Functional analysis of the endothelial cell-specific Tie2/Tek promoter identifies unique protein-binding elements

Bahaa M. FADEL*†§, Stephane C. BOUTET*†§ and Thomas QUERTERMOUS*‡§‖

*Division of Cardiology, Vanderbilt University Medical Center, 315 MRB II, Nashville, TN 37232-6300, U.S.A., ‡Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, 315 MRB II, Nashville, TN 37232-6300, U.S.A., and †The Veterans Administration Medical Center, 1310 24th Avenue South, Nashville, TN, 37212, U.S.A.

To investigate the molecular basis of endothelial cell-specific gene expression, we have examined the DNA sequences and the cognate DNA-binding proteins that mediate transcription of the murine *tie2/tek* gene. Reporter transfection experiments conformed with earlier findings in transgenic mice, indicating that the upstream promoter of Tie2/Tek is capable of activating transcription in an endothelial cell-specific fashion. These experiments have also allowed the identification of a single upstream inhibitory region (region I) and two positive regulatory regions (regions U and A) in the proximal promoter. Electrophoretic mobility-shift assays have allowed further characterization of three novel DNA-binding sequences associated with these regions and have provided preliminary characterization of the protein factors binding to these elements. Two of the elements (U and A) confer increased transcription on a heterologous promoter, with element U functioning in an endothelial-cell-selective manner. By employing embryonic endothelial-like yolk sac cells in parallel with adult-derived endothelial cells, we have identified differences in functional activity and protein binding that may reflect mechanisms for specifying developmental regulation of *tie2/tek* expression. Further study of the DNA and protein elements characterized in these experiments is likely to provide new insight into the molecular basis of developmental- and cell-specific gene expression in the endothelium.

INTRODUCTION

The endothelium occupies a pivotal position at the interface between the circulating humoral and cellular elements of the blood and the solid tissues that constitute the various organs. In this unique position, endothelial cells regulate a large number of critical processes. Such processes include leukocyte adherence and transit through the blood vessel wall, local control of blood vessel tone, modulation of the immune response, and the balance between thrombosis and thrombolysis. It is not surprising that endothelial cell dysfunction has been postulated as a central feature of vascular diseases such as hypertension and atherosclerosis.

The endothelial cell has also been the object of research investigating the developmental processes regulating the formation of blood vessels [1]. Classical embryological studies have shown that most embryonic blood vessels develop through a process termed vasculogenesis, characterized by the *in situ* differentiation and organization of endothelial precursor cells into capillary-like structures [1–5]. The formation of vessels by sprouting and branching from existing vessels is termed angiogenesis and has also been shown to contribute to vascular formation in specialized organs of the embryo such as brain and kidneys [4,6]. Recent studies employing targeted mutations of several endothelial cell-specific tyrosine kinase receptors, including Flk1, Flt1, Tie1 and Tie2/Tek, have provided new insight into early endothelial cell development [7–10]. A targeted mutation of Flk1 causes a failure of differentiation of precursor cells into a more mature endothelial cell phenotype, resulting in a lack of blood vessel development and defective vasculogenesis [7]. In contrast, mice lacking Flt1 have an increased number of endothelial cells with disorganized and dilated vascular networks. This suggests that Flt1 plays a role in the organization of the embryonic vasculature once endothelial cell differentiation has occurred [8]. A targeted mutation of Tie1 results in less dramatic morphological abnormalities; however, significant functional abnormalities are evident, characterized by diffuse oedema and localized haemorrhage due to a 'leakage syndrome' [10]. The *tie2/tek* knock-out mouse also has a characteristic phenotype with a diminished number of endothelial cells and abnormal branching and sprouting of blood vessels, suggesting a role for Tie2/Tek factor in the process of angiogenesis [10]. Thus each of these factors plays a distinct and unique role in the development of a normal vascular system. The normal expression of each is required for endothelial cell functions that regulate lineage development, vascular morphogenesis and the organization and integrity of blood vessels once they are formed.

Despite the significant contribution of these studies to the understanding of vascular development, most of the data remain descriptive. Further interpretation of the phenotype of these genetic models and further understanding of vascular development in general will require greater knowledge of the origin and differentiation of the endothelial cell lineage. To pursue these fundamental questions in endothelial cell biology, this laboratory and others have investigated mechanisms of transcriptional regulation in this cell type. By comparison with experiments in

Abbreviations used: BAEC, bovine aortic endothelial cell; C/EBP, CCAAT-box-binding proteins; DTT, dithiothreitol; EMSA, electrophoretic mobility-shift assay; RLU, relative light unit; rp, ribosomal protein; SV40, simian virus 40; UTR, untranslated region; VSMC, vascular smooth muscle cell; vWf, von Willebrand factor; YSC, yolk salk cell.

§ Present address: Falk Cardiovascular Research Center, Stanford University, 300 Pasteur Drive, Stanford, CA 94305, U.S.A.

‖ To whom correspondence should be addressed.

the myogenic and haematopoietic lineages, such work has aimed at the characterization of cell-specific transcription factors. Toward this end, a large number of studies have investigated the *cis*-acting DNA elements and *trans*-acting protein factors that regulate transcription of the following endothelial cell genes: plasminogen activator inhibitor-1, granulocyte–macrophage colony stimulating factor, endothelin-1, thrombomodulin, von Willebrand factor (vWf), platelet-derived growth factors A and B, intracellular cell adhesion molecule-1, vascular cell adhesion molecule-1, E-selectin, P-selectin, tissue factor, angiotensin-converting enzyme, inducible and constitutive NO synthase, superoxide dismutase, Flk1 and Flt1 [11]. This work has taught us surprisingly little about the fundamental mechanisms that mediate cell-specific transcription in endothelial cells.

A limitation of studies to date is that they have been conducted with promoter sequences that cannot direct endothelial cell-specific expression in transgenic animals, or have not been evaluated in transgenic animals. The probability of identifying cell-specific elements should be greatly increased if experiments *in vitro* are conducted with a minimal cell-specific region as delineated *in vivo*. Promoter regions of four endothelial cell genes have been shown to direct cell-restricted expression in the endothelium in transgenic mice: the murine endothelin-1 promoter, the human vWf promoter and the promoters of two homologous murine tyrosine kinase receptors, Tie1 and Tie2/Tek [12–15]. The murine endothelin-1 promoter also directs transcription in epithelial cells, making it a less desirable gene for study. The vWf promoter targets expression of the transgene only to a subpopulation of endothelial cells limited to the yolk sac and adult brain. While reporter expression in transgenic animals directed by the Tie1 and Tie2/Tek promoters is not as universal in embryonic endothelial cells as the endogenous promoters, expression is seen in a majority of the endothelium throughout most of development. Since the Tie2/Tek promoter has been most extensively investigated in transgenic animals, this promoter was chosen for experiments *in vitro* aimed at identifying the DNA and protein elements that direct cell-specific transcription in endothelial cells.

Here, we report the cloning and initial characterization of the 5′ regulatory region of the murine *tie2/tek* gene. Reporter gene transfection studies reveal that the upstream promoter and first untranslated exon contain a single inhibitory sequence and multiple positive regulatory elements. Footprinting analysis and electrophoretic mobility-shift assays (EMSA) have allowed an initial definition of the protein-binding sequences in these regions and a preliminary evaluation of the interacting proteins. While each of these sites also binds proteins in non-endothelial cells, one element (element B) appears to bind factors that are developmentally regulated in endothelial cells. Another novel element (element U) is capable of activating transcription of a heterologous promoter in an endothelial cell-selective manner. These data reveal the surprising complexity of this endothelial cell-specific promoter and provide for more definitive studies investigating the molecular nature of the factors that regulate transcription of this gene in the endothelium.

