Vol. 211, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
June 6, 1995 Pages 115-122

A REGULATORY ELEMENT IN THE 5'UTR DIRECTS CELL-SPECIFIC EXPRESSION OF THE MOUSE ALPHA$_4$ GENE

Els Schollen, Catherine De Meirsman, Gert Matthijs and Jean-Jacques Cassiman*

Center for Human Genetics, University of Leuven, Gasthuisberg O&N6, Herestraat 49, 3000 Leuven, Belgium

Received April 11, 1995

Summary: Transfection experiments showed that the mouse α_4 promoter contains a downstream element in its 5'UTR which is essential for efficient promoter activity. DNaseI footprinting and electrophoretic mobility shift assays (EMSA) demonstrated that the region from nt +113 to +148 can bind a cell type-specific factor (MIII-3) present in the α_4m expressing cell line L1210 but not in the non-expressing cell lines A9 and LMTK$^-$. Two consensus SP1 sites in this 5'UTR were recognised in L1210, A9 and LMTK$^-$, and with extracts from A9 and LMTK$^-$, an AP2-like protein was shown to bind a downstream AP2 site. Thus both ubiquitous and cell type-specific factors regulate the expression of α_4m.

α4ß1 is a member of the integrin family that mediates cell-extracellular matrix interactions as a receptor for fibronectin and cell-cell interactions via VCAM-1. Interaction of α4ß1, which is expressed on lymphoid and myeloid cells, with both ligands seems to be important for targeting immune cells to sites of inflammation and injury (reviewed in 1,2). When associated to the subunit ß7, α4 directs homing of lymphocytes to Peyer's patch high endothelial venules (HEVs) through binding with mucosal vascular addressin (MAd-CAM-1) (3,4). Besides its role in the immune system, α4 is also involved in myogenesis (5) and probably in embryogenesis since it is expressed in a stage and tissue-specific fashion in mouse embryo (6).

In an attempt to understand the mechanism regulating this differential expression, we previously cloned the mouse α4 (α_4m) promoter (7) and determined its genomic structure (manuscript in preparation). Transfection experiments suggested that the region between the transcription start site and the translation initiation site (5' untranslated region or 5'UTR) had a significant role in the expression of the α_4m gene (7). The human α_4 gene promoter has recently been studied (8,9), showing that transcription of the α_4 gene is controlled through a combination of multiple regulatory elements, including a member of the Ets-family. However, the 5'UTR was not included in these studies. Here, we show that the 5'UTR binds a combination of ubiquitous and cell type-specific factors.

***Corresponding author:** Prof. JJ. Cassiman, Centre for Human Genetics, University of Leuven, Campus Gasthuisberg O&N6, Herestraat, B-3000 Leuven, Belgium. Fax: 32-16-34.59.97.

0006-291X/95 $5.00

MATERIALS AND METHODS

α_4m promoter-luc constructs: PCR was used to generate six 3'deletion fragments of the α_4m promoter from nt -348 to +33, +74, +113, +148, +184 and +221 respectively. These fragments were cloned into the expression vector pGL2-Basic (Promega). pT109-luc (8), containing the Herpes Simplex I thymidine kinase (Tk) promoter, was used as a positive control.

Cell cultures and transfection assays: L1210, a mouse lymphocytic leukaemia cell line, A9, a mouse fibroblast cell line and LMTK$^-$, a mouse connective tissue cell line, were obtained from the American Type Culture Collection (ATCC, MD). The cells were grown in Dulbecco's modified Eagles' medium (DMEM) and Ham's F-12 (1:1) supplemented with 10% Fetal Calf Serum (FCS) (Gibco, BRL). Electroporation of the cell line L1210 and luciferase assays were done as previously described (7).

Footprinting assays: The fragment corresponding to nt -2 to +221 of the mouse α_4 gene was analyzed in DNaseI footprinting assays. To visualize the coding strand, a *BglII/HindIII* fragment was labelled at both ends with Klenow enzyme and [α^{32}P]-dCTP, followed by digestion with *NciI* (see fig 1). Reactions were performed as described (11).

