-Groszpiere B, Guy-Grand D, Griscelli C. A monoclonal antibody (HML-1) defining a novel membrane molecule present on human intestinal lymphocytes. *Eur J Immunol* 1987; **17**:1279–85.
- Kilshaw PJ, Baker KC. A unique surface antigen on intraepithelial lymphocytes in the mouse. *Immunol Lett* 1988; **18**:149–54.
- Kilshaw PJ. Alpha E beta 7. *J Clin Pathol Mol Pathol* 1999; **52**:203–7.
- Kilshaw PJ, Murrant SJ. A new surface antigen on intraepithelial lymphocytes in the intestine. *Eur J Immunol* 1990; **20**:2201–7.
- Parker CM, Cepek K, Russell GJ, Shaw SK, Posnett DN, Schwarting R, Brenner MB. A family of beta 7 integrins on human mucosal lymphocytes. *Proc Natl Acad Sci USA* 1992; **89**:1924–8.
- Kilshaw PJ. Expression of the mucosal T cell integrin  $\alpha M290\beta 7$  by a major subpopulation of dendritic cells in mice. *Eur J Immunol* 1993; **23**:3365–8.
- Baumgart M, Witt C, Hüge W, Müller B. Increase in the expression of alpha E beta 7, characteristic of intestinal intraepithelial lymphocytes, on cells in the lung epithelium of patients with interstitial lung diseases and in synovial fluid of patients with rheumatic diseases. *Immunobiology* 1996; **196**:415–24.
- Robertson H, Wong WK, Burt AD, Kirby JA. Renal allograft rejection: does CD103 hold the clue to tissue-specific cytotoxicity? *Immunology* 1999; **98**(Suppl. 1):125.
- Fujihara T, Fujita H, Tsubota K, Saito K, Tsuzaka K, Abe T, Takeuchi T. Preferential localisation of CD8<sup>+</sup> alpha E beta 7<sup>+</sup> T cells around acinar epithelial cells with apoptosis in patients with Sjogren's syndrome. *J Immunol* 1999; **163**:2226–35.
- Cepek KL, Shaw SK, Parker CM, Russell GJ, Morrow JS, Rimm DL, Brenner MB. Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and alpha E beta 7 integrin. *Nature* 1994; **372**:190–3.
- Karecla PI, Bowden SJ, Green SJ, Kilshaw PJ. Recognition of E-cadherin on epithelial cells by the mucosal T cell integrin alpha E beta 7. *Eur J Immunol* 1995; **25**:852–6.
- Schon MP, Arya A, Murphy EA *et al.* Mucosal T lymphocyte numbers are selectively reduced in integrin  $\alpha E$  (CD103)-deficient mice. *J Immunol* 1999; **162**:6641–9.
- Kilshaw PJ, Murrant SJ. Expression and regulation of  $\beta 7$  ( $\beta p$ ) integrins on mouse lymphocytes: relevance to the mucosal immune system. *Eur J Immunol* 1991; **21**:2591–7.
- Andrew DP, Rott LS, Kilshaw PJ, Butcher EC. Distribution of  $\alpha 4\beta 7$  and  $\alpha E\beta 7$  on thymocytes, intestinal epithelial lymphocytes and peripheral lymphocytes. *Eur J Immunol* 1996; **26**:897–05.
- Shaw SK, Cepek KL, Murphy EA, Russell GJ, Brenner MB, Parker CM. Molecular cloning of the human mucosal lymphocyte integrin  $\alpha E$  subunit. Unusual structure and restricted RNA distribution. *J Biol Chem* 1994; **269**:6016–25.
- Smith TJ, Ducharme LA, Shaw SK, Parker CM, Brenner MB, Kilshaw PJ, Weis JH. Murine M290 integrin expression modulated by mast cell activation. *Immunity* 1994; **1**:393–03.
- Nueda A, Lopez-Cabrera M, Vara A, Corbi AL. Characterization of the CD11 $\alpha$  ( $\alpha L$ , LFA-1 $\alpha$ ) integrin gene promoter. *J Biol Chem* 1993; **268**:19305–11.
- Corbi AL, Lopez-Rodriguez C. CD11c integrin gene promoter activity during myeloid differentiation. *Leuk Lymphoma* 1997; **25**:415–25.
- Lin C-S, Chen Y, Huynh T, Kramer R. Identification of the human  $\alpha 6$  integrin gene promoter. *DNA Cell Biol* 1997; **16**:929–37.