MATERIALS AND METHODS

Cloning and mapping of *tie2/tek* regulatory regions

Using reverse-transcriptase PCR of murine yolk sac cell (YSC) RNA, two PCR primers 5′-CTTGTAAACAAGAGCGAGTGGAC and 5′-CCATTCTCTGGTCACATCTTGAG were used to amplify 369 bp spanning the 5′ end of the *tie2/tek* cDNA and a part of the 5′-untranslated region (5′-UTR). This DNA fragment was employed to screen a 129SV/J mouse genomic library constructed in Lambda Fix II (Stratagene, La Jolla, CA, U.S.A.). Clones thus isolated were restriction mapped and exons were identified by hybridization. Restriction fragments were subcloned into Bluescript II KS (Stratagene, La Jolla, CA, U.S.A.) for further mapping and nucleotide sequence analysis. Dideoxy chain-termination sequencing was performed with Sequenase 2.0 according to the manufacturer's instructions (Amersham, Cleveland, OH, U.S.A.). Each DNA was sequenced at least twice from both directions and sequence analyses were performed using the GCG software package (Genetics Computer Group, Madison, WI, U.S.A.).

Reporter gene construction

For functional analysis of the upstream regulatory region, a series of 5′ deletion constructs was generated. A *Hin*dIII–*Hin*dIII and a *Hin*dIII–*Bam*H1 restriction fragment spanning 2.3 kb and 471 bp of the Tie2/Tek promoter and 5′-UTR respectively were subcloned into pGL2 basic vector (Promega, Madison, WI, U.S.A.), linearized with *Sac*1 and subjected to exonuclease III and S1 nuclease digestion, blunt-ended with T4 DNA polymerase and re-ligated. The sequence of each deletion construct was determined by nucleotide sequence analysis. For analysis of the two regulatory regions U and A in association with a heterologous promoter, oligonucleotides encoding binding sequences were concatamerized and cloned upstream of the minimal simian virus 40 (SV40) promoter in the pGL2 promoter vector (Promega). Single-stranded oligonucleotides (regions A and U) were annealed and multimerized as previously described [16]. All constructs were subjected to DNA sequencing to confirm the size and sequence of the end product.

Cell culture and transfection studies

The murine YSC used in this study were supplied by Progenitor Inc. (Menlo Park, CA, U.S.A.), with culture medium consisting of alpha-minimal essential medium supplemented with 18 % (v/v) fetal bovine serum, leukaemia inhibitory factor, 2-mercaptoethanol and gentamycin [17,18]. Bovine aortic endothelial cells (BAEC) and rat aortic vascular smooth muscle cells (VSMC) were harvested and cultured as previously described [19]. HeLa, HepG2 (human hepatoma), 3T3 (mouse fibroblast), C6 (rat glioma), LLC1 (rat lung carcinoma), 143B (human osteosarcoma), C2C12 (myoblast) and JEG-3 (human choriocarcinoma) cell lines were obtained from American Type Culture Collection (Rockville, MD, U.S.A.) and grown in Dulbecco's modified Eagle's medium supplemented with 10 % (v/v) fetal bovine serum, penicillin (100 units/ml), streptomycin (100 μg/ml) and glutamine (2 mM). All media and tissue culture reagents were purchased from Life Technologies (Grand Island, NY, U.S.A.).

Transient transfection of all cell lines was performed using lipofectamine according to the manufacturer's instructions (Life Technologies), except for 3T3 fibroblasts which were transfected using the calcium phosphate method [20]. Briefly, cells were seeded in six-well plates at approximately 4×10^5 cells per well in 2 ml of medium. Twenty-four hours later at 50–70 % confluence, cells were washed twice with PBS and co-transfected with 1 μg of pGL2 plasmid and 1 μg of control SV40/LacZ plasmid. For the lipofectamine method, 6 μl of lipofectamine was incubated with DNA in a 210 μl Optimem (Life Technologies) for 15 min at 37 °C and cells were transfected for a period of 4 h, after which the transfection solution was removed and replaced with fresh medium. Forty-eight hours after transfection, cells were washed twice with PBS and lysed in 140 μl of cell culture lysis reagent (Promega). The lysate was vortexed, subjected to one freeze–thaw

Table 1 Activity of the Tie2 promoter in three non-endothelial cell lines as compared with activity in BAEC

Activity is given as relative luciferase activity in each of the three non-endothelial cell lines divided by the relative luciferase activity in BAEC for the same reporter plasmid construct.

Cells	Activity*											
Construct ...	−2000†	−153	−105	−96	−68	−65	−46	−35	−31	+107	+233	+255
HeLa	0.04	0.03	0.03	0.08	0.06	0.13	0.09	0.19	0.14	0.06	0.05	0.03
$HepG_2$	0.01	0.02	0.02	0.02	0.05	0.06	0.05	0.05	0.06	0.04	0.02	0.02
VSMC	0.03	0.03	0.03	0.08	0.06	–†	–†	0.16	–†	–†	–†	–†

* Indicates the reporter vector construct p( )tek-luc used in the transfection experiments. These constructs are identical with those used in Figure 2.
† Indicates that the reporter vector construct was not tested in VSMC.

Table 2 DNA sequence of the wild-type and mutant oligonucleotides used in EMSA

DNA sequences representing the binding site, or that are part of the binding site, are in bold. The binding sites sharing high similarity, i.e. *tie2* region A with CAAT consensus, *tie2* region B with *flk1* and *flt1*, are grouped together. The 12 bp sequence of *tie2* region U represents the binding site with *tie2* region U_2 representing a longer DNA sequence spanning region U.

Binding site	Wild-type DNA sequence	Location (bp)*	Mutant DNA sequence†
tie2 region A	GATCGGCTTGTTATT**CAATTC**CTGGCCTATGA	−79 to −51	GATCGGCTTGTTATTGTCTTCCTGGCCTATGA
CAAT consensus	TGCAGATTGCG**CAATCT**GGA		TGCAGAGACTAGTCTCTGGA
tie2 region U	**GCATACCATACA**	−100 to −88	GCGATCCACGAA
tie2 region U_2	GATCTCCTCATC**GCATACCATACA**TAGGT	−108 to −83	
tie2 region B	GATCCTATTGTTCC**TGAAAATG**CTGACCAG	−41 to −15	
flk1 sequence	GATCAGCTCTGCTC**TGAAAAGG**GGCATGG	−744 to −719	
flt1 sequence	GATCTCAATGCGGC**CGAAAAAG**ACACGGAC	−704 to −678	

* Location of the wild-type oligonucleotide relative to the transcription start site of the corresponding gene.
† Underlined nucleotides represent the bp mutations.

cycle and briefly centrifuged and 20 μl of the supernatant was assayed for luciferase and β-galactosidase activity. Relative light units (RLU) were calculated as the ratio of light units obtained with the Tie2/Tek promoter construct divided by the value obtained with the promoterless pGL2 basic plasmid, and the resulting number was divided by absorbance obtained in the β-galactosidase assay. For HeLa, HepG2 and VSMC, transcriptional activity is expressed as the ratio of RLU in these cells compared with RLU measured in BAEC (Table 1). Transfections were done in triplicate and conducted a minimum of twice with each construct.