Electrophoretic mobility shift assay (EMSA): Double stranded oligomers, shown schematically in figures 2 and 4C, were labelled by annealing complementary oligomers and filling the protruding ends with Klenow fragment of DNA polymerase in the presence of [α^{32}P]-dCTP (12). Commercial double stranded oligomers were labelled with T_4-PNK and [γ^{32}P]-ATP. Nuclear extracts were prepared (13) and EMSA was done (11) as described. In a supershift assay for AP2, 1µg of the AP2 antibody C-18 (Santa Cruz Biotechnology, Inc.) was added halfway the incubation time. AP2 and SP1 oligomers and a commercial AP2 extract were obtained from Promega.

RESULTS

Functional analysis of the 5'UTR

A previous study indicated that the 5'UTR of the α_4m gene contains elements which are required for high promoter activity. In the present study, 3' deletion constructs were used to narrow down the region of interest. In transfection experiments in the α_4m expressing mouse cell line L1210, construct pGL-221 (-348 to +221) directed the highest level of luciferase activity and was taken as 100%. Sequential deletions from +221 to +184 and +148 resulted in a decrease of promoter activity from 100% to 82%±38 and 74%±20 respectively. Deletion of the fragment from nt +113 to +148 (pGL-113) caused a two fold decrease (to 33%±8) of activity. Further deletions had only a negligible effect on promoter activity. These results suggest that the region from +113 to +148 contains an element with a strong regulatory influence on the α_4m promoter.

The 5'UTR is recognised by several tissue-specific and ubiquitous transcription factors

The entire 5'UTR was analyzed for the binding of trans-acting factors in DNaseI footprinting experiments. In order to identify tissue-specific factors, nuclear extracts from L1210 and from two non-expressing cell lines (A9 and LMTK$^-$)(7) were incubated with a probe covering nt -2 to +221 (fig 1). Two areas (FPA and FPB) were protected from DNaseI digestion with extracts from L1210, A9 and LMTK$^-$. FPA, from nt +12 to +31, was equally protected in all three cell lines. This region is probably not involved in differential expression of α_4m and was not further analyzed. The region from nt +51 to nt +147 (FPB) was extensively protected from DNaseI digestion with extracts from the three cell lines but the distribution of hypersensitive

