Isolation of a Regulatory Region of Activin Receptor-Like Kinase 1 Gene Sufficient for Arterial Endothelium-Specific Expression

Tsugio Seki, Kwon-Ho Hong, Jihye Yun, Seong-Jin Kim, S. Paul Oh

Abstract—Activin receptor-like kinase 1 (Acvrl1; Alk1) is a type I receptor for transforming growth factor-β (TGF-β). ALK1 plays a pivotal role in vascular development and is involved in the development of hereditary hemorrhagic telangiectasia 2 (HHT2), a dominantly inherited vascular disorder, and pulmonary hypertension. We have previously shown that Alk1 is expressed predominantly in arterial endothelial cells (ECs). Despite recent discoveries of a number of artery-specific genes, the regulatory elements of these genes have not been characterized. To investigate the cis-acting elements essential for the artery-specific Alk1 expression, we have generated a series of transgenic constructs with various lengths and regions of Alk1 genomic fragments connected to a LacZ reporter gene, and analyzed the reporter gene expression in transgenic mice. We found that a 9.2-kb genomic fragment, which includes 2.7-kb promoter region and the entire intron 2, is sufficient to drive arterial endothelium-specific expression. The defined regulatory region, as well as the transgenic mouse lines, would be invaluable resources in studying the mechanisms underlying angiogenesis, arteriogenesis, and vascular disorders, such as HHT and pulmonary hypertension. The full text of this article is available online at http://circres.ahajournals.org. (Circ Res. 2004;94:e72-e77.)

Key Words: activin receptor-like kinase 1 ▪ artery-specific gene expression ▪ angiogenesis ▪ hereditary hemorrhagic telangiectasia ▪ transforming growth factor-β

Blood vessels consist of a network of arteries, capillaries, and veins. Recent discoveries of artery- or vein-specific genes, such as ephrinB2 and its receptor EphB4, and some molecules involved in the Notch-delta signaling pathway have contributed to a significant advancement in understanding the mechanisms by which arteries and veins gain their distinct identities during vascular development. In contrast to the long-standing belief that the acquisition of arterial and venous identities is largely determined by different physiological parameters (eg, blood flow, blood pressure, and shear stress), genetic studies in mice and fish suggest that arterial and venous ECs acquire their distinct molecular identities before the establishment of blood flow. Studies have shown that the aforementioned artery- or vein-specific genes play crucial roles in angiogenesis and segregation of the two blood vessel types. It remains unclear, however, whether those genes are directly involved in the morphogenesis of vessel type-specific architecture during development.

ALK1 is a transforming growth factor-β (TGF-β) type I receptor in vascular ECs. Haploinsufficiency of ALK1 in humans causes HHT2, which is characterized by recurrent epistaxis, localized mucocutaneous telangiectases, and arteriovenous malformations (AVM) in the lungs, liver, and brain. Similarly, heterozygous Alk1 mutant mice exhibit HHT-like vascular lesions, whereas Alk1-null homozygotes die at midgestation with hyperdilation of vessels, AVMs, and defects in the differentiation/recruitment of vascular smooth muscle cells (VSMCs). Using Alk1-LacZ “knock-in” mice (Alk1<sup>+/ Δαα<sup>2</sup>), we have previously observed a dynamic spatiotemporal expression pattern of Alk1: Alk1 is expressed predominantly in arterial ECs throughout the developmental and postnatal growth stages, and its expression is diminished in the adult stage; Alk1 expression is induced in nascent and remodeling arteries during angiogenesis prompted by either wound healing or tumorigenesis. These expression data together with the VSMC defect in Alk1-deficient embryos suggest that ALK1 may be directly involved in the morphogenesis of arterial development and remodeling.