DNAse I footprinting and EMSA

Nuclear extracts for both footprinting and EMSA were prepared from different cell lines according to the method of Dignam et al. [21]. For DNAse I footprinting analysis, the *Pst*1–*Sty*1 restriction fragment of the Tie2/Tek promoter was dephosphorylated, end-labelled and digested with *Bam*H1 and the *Bam*H1–*Sty*1 fragment was used as a probe for the footprinting reaction of the non-coding strand of the promoter. Binding reaction mixtures consisted of 50 mM NaCl, 10 mM Hepes (pH 7.9), 1 mM $MgCl_2$, 1 mM $CaCl_2$, 100 μg BSA, 1.4 mM 2-mercaptoethanol, 10% (v/v) glycerol, 5 μg poly(dI-dC) and 40 μg of YSC nuclear extract in a final volume of 50 μl. After digestion with DNAse I (Worthington Biochemical Corporation, Freehold, NJ, U.S.A.) at a final concentration of 40 units/ml for 30 s at 4 °C, the product was electrophoresed on a 7% polyacrylamide gel. Maxam–Gilbert sequencing was performed as previously described [20].

The oligonucleotides used in EMSA are listed in Table 2. These oligos were end-labelled with [^{32}P]ATP using T_4 polynucleotide kinase (New England Biolabs, Beverly, MA, U.S.A.). Two conditions were used for binding reactions (20 μl total volume). Condition 1 (medium salt, MS) consisted of 40 mM NaCl, 50 mM KCl, 10% (v/v) glycerol, 10 mM Tris/HCl (pH 7.5), 2 mM dithiothreitol (DTT), 2 μg poly(dI-dC) and 0.8 mg/ml BSA. Condition 2 (medium salt with Hepes and magnesium, HM) consisted of 40 mM NaCl, 60 mM KCl, 5 mM $MgCl_2$, 10 mM Hepes (pH 7.8), 1 mM EDTA (pH 8.0), 2 μg poly(dI-dC) and 1 mM DTT. For EMSA in Figure 4 (element U), condition HM was used and electrophoresis was performed with 0.5 × Tris/borate/EDTA running buffer. For EMSA in Figure 5 (element A), binding was conducted with condition MS, and electrophoresis was performed in 1 × Tris/glycine buffer [20]. For EMSA in Figure 6 (element B), binding condition MS was used, and samples were electrophoresed in 0.5 × Tris/borate/EDTA running buffer. Nuclear extracts (7.5 μg) from the various cell lines were incubated with the binding reaction mixture for 10 min at room temperature. Labelled probe (0.2–0.5 ng) and excess cold oligonucleotide (100-fold unless otherwise indicated) were added and reactions were incubated at room temperature for an additional 20 min before being loaded onto the gel and electrophoresed at 200 V at 4 °C for 4–6 h, fixed, dried and exposed to film.

RESULTS

Cloning of the Tie2/Tek promoter region

The cloning of the 5′ portion of the murine *tie2/tek* gene has been reported previously [13]. We have extended the characterization of this gene by cloning and performing detailed

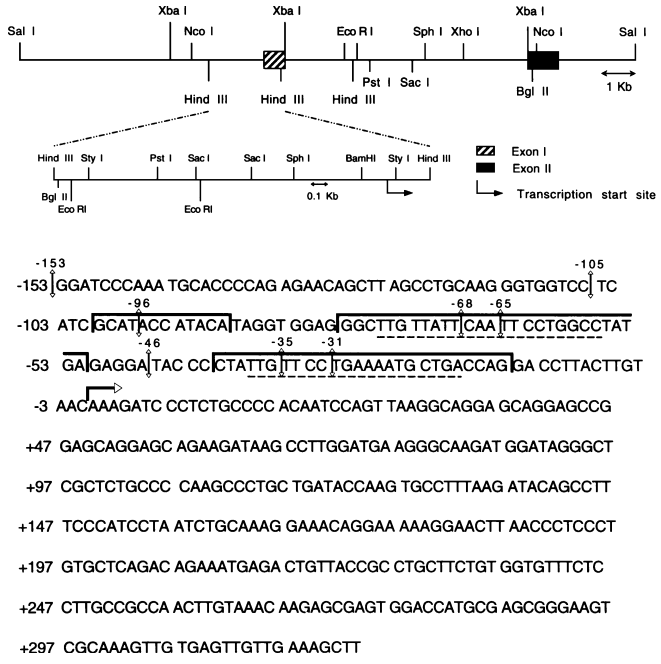
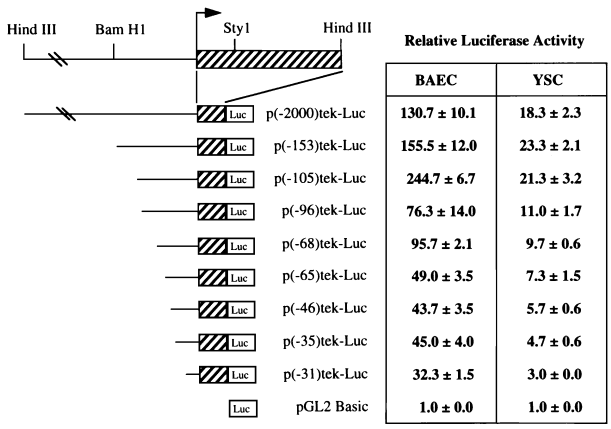


Figure 1 Genomic organization of the murine *tie2/tek* locus

Top: restriction mapping analysis of the 18 kb *tie2/tek* locus and a 2.3 kb region (*Hin*dIII–*Hin*dIII) spanning the proximal promoter and the 5′-UTR. The first exon is indicated by a hatched box, the second exon by a black box and the transcription start site by an arrow. Bottom: nucleotide sequence of the immediate upstream promoter region and the 5′-UTR employed in these studies. The transcription start site is indicated by an arrow and the first nucleotide is designated +1. The 5′-ends of reporter constructs containing variable lengths of promoter sequence are indicated by vertical arrows and negative numbers. All of these constructs terminate at the *Hin*dIII site at +318. Sequences identified by DNAse I footprinting are underlined with broken lines and oligonucleotide probes employed in protein binding and enhancer trap experiments are indicated by brackets.