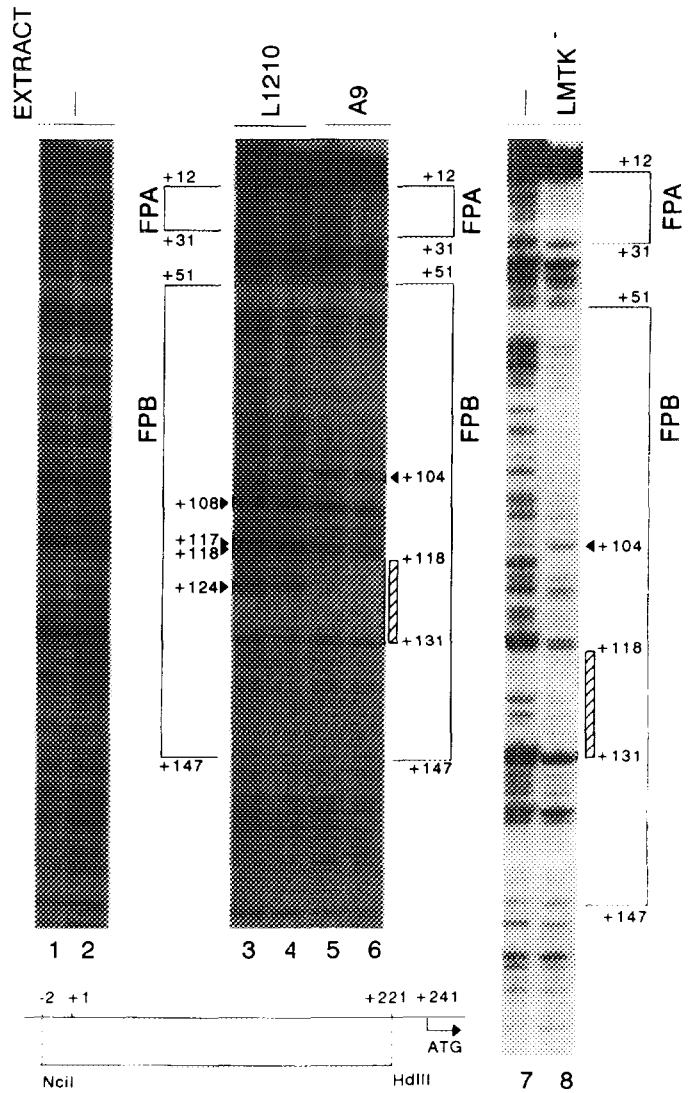

Figure 1. Footprinting analysis of the 5'UTR of the α_4m gene. A probe ranging from -2 to +221 was digested with 1 and 2 U DNaseI in the absence (lanes 1,2,7) or presence of 50 µg nuclear extract from the cell lines L1210 (lanes 3,4), A9 (lanes 5,6) and LMTK$^-$ (lane 8). Differences in footprinting pattern between L1210 and A9 and LMTK$^-$ are indicated with arrowheads or shaded regions.

sites and the intensity of the footprints varies with L1210 versus A9 and LMTK$^-$. Most differences were located in the region which had a major influence on luciferase activity (+113 to +148): 1) positions +117 and +118, which subdivide FPB in L1210 in two parts, were protected in A9 and LMTK$^-$; 2) the region +118 to +131 was more strongly protected in A9 and LMTK$^-$ as compared to L1210; 3) a hypersensitive site at nt +124 in L1210 was not seen in A9 and LMTK$^-$. These results suggest that tissue-specific factors which interact with the FPB region might have a role in differential expression of α_4m.

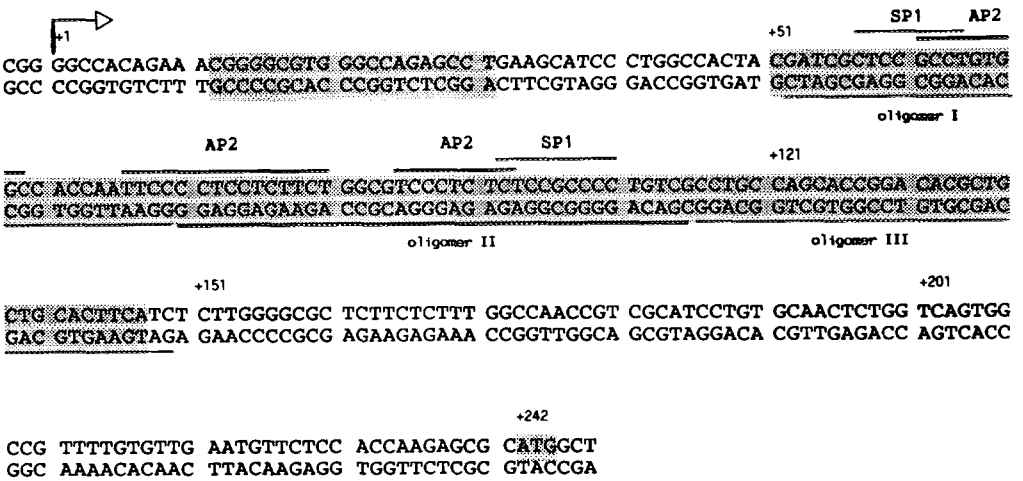

Figure 2. Sequence of the 5'UTR of the α_4m gene. Footprinted areas are shaded and consensus binding sites for transcription factors are indicated above the sequence. The oligomers I, II and III, used in EMSA, are underlined.