To investigate the cis-acting elements required for the spatiotemporal expression of Alk1 gene (Figure 1a), we have generated three transgenic constructs containing various regions/lengths of putative regulatory sequences connected to the LacZ gene (Figure 1b). The first construct, pXh4.5-SIBN, contained a 2.7-kb promoter, exon 1, and a part of 5′ region of intron 2. We established five independent founder lines...
from this construct, and the transmission of the transgene to 
F1 offspring was confirmed by genomic Southern blot anal-
yses. The embryos from none of these lines, however, 
displayed the vascular-specific expression of the 
\( \text{LacZ} \) gene.

To investigate whether the essential regulatory elements 
resided in upstream sequences, the 2.7-kb promoter region 
was extended to a total length of 8 kb in the second construct, 
pBam9-SIB. Nevertheless, a mouse line from this construct 
did not exhibit vascular-specific expression pattern of the 
transgene.

These results prompted us to search for the potential 
regulatory elements by comparing DNA sequences between 
the human and mouse ALK1 gene loci. Intriguingly, dot plot 
and blast analyses revealed the presence of highly conserved 
sequences at the 3′/H11032 
region of intron 2 (Figure 1c). The span of 
1.8 kb conserved region in the 3′/H11032 
region of intron 2 consisted 
of at least five homologous regions ranging from 37 to 145 bp 
in length, and each sequence cluster showed more than 80% 
identity (data not shown). Based on this result, we generated 
the third construct, pXh4.5-in2-SIB, which is similar to the 
first construct, pXh4.5-in2-SIB, but with the extension of 3′ 
region to include the remaining intron-2, exon 3, and approx-
imately 300 bp of 5′ region of intron 3, for a total length of 
9.2 kb (Figure 1b). Two founder lines from this construct 
showed artery-specific expression of the \( \text{LacZ} \) gene (Figure 
2a), which recapitulated the pattern seen in \( \text{Alk1}^{-/\text{Alk1}} \) embryos (Figure 2b).

To examine detailed regulatory activity of the 9.2-kb \( \text{Alk1} \) sequences, we analyzed the \( \text{LacZ} \) expression in various stages of embryos, adult organs, as well as in wounded skins, and compared the results to those of \( \text{Alk1}^{-/\text{Alk1}} \) mice. Two trans-
genic mouse lines from pXh4.5-in2-SIB construct, \( \text{Tg(Alk1-LacZ)C2} \) and \( K1 \), showed identical expression patterns with 
slightly different intensity (thus, will be collectively referred 
as \( \text{Tg(Alk1-LacZ)} \) hereafter). In embryos, the transgene 
expression pattern of \( \text{Tg(Alk1-LacZ)} \) was virtually indistin-
guishable from \( \text{Alk1}^{-/\text{Alk1}} \) except for the intensity levels 
(Figure 2). Strong expressions in the capillaries of perineural 
tissues in embryonic days 10.5 to 13.5 (E10.5 to E13.5) were 
consistently observed (Figures 2c and 2d). The transgene 
expression in \( \text{Tg(Alk1-LacZ)} \) was detected in developing 
arterial ECs, but neither in VSMCs nor in venous ECs 
(Figures 2e and 2f). The artery-specific expression was also 
clearly shown in internal thoracic arteries (Figures 2g and 
2h), mesenteric arteries (Figure 2i), and descending aorta and 
tercostal arteries (Figures 2j and 2k). Strong expression in 
the pulmonary vessels (Figures 2l and 2m), and the lack of 
staining in the liver (data not shown) were also consistent 
with those in \( \text{Alk1}^{-/\text{Alk1}} \).