Relative Luciferase Activity

	BAEC	YSC
p(-2000)tek-Luc	130.7 ± 10.1	18.3 ± 2.3
p(-153)tek-Luc	155.5 ± 12.0	23.3 ± 2.1
p(-105)tek-Luc	244.7 ± 6.7	21.3 ± 3.2
p(-96)tek-Luc	76.3 ± 14.0	11.0 ± 1.7
p(-68)tek-Luc	95.7 ± 2.1	9.7 ± 0.6
p(-65)tek-Luc	49.0 ± 3.5	7.3 ± 1.5
p(-46)tek-Luc	43.7 ± 3.5	5.7 ± 0.6
p(-35)tek-Luc	45.0 ± 4.0	4.7 ± 0.6
p(-31)tek-Luc	32.3 ± 1.5	3.0 ± 0.0
pGL2 Basic	1.0 ± 0.0	1.0 ± 0.0

Figure 2 Luciferase activity of various Tie2/Tek reporter plasmids

A series of constructs derived by 5′-nested deletion were evaluated by transfection into BAEC and embryonic YSC. Relative luciferase activity is the luciferase activity of the Tie2/Tek reporter normalized to the promoterless luciferase vector (pGL2 basic) and to the β-galactosidase activity resulting from co-transfection with a constitutive β-gal vector. The increase in transcriptional activity in BAEC following the deletion of the DNA sequence from −153 bp to −105 bp suggests that this region (region I) contains one or more negative regulatory elements. The decrease in transcriptional activity detected in both BAEC and YSC with deletion of sequences from −105 bp to −96 bp and from −68 bp to −65 bp suggests the presence of two positive regulatory regions, regions U and A respectively.

mapping of 18 kb of the genomic sequence. Included in this region are approximately 8 kb of upstream sequences, the first and second exons and the first 7 kb intron (Figure 1, top). Similar to many genes expressed in early embryogenesis, *tie2/tek* lacks a consensus TATA sequence [22]. Sequences containing the functional regions of the promoter are presented in Figure 1, bottom. Numbering is relative to the transcription-initiation site as previously reported [13]. As indicated, reporter constructs resulting from a 5′ deletion series have the *Hin*dIII site at +318 bp as their 3′ end, and their 5′ end is indicated by an arrow and a negative number (Figure 1, bottom).

Reporter transfection experiments

As a first approach to identifying the DNA sequences that regulate transcription of the *tie2/tek* gene, a series of deletion constructs were generated and transfected into adult-derived BAEC, endothelial-like embryonic YSC and several non-endothelial cell types. YSC are derived from murine yolk sac tissue at approximately 8.0 days of embryonic development and are used in these studies as a model of embryonic endothelial cells. These cells express endothelial markers such as Flk1 and Tie2/Tek and are capable of forming microvascular-like tubular structures *in vitro*, even without being plated on Matrigel ([17] and results not shown). Furthermore, these cells are clonal and long-lived in culture, and have high transfection efficiency. Non-endothelial cells include the epithelial cell line HeLa, the hepatoma cell line HepG2 and primarily cultured rat aortic VSMC.

Transfection studies using the Tie2/Tek promoter constructs revealed cell-type-specific expression in cultured endothelial cells. The largest restriction fragment employed in reporter constructs was a 2.3 kb *Hin*dIII–*Hin*dIII fragment containing approximately 2.0 kb of upstream sequence [Figure 1 (top) and Figure 2]. This sequence includes the region that has been shown to direct endothelial cell-specific gene expression in transgenic mice [13]. This 2.3 kb region of the Tie2/Tek promoter was proven to be cell-specific in cultured cells (Figure 2). Transfection of this reporter construct resulted in high levels of luciferase activity in BAEC. While the level of activity of the reporter constructs was considerably less in YSC than in BAEC, it was still significantly higher than in non-endothelial cells, which had levels of activity barely greater than the promoterless pGL2 basic plasmid (Table 1). This cell-specificity *in vitro* was demonstrated for all of the constructs evaluated, indicating that even the most proximal promoter elements may be involved in restricting expression in cultured endothelial cells.

Transfection of the series of upstream deletion constructs identified a number of regulatory regions. Since the transcriptional activity obtained with p(−2000)tek-luc was equivalent to that obtained with the p(−153)tek-luc construct in both BAEC and YSC, detailed analysis of upstream sequences focused on the region between −153 bp and the transcription start site. While activity of the −2000 bp and −153 bp constructs was similar in both BAEC and YSC, there was a significant increase in activity in BAEC when the region between −153 bp and −105 bp was removed from the reporter construct (Figure 2). As is evident from the error bars, the variance in these experiments was extremely low, indicating that the almost 2-fold increase in activity is statistically significant. This increase most likely represents the deletion of sequences that function in this context as a transcriptional repressor. Interestingly, there is no evidence for such inhibitory activity conferred by this sequence in YSC, suggesting that the DNA-binding proteins that confer this inhibition may not be present, or may not be active, in the

embryonic YSC. Also, deletion of the sequence between −153 bp and −105 bp did not affect transcription in non-endothelial cells (Table 1). Inspection of the nucleotide sequence between −153 bp and −105 bp reveals no known consensus DNA-binding sites.

Analysis of more proximal deletion constructs identified two positive regulatory domains in the upstream Tie2/Tek promoter. The construct containing 96 bp of promoter region, p(−96)tek-luc, exhibited less than half the activity of the construct containing 105 bp of the promoter sequence, p(−105)tek-luc (Figure 2). A similar loss of activity was observed in YSC with deletion of this region of promoter sequence (Figure 2). While there is no known consensus DNA-binding site in this region, there is a repeated 5 bp sequence (CATAC) between −99 bp and −90 bp, and one of these repeats is missing in the p(−96)tek-luc construct. A second decrease in luciferase activity was detected in both BAEC and YSC, with deletion of sequences between −68 bp and −65 bp. As is evident from the promoter sequence, the three nucleotides which are not present in the p(−65)tek-luc construct are part of a CAAT sequence. This region has a similar DNA sequence to the consensus CAAT motif that interacts with members of the C/EBP (CCAAT-box-binding protein) family of transcription factors [23]. Further deletion of sequences between −65 bp and −31 bp did not result in any significant change in transcriptional activity in either BAEC or YSC (Figure 2).

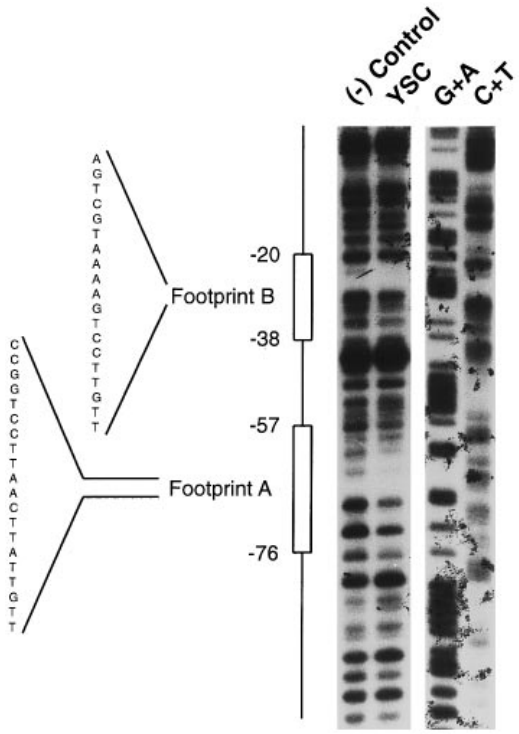


Figure 3 DNAse I footprinting of the non-coding strand of the proximal Tie2/Tek promoter region

End-labelled DNA was subjected to DNAse I digestion in the presence and absence of YSC nuclear extract as detailed in the Materials and methods section. Maxam–Gilbert G+A and C+T sequencing reactions were conducted with the same labelled DNA fragment. Region A, previously identified in transfection studies, is noted as Footprint A. Region U, also identified in transfection studies, is not evident on DNAse footprinting analysis. However, a previously unidentified protein-binding site, Footprint B, is seen and extends from −38 bp to −20 bp.