Characterisation of the trans-acting factors of FPA

In figure 2, the nucleotide sequence of the 5'UTR of the α_4m gene is shown and the footprinted regions have been indicated. Several consensus sequences for putative transcription factors were identified in the region +55 to +114. In order to identify the trans-acting factors interacting with FPB, 3 sets of double-stranded oligomers were synthesised which cover this region (see fig 2). They were used in EMSA with nuclear extracts of the cell lines L1210, A9 and LMTK$^-$.

When labelled oligomer I (+52 to +79) was incubated with nuclear extract from the α_4m expressing cell line L1210, a single complex MI-1 (fig 3A, lane 1) was observed. Two complexes MI-1 and MI-2 (fig 3A, lanes 5 and 9) were seen after incubation with nuclear extracts derived from the non-expressing cell lines A9 and LMTK$^-$. MI-1 is probably an SP1 complex since an oligomer known to bind SP1 (14) specifically competed with MI-1 (fig 3A, lanes 3 and 7). The reverse experiment, with the SP1 oligomer as a probe, confirmed that extracts from L1210 and A9 cells contained an SP1-binding activity that was competed with cold oligomer I (data not shown). Therefore oligomer I is recognised by SP1. The strong competition of the AP2 oligomer (15) for MI-2 in A9 and LMTK$^-$ extracts indicated that the MI-2 complex might contain AP2 or an AP2-like protein. The cross-competition of the AP2 oligomer with MI-1 in L1210, A9 and LMTK$^-$ nuclear extracts is most likely due to the presence of a consensus SP1 site in the AP2 oligomer. When labelled AP2 oligomer was used, the same binding pattern was obtained as with oligomer I and binding was specifically competed with cold oligomer I (fig 3B). However, a supershift experiment with the polyclonal AP2 antibody C-18 showed that the antibody did not affect the mobility of the MI-2 complex in contrast to the complex formed between oligomer I and a commercial AP2 extract (fig 3B). Moreover, the complex formed with the AP2 extract showed a higher mobility than MI-2. This indicated that the factor binding to oligomer I was not AP2 but may be an AP2-like factor which is not recognised by C-18.