- Zutter MM, Santoro SA, Painter AS, Tsung YL, Gafford A. The human  $\alpha 2$  integrin gene promoter. Identification of positive and negative regulatory elements important for cell-type and developmentally restricted gene expression. *J Biol Chem* 1994; **269**:463–9.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. *Current Protocols in Molecular Biology*. USA: John Wiley and Sons Inc., 1998.
- Dennler S, Itoh S, Vivien D, ten Dijke P, Huet S, Gauthier J-M. Direct binding of Smad3 and Smad4 to critical TGF $\beta$ -inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J* 1998; **17**:3091–00.
- Higgins JM, Cernadas M, Tan K, Irie A, Wang J, Takada Y, Brenner MB. The role of alpha and beta chains in ligand recognition by  $\beta 7$  integrins. *J Biol Chem* 2000; **275**:25652–64.
- Tuckwell D. Evolution of von Willibrand factor A (VWA) domains. *Biochem Soc Trans* 1999; **27**:835–40.
- Philpott A, Krude T, Laskey RA. Nuclear chaperones. *Semin Cell Dev Biol* 2000; **11**:7–14.
- Kuo CT, Leiden JM. Transcriptional regulation of T lymphocyte development and function. *Annu Rev Immunol* 1999; **17**:149–87.
- Amara FM, Entwistle J, Kuschak TI, Turley EA, Wright JA. Transforming growth factor beta-1 stimulates multiple protein interactions at a unique *cis*-element in the 3'-untranslated region of the hyaluronan receptor RHAMM mRNA. *J Biol Chem* 1996; **271**:15279–84.
- Amara FM, Junaid A, Clough RR, Liang B. TGF-beta (1) regulation of alzheimer amyloid precursor protein mRNA expression in a normal human astrocyte line: mRNA stabilization. *Brain Res Mol Brain Res* 1999; **71**:42–9.
- Werkmeister JR, Blomme EA, Weckmann MT *et al.* Effect of transforming growth factor on parathyroid hormone-related protein secretion and mRNA expression by normal human keratinocytes *in vitro*. *Endocrine* 1998; **8**:291–9.
- Hirsch E, Balzac F, Pastore C, Tarone G, Silengo L, Altruda F. The beta 1 integrin distal promoter is developmentally regulated in transgenic mice. *Cell Adhes Commun* 1993; **1**:203–12.
- Ritchie KA, Aprikian A, Gollahon KA, Hickstein DD. The human leukocyte integrin CD11a promoter directs expression in leukocytes of transgenic mice. *Blood* 1995; **86**:147–55.
- Kwiatkowski BA, Embree LJ, Ritchie KA, Hickstein DD. Human leukocyte integrin CD18 promoter directs low levels of expression of a mutated human CD4 reporter gene in leukocytes of transgenic mice. *Biochem Biophys Res Commun* 1996; **222**:601–6.
- Krebs JE, Peterson CL. Understanding 'active' chromatin: an historical perspective of chromatin remodeling. *Crit Rev Eukaryot Gene Expr* 2000; **10**:1–12.



- 34 Shi X, Yang X, Chen D, Chang Z, Cao X. Smad1 interacts with homeobox DNA-binding proteins in bone morphogenic protein signaling. *J Biol Chem* 1999; **274**:13711–7.
- 35 Verschueren K, Remacle JE, Collart C *et al.* SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *J Biol Chem* 1999; **274**:20489–98.
- 36 Brew R, West DC, Burthem J, Christmas SE. Expression of the human mucosal lymphocyte antigen, HML-1, by T cells activated with mitogen or specific antigen *in vitro*. *Scand J Immunol* 1995; **41**:553–62.
- 37 Schieferdecker HL, Ullrich R, Weiss-Breckwoldt AN, Schwarting R, Stein H, Riecken EO, Zeitz M. The HML-1 antigen of intestinal lymphocytes is an activation antigen. *J Immunol* 1990; **144**:2541–9.
- 38 Miyazono K, tenDijke P, Heldin CH. TGF $\beta$  signalling by Smad proteins. *Adv Immunol* 2000; **75**:115–57.
- 39 Massague J, Wotton D. Transcriptional control by the TGF $\beta$ /Smad signalling system. *EMBO J* 2000; **19**:1745–54.
- 40 Attisano L, Wrana JL. Smads as transcriptional co-modulators. *Curr Opin Cell Biol* 2000; **12**:235–43.
- 41 Rih S, Walker C, Virchow JC, Boer C, Kroegel C, Giri SN, Braun RK. Differential expression of alphaE $\beta$ 7 integrin on bronchioalveolar lavage T lymphocyte subsets: regulation by  $\alpha$ 4 $\beta$ 1 integrin cross linking and TGF $\beta$ . *Am J Respir Cell Mol Biol* 1996; **15**:600–10.