Protein-binding experiments

To obtain an initial correlation of protein–DNA interactions with these functional observations, DNAse I footprinting was conducted *in vitro*. A DNA fragment containing the transcription start site and the upstream promoter region (−153 bp to +67 bp) was end-labelled and subjected to DNAse I digestion in the presence or absence of YSC nuclear extract. These experiments revealed protection from DNAse I digestion in two regions of the Tie2/Tek promoter (Figure 3). The most 5′ footprinted region, region A, extends from bps −76 to −57, and a more proximal region, region B, extends from bps −38 to −20. Region A includes the functional sequence between bps −68 and −65 identified in earlier transfection experiments [Figure 1 (top) and Figure 2]. Information from reporter gene and footprinting experiments was employed to direct EMSA analysis of DNA binding sites. In order to investigate the pattern of protein binding to the oligonucleotide probes, two different binding conditions were employed, along with two different electrophoresis conditions.

Initial EMSA studies investigated the functional region between −105 bp and −96 bp, designated region U (Table 2). These experiments revealed specific interaction with binding protein present in nuclear extract derived from YSC (Figure 4). To define further the binding sequence, a probe was constructed with mutations in both of the CATAC repeats (GCATACCATACA → GCGATCCACGAA). This mutU probe was not able to compete for binding to the native sequence and was not able to bind protein when radiolabelled and employed as a probe under various conditions (Figure 4). This 12 bp element defined by these studies has not been previously characterized, suggesting that the protein(s) identified represent a new class of DNA-binding factor(s). This element binds the same protein(s) in YSC, BAEC and HeLa, indicating that the interacting factor(s) is not restricted to endothelial cells (results not shown). Under other binding conditions, two bands were apparent and both bands could be competed out by the addition of excess cold probe (results not shown). It seems most likely that these bands represent binding by protein to one or both of the repeated sequences in the probe, suggesting that CATAC represents the basic binding sequence.

EMSA were also employed to investigate protein interactions with the sequence identified as region A. Since functional studies had indicated an important role for the sequence between −68 bp and −65 bp, probes were designed to contain these bps and a minimal amount of flanking sequence (Table 2). This oligonucleotide was capable of binding proteins present in YSC, BAEC and HeLa cells (results not shown). Analysis of the DNA sequence of this oligonucleotide showed that it contains a CAAT motif that may represent the binding site for members of the C/EBP family of DNA binding proteins. However, excess of cold oligonucleotides encoding a CAAT consensus sequence were not able to compete for binding to this Tie2/Tek promoter element (results not shown).

Additional experiments were employed to characterize the binding sequence in region A as well as the proteins that bind this site. To determine whether the CAAT sequence is a part of the binding site, three point-mutations were introduced in the oligonucleotide (CAA → GTC) (Table 2). This mutA probe appeared to compete partially for binding to the native sequence, so the stoichiometry of this competition was evaluated in more detail. EMSA indicated that the mutant was 5-fold less effective than the native sequence at competing for binding, but was capable of completely inhibiting binding at high (100-fold) molar excess (Figure 5, bottom). More significantly, a radiolabelled mutA probe did not exhibit any specific binding (Figure 5, top).

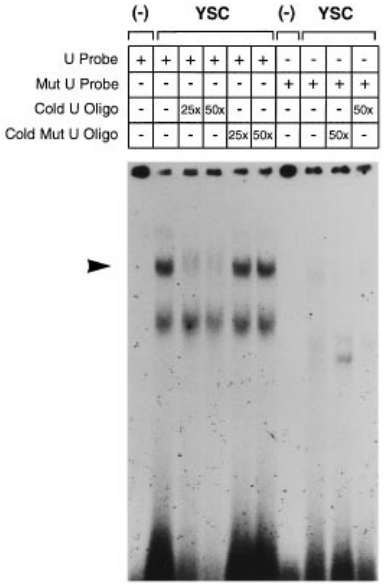


Figure 4 EMSA employing oligonucleotide probes encoding functional region U

Labelled oligonucleotides encoding wild-type and mutant region U sequences (see Table 2) were employed in EMSA with nuclear extract derived from YSC. Under the conditions employed in this experiment, specific binding is represented by a single retarded band indicated by an arrowhead. A cold probe containing mutations in the CATAC repeats is unable to compete for binding and a labelled mutant probe is unable to bind protein.

Further experiments were thus conducted to determine whether this sequence binds members of the C/EBP family of transcription factors. First, oligonucleotides encoding the classical CAAT consensus sequence could not compete for binding to the wild-type region A probe (Figure 5, top). Also, when the consensus CAAT motif was employed as a probe, C/EBP protein binding was noted to produce a much slower-migrating complex. While binding of this complex could be competed out by cold consensus CAAT oligonucleotide, binding was not competed out by cold region A oligonucleotide. Taken together, these data suggest that the CAAT sequence constitutes part of the binding site in region A, but it does not represent a classical CAAT box consensus sequence; i.e. it does not bind C/EBP factors.

EMSA experiments were also conducted using oligonucleotide probes encoding region B, as identified by DNAse I footprinting analysis. The region B probe produced a highly retarded complex in YSC, and a complex of similar size was seen with extracts from HeLa, C2C12 myoblast and myotube cell lines (Figure 6, top). Interestingly, this complex was not seen with BAEC. Instead, the BAEC extract produced a specific and more rapidly migrating complex. This complex was also observed with extracts from all cell types except YSC. With YSC likely to represent a less differentiated endothelial cell phenotype, the differences in the binding patterns may suggest differences in binding proteins between embryonic and adult endothelial cells.

Since region B binds developmentally regulated proteins, we searched for similar sequences in other early endothelial-specific genes. Both Flk1 and Flt1 were found to have similar sequences in their upstream promoter regions (Table 2) [24,25]. Oligonucleotides encoding these sequences were employed in EMSA studies. These probes were neither able to compete with the Tie2/Tek probe for protein binding nor bind protein in YSC nuclear extract (Figure 6, bottom). A search of

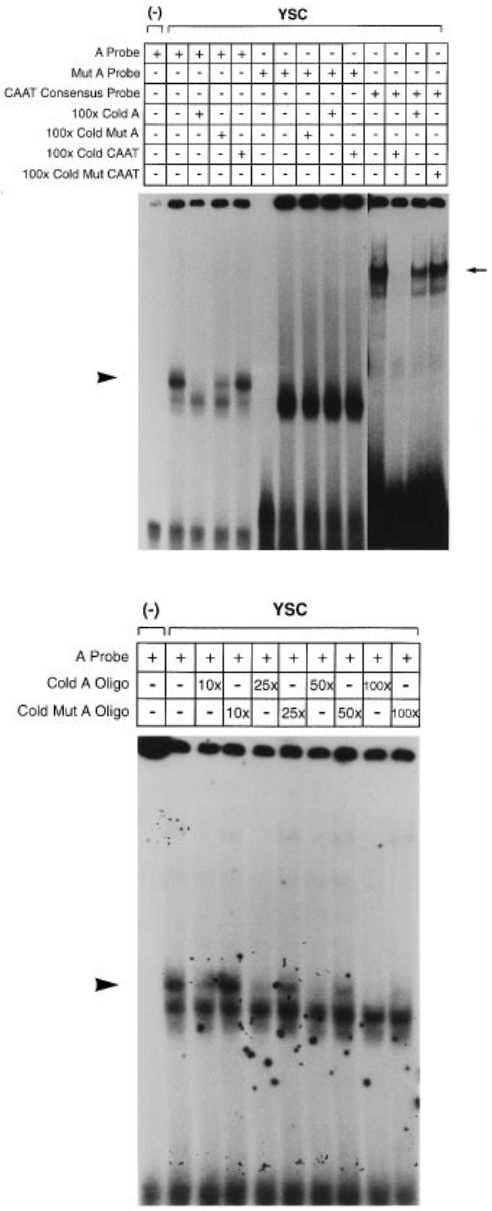


Figure 5 EMSA employing oligonucleotide probes encoding region A identifies specific binding proteins, and these proteins are distinct from those which bind consensus CAAT box sequences