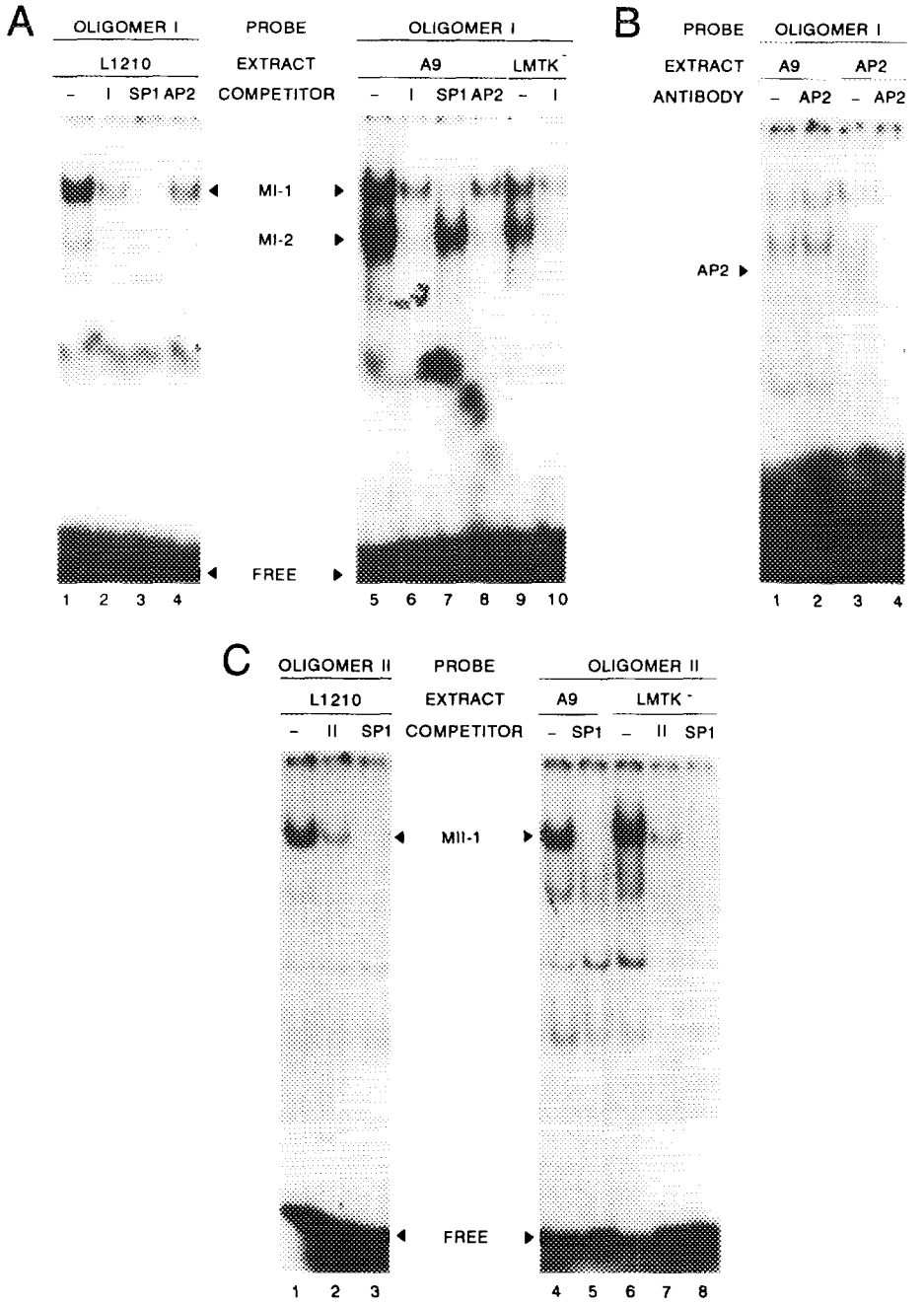

Figure 3. EMSA with oligomer I (A,B) and oligomer II (C). (A) shows the gel retardation pattern of oligomer I with extract from L1210 (lanes 1-4), A9 (lanes 5-8) and LMTK- (lanes 9,10). Free probe and the two retarded complexes (MI-1 and MI-2) are indicated. In competition experiments, an excess of 10 pmol cold oligomer I (lanes 2,6,10), SP1 oligomer (lanes 3,7) or AP2 oligomer (lanes 4,8) was added. In (B) the complex retarded by oligomer I with A9 extract (lane 1) is compared with commercial AP2 extract (lane 3). Addition of an AP2 antibody (C-18) shifted the complex formed with AP2 extract (lane 4) but not complex MI-2 (lane 2). (C) represents the retardation experiments of oligomer II with extracts from L1210 (lanes 1-3), A9 (lanes 4,5) and LMTK- (lanes 6-8). Competition experiments with 10 pmol cold oligomer II (lanes 2,7) and SP1 oligomer (lanes 3,5,8) are shown.