Top: labelled oligonucleotides encoding region A, mutated region A and a consensus CAAT motif were employed in EMSA with nuclear extract derived from YSC. Probe A produces a single specific retarded band under these experimental conditions, as indicated by an arrowhead. Excess of cold probe A is able to compete out protein binding completely, whereas cold mutant probe A can compete out part of the binding, and cold CAAT consensus probe cannot compete out any of the binding. Furthermore, radiolabelled mutant probe A is not able to produce a specific binding pattern. However, the labelled CAAT consensus probe does produce a specific binding pattern, indicated by an arrow, with bands more retarded than those produced with probe A. Bottom: titration competition experiments with probe A and the mutant probe A. A 25-fold excess of the cold probe A is able to compete effectively for all binding to the radiolabelled probe, while a 100-fold excess of cold mutant is required to compete for all binding to the labelled probe A. Specific binding is indicated by an arrowhead.

the various transcription-factor databases failed to identify a known consensus binding sequence in this region. On further analysis, region B is found to contain a TGAAAAT sequence that shares considerable homology with an element, TGA-

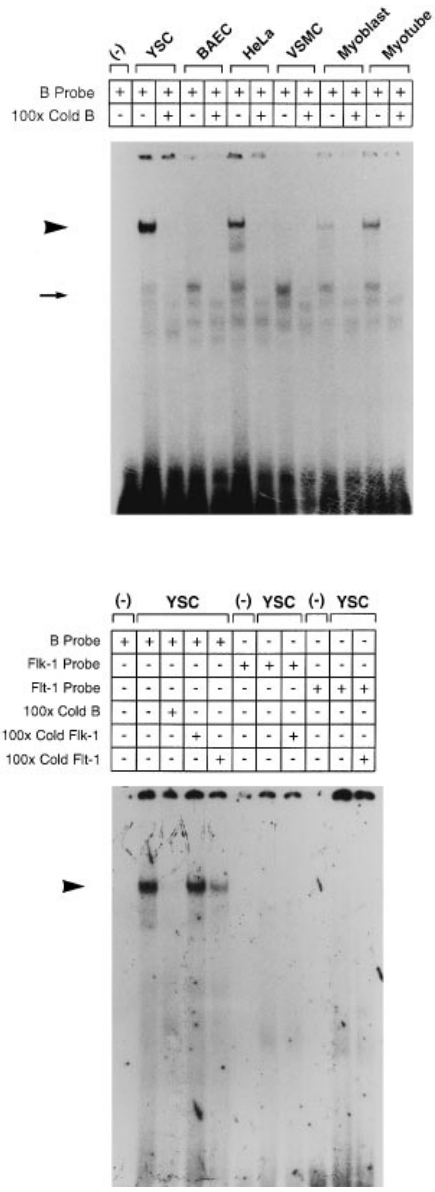


Figure 6 EMSA employing oligonucleotide probes encoding region B

Top: the oligonucleotide probe B detects two specific retarded complexes in HeLa, myoblasts and myotubes. YSC have primarily the larger complex, indicated by an arrowhead, and BAEC have almost exclusively the smaller complex, indicated by an arrow. Bottom: binding to region B is compared with binding by similar sequences in the Flk1 and Flt1 promoter regions. The Flk1 oligonucleotides showed no competition with labelled probe B, and the Flt1 oligonucleotides showed only minimal ability to compete (specific binding indicated by an arrowhead). Neither the Flk1 nor the Flt1 labelled probes were able to bind factors in the YSC extract.

AAAAT, previously characterized in the TATA-less ribosomal protein S16 (rpS16) promoter [26]. Also, the location of this sequence in the Tie2/Tek promoter at approximately −30 bp is identical with the location of the homologous sequence in the rpS16 gene [26].

Earlier experiments in transgenic mice have suggested the presence of important positive regulatory element(s) in the upstream *Sac*1–*Sph*1 region (region S, extending from −753 bp to −537 bp). Transgenes lacking this region fail to demonstrate any significant endothelial expression of the reporter gene, thus

Table 3 Transcriptional enhancement by regions U and A in the context of a heterologous promoter

Results are presented as relative luciferase activity in different cell types. Oligonucleotides encoding four copies of region U or region A were cloned upstream of the minimal SV40 promoter in the plasmid pGL2 promoter and evaluated by transfection into BAEC, 3T3 (fibroblasts), HeLa, C6 (glioma), LLC1 (lung tumour), 143B (osteosarcoma) and JEG3 (choriocarcinoma) cells. Concatamers of region U increase transcription 6.5-fold in BAEC, with only a minimal increase in most other cell types. Concatamers of region A provide a modest increase in transcription in all cell types except JEG3, where they confer a 5.7-fold increase.

		Relative luciferase activity		
Cells	Construct ...	pGL2 promoter	pGL2 promoter + U	pGL2 promoter + A
BAEC		1.0	6.5 ± 1.4	2.7 ± 0.1
3T3		1.0	1.8 ± 0.2	2.4 ± 0.0
HeLa		1.0	1.4 ± 0.3	2.6 ± 0.2
C6		1.0	1.3 ± 0.2	2.1 ± 0.1
LLC1		1.0	1.4 ± 0.6	2.9 ± 0.7
143B		1.0	2.6 ± 0.4	1.8 ± 0.2
JEG3		1.0	2.4 ± 0.6	5.7 ± 2.2

indicating the presence of one or more elements that are critical for the transcriptional activation and/or cell-restricted expression of *tie2/tek*. Using this 223 bp fragment as a probe in EMSA, a single retarded DNA–protein complex was observed in nuclear extracts from YSC (results not shown). Analysis of the DNA sequence of the *Sac*1–*Sph*1 fragment identified two potential binding sites for regulatory factors. A CCAAT sequence is located at −654 bp, which may represent the binding site for a variety of C/EBP proteins. Another consensus sequence, ATTTGCAT, located at −677 bp represents the binding site for members of the octamer binding family of transcription factors.

Heterologous promoter–reporter gene transfection experiments with regions U and A

To verify further that the sequences identified in functional studies as regions U and A are capable of independently regulating transcription, their activity was evaluated in a luciferase reporter plasmid employing a minimal viral promoter. For these experiments, oligonucleotides encoding the sequence of interest were concatamerized, and resulting fragments containing four copies of the region of interest were subcloned into the reporter plasmid. The activity of the resulting vectors was compared with the parent pGL2 promoter plasmid (Table 3). Region U oligonucleotides increased the rate of transcription approximately 7-fold in BAEC, and this increase was cell-restricted. In most other cell types there was only a minimal 1.5–2-fold increase in transcription, with a moderate increase of 3-fold in 143B osteosarcoma cells and JEG3 choriocarcinoma cells. Region A oligonucleotides had a more universal effect, increasing transcription approximately 2–3-fold in all cell types, except JEG3 where they produced a 6-fold increase. Thus both regions A and U increase transcription of a heterologous promoter, and each region has a unique pattern of transcriptional activity, as exhibited by the relative activity in these different cell types.

DISCUSSION

Identification of multiple regulatory elements in the *tie2/tek* gene

The current study represents a detailed analysis of a promoter region that is capable of directing endothelial cell-specific gene

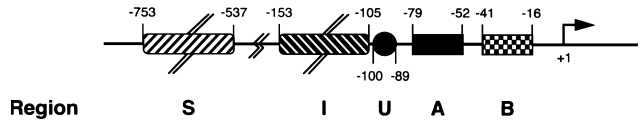


Figure 7 Illustrated summary of the regulatory regions of the Tie2/Tek promoter

Numbers indicate the location of the regulatory regions of *tie2* in relation to the transcription start site which is shown as an arrow. Region S has been previously identified in transgenic experiments. Region I (inhibitory region), regions U and A (positive regulatory regions) and region B have been identified in the current study.

expression *in vivo*. Using the Tie2/Tek promoter, we have identified several novel regulatory elements that are important for its transcriptional activity (summarized in Figure 7). Similar to other genes that are expressed during early development, *tie2/tek* has a TATA-less promoter. Most of the transcriptional activity of this promoter in cultured adult-derived BAEC and endothelial-like embryonic YSC is contained within the proximal promoter and the 5′-untranslated region spanning −153 bp to +324 bp. Reporter transfection experiments have identified an upstream inhibitory region (region I) and two upstream positive regulatory elements (U and A). Protein-binding experiments have allowed the characterization of an additional region, region B, located just upstream of the transcription start site.

The DNA element within region U has been defined at the nucleotide level and consists of a 12 bp motif containing a CATAC repeat (GCATACCATACA). While this element binds nuclear factor(s) derived from both endothelial and HeLa cells, it is capable of activating the transcription of a heterologous promoter, preferentially in endothelial cells rather than in non-endothelial cell types. The consensus DNA-binding sequence of this element is novel, which suggests that the cognate DNA-binding proteins are likely to represent a new family of such factors.

Another binding site, region A, has been identified within the proximal promoter and consists of a 28 bp DNA sequence. The binding element contained in this region has not been completely defined; however, a CAAT sequence contributes to the recognition site. Despite similarities in the binding motif, region A does not represent a consensus CAAT box and does not bind C/EBP factors, as demonstrated by EMSA. This suggests that the factor(s) that bind this element may also represent a new family of DNA-binding proteins. Data from transfection and protein-binding experiments suggest that this element interacts with ubiquitously expressed factor(s).

The most proximal element, element B, consists of a 26 bp DNA sequence identified by footprinting analysis. The precise binding motif within this region has not been fully defined. Of interest is the presence of a TGAAAAT sequence within this element located 30 bp upstream of the transcription start site. Hariharan and Perry [26] have previously identified a homologous sequence, TGAAAAAT, located 30 bp upstream of the transcription start site of the TATA-less rpS16 gene [26]. This element was shown to bind a yet unidentified nuclear factor distinct from conventional TATA-binding proteins and is likely to function as a TATA box counterpart. While mutation of this element in the rpS16 gene abolished binding by EMSA, identical mutations within region B of the Tie2/Tek promoter (TGAAAATG → TGCTCCTC) caused a significant decrease in, but not a total loss of, binding (results not shown). These data suggest that this 8 bp sequence is likely to be a part of a larger protein-binding domain. Because of the location of the binding site and its homology to the rpS16 element, it would seem likely that element B functions as a TATA box counterpart for the TATA-less *tie2/tek* gene. This also suggests that this region may interact with the basal transcription apparatus either directly or via a 'tethering' factor [26]. The fact that two different proteins bind this element in YSC and BAEC suggests developmental differences in the expression of *tie2/tek*. This is likely to result in a switch of the rate of transcription at different stages of endothelial cell development and may explain the differences in the transcriptional activity of the *tie2/tek* construct in BAEC compared with YSC.

Correlation of findings *in vitro* with transgenic experiments *in vivo*

Some of the *in vitro* findings demonstrated in the current study correlate well with previous data in transgenic mice, while in some cases there are differences. Most significantly, the 2000 bp region upstream of the transcription start site of *tie2/tek*, shown to be cell-specific in transgenic mice, is also cell-specific in transfection experiments. However, differing from the transgenic data, experiments *in vitro* indicate that even the minimal promoter contains elements that can specify endothelial cell expression. Another finding in common is the presence of important positive regulatory elements within the first untranslated exon. In transgenic animals, constructs lacking +67 bp (*Sty*1 site) to +323 bp (*Hin*dIII site) were noted to have diminished reporter gene expression [13]. Using reporter transfection experiments, we have identified positive regulatory elements between +107 bp and +233 bp and between +233 bp and +255 bp (results not shown). Further delineation of the DNA sequences responsible for this activity and correlation with protein binding experiments are currently underway.

Experiments in transgenic animals have identified an upstream enhancer located between −753 bp (*Sac*1 site) and −537 bp (*Sph*1 site) (region S), which proved to be critical for *in vivo* transcription of the LacZ reporter transgene in embryonic endothelial cells (Figure 7) [13]. Using protein-binding experiments, we have identified protein factor(s) that specifically interact with this region (results not shown). However, experiments *in vitro* reported here showed that the deletion of this region had no effect on the transcriptional activity of the Tie2/Tek promoter in either BAEC or YSC. These findings may suggest that the proteins binding this region are unable to function outside of the chromosome, or that these regulatory factors need to cooperate with additional factors to modulate endothelial cell-specific expression. There are examples of DNA elements that have been characterized in association with other genes that are capable of regulating transcription *in vivo*, but have little effect on transcription of transfected reporter constructs [27]. In some cases, these elements appear to function simply to open the chromatin, and in other cases they have enhancer-like activity which requires chromosomal integration for productive interactions with a promoter region [27,28].

Possible mechanisms of endothelial cell-specific gene expression

Despite the significant new information provided by this study, the mechanisms underlying cell-specific expression of *tie2/tek* are not fully defined. Possible mechanisms for determining cell-specific expression involve binding of a cell-restricted transcriptional regulator to its target sequence or the interaction of a ubiquitous DNA-binding protein with a cell-specific transcriptional coactivator. The tissue-restricted expression of *tie2* may be mediated in part through the GCATACCATACA

element characterized within region U. This element is capable of significantly activating transcription of a heterologous promoter in endothelial cells and to a much lesser degree in several non-endothelial cell types. However, the protein binding this element is not cell-restricted, as is evident in a similar pattern of binding in endothelial and HeLa cells using EMSA. These findings suggest that if element U contributes to cell-specific expression of *tie2/tek*, this has to reside in the interaction of promoters and enhancers with histones, and the potential for this interaction to determine cell-restricted transcription. It is well established that removal of histone-controlled repression is a necessary step for the activation of many silent eukaryotic genes [29]. Thus, despite the presence of the same transcription factor in a variety of cell types, activation of transcription may take place in some but not in others, depending on the state of the chromatin and whether it can be 'opened' to uncover essential regions of DNA for interaction with these factors [29]. Analysis of DNA–protein interactions using footprinting *in vivo* in endothelial and non-endothelial cells may help determine the sites of endothelial cell-specific interaction.