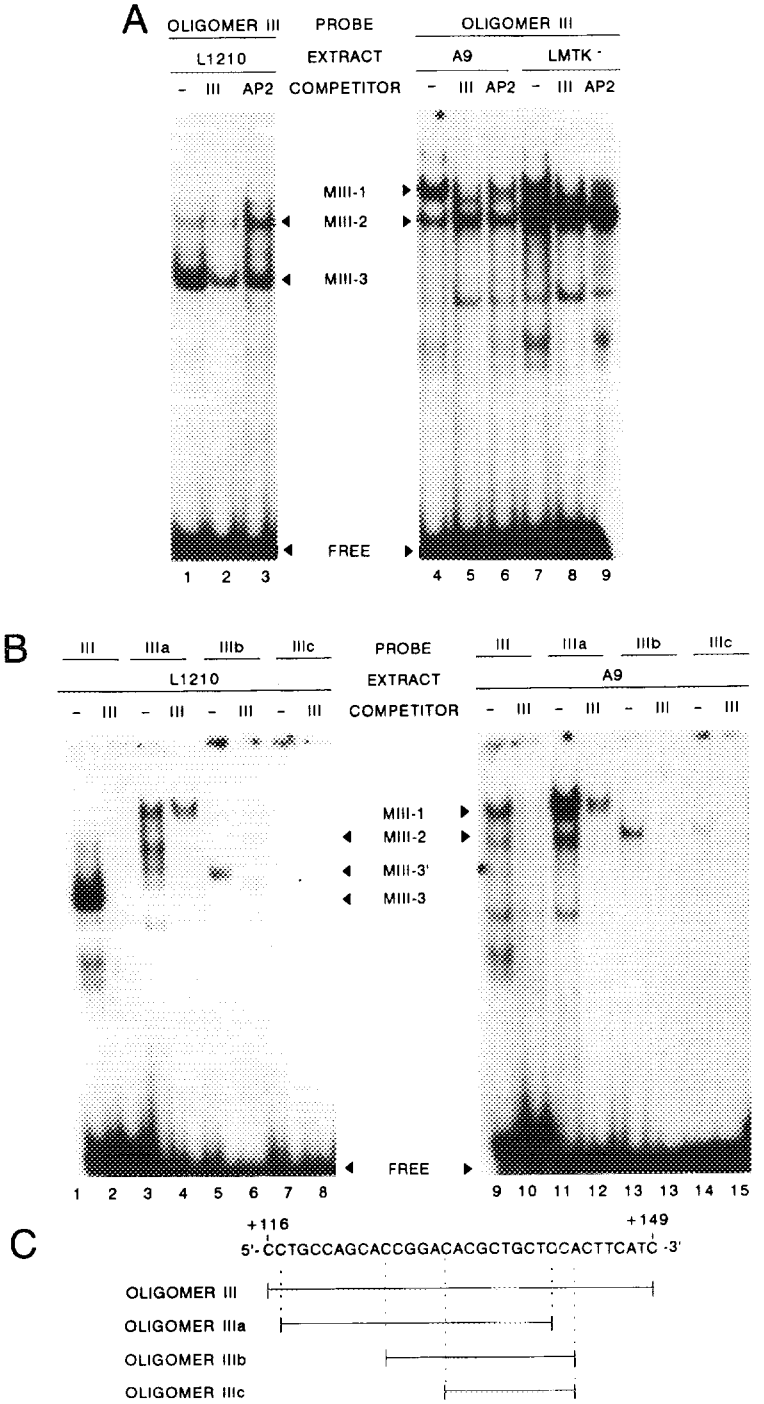

Figure 4. Mobility shift assays of oligomer III with extracts from L1210 (lanes 1-3), A9 (lanes 4-6) and LMTK⁻ (lanes 7-9) are shown in A. Excess of cold oligomer III inhibits formation of complex MIII-3 (lane 2) and complex MIII-1 (lanes 5, 9) but has no effect on complex MIII-2. Addition of 10 pmol cold AP2 oligomer did not influence the binding (lanes 3,6,9). (B) shows EMSA with smaller oligomers derived from oligomer III as indicated in (C). With extract from L1210, only oligomer IIIb retarded a specific complex which corresponds to MIII-3 but with a lower mobility (indicated as MIII-3', lane 5). Complex MIII-1, with A9 extract, was only seen with oligomer IIIa (lane 11).