Another possible mechanism for determining tissue-specific expression resides in region S, which is known to be essential for the endothelial expression of *tie2/tek* reporter transgenes [13]. We have shown that this region is capable of binding factors present in embryonic endothelial-like YSC (results not shown). Analysis of DNA sequences within this region have identified two potential sites for interaction with regulatory factors. A CCAAT sequence, located at −654 bp, may serve as the binding site for a variety of C/EBP factors. A member of this family, C/EBPα, is known to direct terminal differentiation and cell-specific transcription in both adipocytes and liver cells [30,31]. Region S also contains a consensus octamer site, ATTTGCAT, known to bind members of the POU domain family of transcription factors [32]. One of these factors, Oct-1, has a ubiquitous pattern of expression and regulates the transcriptional activity of many genes, including the immunoglobulin genes [32]. B-Cell-restricted expression of the immunoglobulin genes is determined by Bob-1 (also termed OCA-B and OBF-1), a specific coactivator of Oct-1 that is expressed exclusively in B-cells. A second mechanism of cell-specific expression mediated by the octamer motif may involve its interaction with cell-restricted POU-domain transcription factors that are distinct from Oct-1. Several of these transcriptional regulators such as Pit-1, Sprm-1, Skn-1 and N-Oct factors have been well characterized and are expressed in a cell-specific manner. These factors are known to regulate the cell-restricted expression of genes in the pituitary gland, male germ cells, epidermis/hair follicles and neurons respectively [32–35]. Since region S is required for optimal expression of *tie2/tek in vivo*, it remains possible that a cell-restricted coactivator of Oct-1 or a cell-restricted POU-domain transcription factor is responsible for mediating endothelial cell-specific expression.

This work was supported by an RAGS grant from the Veterans Administration (B.M.F.), grant HL52168 from the NHLBI (T.Q.) and an Established Investigator Award from the American Heart Association (T.Q.). The authors extend their appreciation to Dr. Yie-Teh Yu for her thoughtful comments and insights.

Received 29 July 1997/23 September 1997; accepted 8 October 1997

REFERENCES

1 Coffin, J. D. and Poole, T. J. (1988) Development **102**, 735–748
2 Noden, D. M. (1989) Am. Rev. Respir. Dis. **140**, 1097–1103
3 Coffin, J. D. and Poole, T. J. (1991) Anat. Rec. **231**, 383–395
4 Pardanaud, L., Yassine, F. and Dieterlen Lievre, F. (1989) Development **105**, 473–485
5 Pardanaud, L., Altmann, C., Kitos, P., Dieterlen Lievre, F. and Buck, C. A. (1987) Development **100**, 339–349
6 Noden, D. M. (1990) Ann. NY Acad. Sci. **588**, 236–249
7 Shalaby, F., Rossant, J., Yamaguchi, T. P., Gertsenstein, M., Wu, X. F., Breitman, M. L. and Schuh, A. C. (1995) Nature (London) **376**, 62–66
8 Fong, G. H., Rossant, J., Gertsenstein, M. and Breitman, M. L. (1995) Nature (London) **376**, 66–70
9 Dumont, D. J., Gradwohl, G., Fong, G. H., Puri, M. C., Gertsenstein, M., Auerbach, A. and Breitman, M. L. (1994) Genes Dev. **8**, 1897–1909
10 Sato, T. N., Tozawa, Y., Deutsch, U., Wolburg-Buchholz, K., Fujiwara, Y., Gendron-Maguire, M., Gridley, T., Wolburg, H., Risau, W. and Qin, Y. (1995) Nature (London) **376**, 70–74
11 Lee, M. E., Block, K. D., Clifford, J. A. and Quertermous, T. (1990) J. Biol. Chem. **265**, 10446–10450
12 Aird, W. C., Jahroudi, N., Weiler Guettler, H., Rayburn, H. B. and Rosenberg, R. D. (1995) Proc. Natl. Acad. Sci. U.S.A. **92**, 4567–4571
13 Schlaeger, T. M., Qin, Y., Fujiwara, Y., Magram, J. and Sato, T. N. (1995) Development **121**, 1089–1098
14 Harats, D., Kurihara, H., Belloni, P., Oakley, H., Ziober, A., Ackley, D., Cain, G., Kurihara, Y., Lawn, R. and Sigal, E. (1995) J. Clin. Invest. **95**, 1335–1344
15 Korhonen, J., Lahtinen, I., Halmekyt, M., Alhonen, L., Jänne, J., Dumont, D. and Alitalo, K. (1995) Blood **86**, 1828–1835
16 Kadonaga, J. T. and Tijan, R. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 5889–5893
17 Corn, B. J., Reed, M. A., Dishong, S. L., Li, Y. and Wagner, T. E. (1991) Clin. Biotechnol. **3**, 15–19
18 Quertermous, E. E., Hidai, H., Blanar, M. A. and Quertermous, T. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 7066–7070
19 Lee, M. E., Temizer, D. H., Clifford, J. A. and Quertermous, T. (1991) J. Biol. Chem. **266**, 16188–16192
20 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (eds.), (1987) John Wiley and Sons, New York
21 Dignam, J. D., Levovitz, R. M. and Roeder, R. G. (1983) Nucleic Acids Res **11**, 1475–1489
22 Nothias, J. Y., Majumber, S., Kaneko, K. J. and DePampilis, M. L. (1995) J. Biol. Chem. **270**, 22077–22080
23 Osada, S., Yamamoto, H., Nishihara, T. and Imagawa, M. (1996) J. Biol. Chem. **271**, 3891–3896
24 Patterson, C., Parrella, M. A., Hsieh, C. M., Yoshizumi, M., Lee, M. E. and Haber, E. (1995) J. Biol. Chem. **270**, 23111–23118
25 Morishita, K., Johnson, D. E. and Williams, L. T. (1995) J. Biol. Chem. **270**, 27948–27953
26 Hariharan, M. and Perry, R. P. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 1526–1530
27 Tapscott, S. J., Lassar, A. B. and Weintraub, H. (1992) Mol. Cell Biol. **12**, 4994–5003
28 Ellis, J., Tan-Un, K. C., Harper, A., Michalovich, D., Yannoutsos, N., Philipsen, S. and Grosveld, F. (1996) EMBO J. **15**, 562–568
29 Kingston, R. E., Bunker, C. A. and Imbalzano, A. N. (1996) Genes Dev. **10**, 905–920
30 Zhaodan, C., Umek, R. M. and McKnight, S. L. (1991) Genes Dev. **5**, 1538–1552
31 Freytag, S. O., Paielli, D. L. and Glibert, J. D. (1994) Genes Dev. **8**, 1654–1663
32 Herr, W. and Cleary, M. (1995) Genes Dev. **9**, 1679–1693
33 Anderson, B., Pearse, R. V., Schlegel, P. N., Cichon, Z., Schonemann, M. D., Bardin, C. W. and Rosenfeld, M. G. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 11084–11088
34 Anderson, B., Schonemann, M. D., Flynn, S. D., Pearse, R. V., Singh, H. and Rosenfeld, M. G. (1993) Science **260**, 78–82
35 Schreiber, E., Harshman, K., Kemler, I., Malipiero, U., Schaffner, W. and Fontana, A. (1990) Nucleic Acids Res. **18**, 5495–5503