Oligomer II, corresponding to a region that showed differences between L1210 and A9 cell lines in DNaseI footprinting, retarded only one complex (MII-1, fig 3C) with extracts from both expressing and non-expressing cell lines. It was also related to SP1, because of its competition with the SP1 oligomer. Therefore, the consensus SP1 site at position nt +104 to +109 must be recognised by this transcription factor. The reverse experiment (labelled SP1 oligomer) again supported this observation (data not shown). Oligomer III, derived from nt +116 to nt +149, formed two complexes with nuclear extracts from A9 or LMTK$^-$ (MIII-1 and MIII-2) and two complexes with extracts from L1210 (MIII-2 and MIII-3, fig 4A). This result is in agreement with the differences seen in FPB in footprinting experiments in this region for the expressing versus non-expressing cell lines. Complex MIII-2, common to all cell lines, must be aspecific because it was not competed with cold oligomer III. MIII-1 was only present in non-expressing (A9 and LMTK$^-$) cell lines and MIII-3 in the expressing cell line (L1210). Since no consensus binding sites for a known transcription factor could be found in oligomer III, smaller oligomers derived from oligomer III (as indicated in fig 4C) were used to delineate the exact binding sites of the cell type-specific factors. The binding region for MIII-3 is located between nt +126 and +142 (fig 4B). The region which is essential for the formation of complex MIII-1 extends from nt +119 to nt +135 (fig 4B and data not shown). Although the mobility of the complex formed with oligomer IIIb in L1210 (MIII-3') is lower than the mobility of complex MIII-3, both oligomers III and IIIb seem to be recognised by the same nuclear factor since binding of this protein to oligomer IIIb was inhibited by oligomer III (fig 4B) and vice versa (data not shown).

DISCUSSION

Transfection experiments have shown that the 5'UTR of the α_4m gene has a major influence on promoter activity (7). In this study we confirmed this observation and showed that the region between nt +113 and +148 contains an element with a strong regulatory role in α_4m expression. Although this is not a usual feature for RNA polymerase II-driven genes, a number of genes have already been shown to contain downstream elements (16-18). The region +113 to +148 showed major differences in the banding pattern between expressing and non-expressing cell lines in footprinting experiments. This indicates the binding of a transcription factor to this region and suggests the involvement of a tissue-specific factor. In EMSA, a cell type-specific factor in complex MIII-3, seen with nuclear extracts from L1210, bound to the region +126 to +142 whereas the factor MIII-1 from A9 bound to nt +119 to +135. The partially overlapping binding sites reflect the differential DNaseI protection at nt +117 and + 118 in L1210 versus A9. The region containing the binding site of MIII-3 had a strong regulatory influence on promoter activity in transfection experiments. Thus, the protein(s) present in the MIII-3 complex are probably involved in efficient and cell type-specific promoter activity. However, the exact nature of the factors MIII-1 and MIII-3 remains unknown since no homology with consensus sequences for known transcription factors could be found.
The extended size of FPB (nt +51 to +147) suggests that a combination of transcription factors may bind to the 5'UTR of α_4m. EMSA showed that the two SP1 consensus sites, at nt +57 to +64 and + 102 to +110, are recognised by the SP1 transcription factor with extracts from both expressing and non-expressing cell lines. It has been shown that SP1 is required to sequester the transcription apparatus in some cases of TATA-less promoters (19) which may also be true for α_4m. Although such observations have thus far only

been made for SP1 sites upstream from the capping site, it is possible that downstream SP1-sites preserve this activator function. The presence of a functional TATA-box downstream of the transcription start site has been documented in the Ad IVa2 gene (20).

An AP2-like protein, only present in the non-expressing cell lines A9 and LMTK$^-$, is able to bind the AP2 consensus sequence at position +95 to +103. Although AP2 normally functions as a transcriptional activator (21), an alternatively spliced form of AP2, which inhibits activation by blocking of sequence-specific DNA binding of AP2, has been described (22). Therefore it is not unlikely that other members of the AP2 family might exist which exert an inhibitory function. Although the binding site of the AP2-like protein overlaps with an SP1 binding site, we showed binding to both sites *in vitro*. However, if *in vivo* the AP2-like factor has a higher affinity for its binding site than SP1, AP2 could have an inhibitory effect on α_4m expression by preventing binding of SP1.

In conclusion, the 5'UTR of the α_4m gene is recognised by a combination of ubiquitous (SP1) and cell type-specific factors (an AP2-like factor and MIII-1 for non-expressing cell lines versus MIII-3 for expressing cell lines respectively). The region containing the binding site for the cell type-specific factors MIII-3 (+126 to +142) and MIII-1 (+119 to +135) is indeed necessary for efficient α_4m promoter activity.

REFERENCES

1. Hemler, M.J. (1991) In Receptors for Extracellular Matrix Proteins (R.P. Mecham and J.A. McDonald, Ed.), pp. 255-299. Academic Press, San Diego.
2. Hynes, R.O. (1992) Cell 69, 11-25.
3. Holzmann, B., McIntyre, B.W., and Weissman, I.L. (1989) Cell 56, 37-46.
4. Berlin, C., Berg, E.L., Briskin, M.J., Andrew, D.P., Kilshaw, P.J., Holzmann, B., Weissman, I.L., Hamann, A., and Butcher, E.C. (1993) Cell 74, 185-195.
5. Rosen, G.D., Sanes, J.R., Lachance, R., Cunningham, J.M., Roman, J., and Dean, D.C. (1992) Cell 69, 1107-1119.
6. Jaspers, M., Wu, R.R., Van der Schueren, B., and Cassiman JJ. (submitted) Differentiation.
7. De Meirsman, C., Schollen, E., Jaspers, M., Ongena, K., Matthijs G., Marynen P., and Cassiman, J.J. (1994) DNA and Cell Biol. 13 (7), 743-754.
8. Audet, J.-F., Masson, J.-Y., Rosen, G.D., Salesse, C., and Guerin, S.L. (1994) DNA and Cell Biol. 13 (11), 1071-1085.
9. Rosen, G.D., Barks, J.L., Iademarco, M.F., Fisher, R.J., and Dean, D.C. (1994) Jour. Biol. Chem. 269 (22), 15652-15660.
10. Nordeen, S.K. (1988) BioTech. 6, 454-457.
11. Matthijs, G., Cassiman, J.J., Van den Berghe, H., and Marynen, P. (1994) BBRC 202 (1), 65-72.
12. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) In Molecular Cloning: A Laboratory Manual, 2nd ed.. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. Shapiro, D.J., Sharp, P.A., Wahli, W.W., and Keller, M.J. (1988) DNA 7 (1), 47-55.
14. Briggs, M.R., Kadonaga, J.T., Bell, S.P., and Tjian, R. (1986) Science 234, 47-52.
15. Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) Genes and Dev. 2, 1557-1569.
16. Chen, R., Ingraham, H.A., Treacy, M.N., Albert, V.R., Wilson, L., and Rosenfeld M.G. (1990) Nature 346, 583-586.
17. Tamura, T., and Mikoshiba, K. (1991) FEBS 280 (1), 75-78.
18. Biggin, M.D., and Tjian, R. (1988) Cell 53, 699-711.
19. Pugh, B.F. and Tjian, R. (1990) Cell 61, 1187-1197.
20. Carcamo, J., Maldonado, E., Cortes, P., Ahn, M.-H., Ha, I., Kasai, Y., Flint, J., and Reinberg, D. (1990) Genes and Dev. 4, 1611-1622.
21. Williams, T., and Tjian, R. (1991) Genes and Dev. 5, 670-682.
22. Beuttner, R., Kannan, P., Imhof, A., Bauer, R., Yim, S.O., Glockshuber, R. Van Dyke M.W., and Tainsky M.A. (1993) Mol. Cell. Biol. 13 (7), 4174-4185.