The X-gal staining patterns of \( \text{Tg(Alk1-LacZ)} \) mice during 
postnatal life were also identical in a large extent to the ones 
of \( \text{Alk1}^{-/\text{Alk1}} \) mice, in which the intense \( \text{Alk1} \) expression in 
arterial vessels during the postnatal growth phase is greatly 
diminished in most of the tissues except for the lungs at the 
adult stage. Consistent with these observations, \( \text{Tg(Alk1-}
\text{LacZ)} \) mice showed intense vascular staining, such as in 
the brain and the iris (Figures 3a and 3b), during the early 
postnatal growth phase. In the adult stage, however, positive

Figure 1. Transgenic constructs and the 
dot plot analysis between human and 
mouse ALK1 genomic sequences. a, Full 
mouse \( \text{Alk1} \) gene structure. Restriction 
enzyme sites used for generating the 
transgenic constructs are indicated by X 
(\( \text{XhoI} \)) and B (\( \text{BamHI} \)). Boxes represent 
exons, and exon numbers are indicated 
above. b, Schematic representations of 3 
transgenic constructs. Detailed construct 
information is available in the Materials 
and Methods. c, Dot plot analysis 
between partial genomic sequences of 
human (y-axis) and mouse (x-axis) ALK1 
gene. Exon 1 of the human ALK1 gene is 
split in 2 exons (1 and 2) in the mouse 
\( \text{Alk1} \) gene. Continuous sequence homol-
}
staining was barely detectable in capillary-like small vessels throughout the body except for the moderate expression in the lungs (Figures 3c through 3e). In the skin, the 
\text{Tg(Alk1-LacZ)\text{\textregistered}} mice showed intense staining of arteries at newborn stage (Figure 3f) as observed in \text{Alk1/H11001/LacZ\textregistered}. Unlike \text{Alk1/H11001/LacZ\textregistered} mice, however, the expression was faded away in the 2-week-old mice.

We also examined the transgene expression during induced angiogenesis in adult mice using skin wound healing model. Although overall X-gal staining intensity in the skin was lower in \text{Tg(Alk1-LacZ)} mice when compared with the \text{Alk1/H11001/LacZ\textregistered} mice, the expression pattern was comparable. During wound healing in the adult stage, the transgene expression was induced in arterial vessels in a similar fashion as the \text{Alk1/H11001/LacZ\textregistered} mice, although the staining was significantly lighter and limited in both duration and area (Figures 3g through 3j). Although the vascular staining in the wounded area of the \text{Alk1/H11001/LacZ\textregistered} mice was present from days 3 to 12 after wounding, staining of the transgenic mice was observed only from days 3 to 8. Also, the staining was restricted to the vessels adjacent to the wounds and did not extend to distant arteries that fed the wound lesion in the transgenic mice.

Our transgenic studies indicate that the conserved intronic regions may contain the enhancer element(s) for the artery-specific \text{Alk1} expression. To investigate whether the 9.2-kb regulatory fragment of \text{Alk1} gene contains common regulatory elements, we have performed DNA sequence homology alignment between the 9.2-kb regulatory sequence and regulatory regions of other known endothelium- or artery-specific genes (for details, see the Materials and Methods section). Extensive homology searches, however, did not show any significant continuous homology region.

To identify potential transcriptional factor binding sites (TFBSs) in the 9.2-kb regulatory region of \text{Alk1} gene contains common regulatory elements, we have performed DNA sequence homology alignment between the 9.2-kb regulatory sequence and regulatory regions of other known endothelium- or artery-specific genes (for details, see the Materials and Methods section). Extensive homology searches, however, did not show any significant continuous homology region.

To identify potential transcriptional factor binding sites (TFBSs) in the 9.2-kb regulatory fragment, comprehensive in silico analysis with the rVISTA 2.0 program was performed. We found 16 conserved TFBSs for 12 transcription factors known to regulate EC-specific genes within the 9.2-kb regulatory fragment (Table; Figure 4). The in vivo relevance of the conserved intronic regions and TFBSs remains to be investigated.

We have demonstrated using transgenic mice that the 9.2-kb fragment of \text{Alk1} gene contains essential regulatory elements for its spatiotemporal expression. We have also presented an intriguing possibility that the conserved region

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Artery-specific LacZ expression in the blood vessels of \text{Tg(Alk1-LacZ)} embryos. a and b, Whole-mount X-gal–stained E11.5 \text{Tg(Alk1-LacZ)} (a) and \text{Alk1/H11001/LacZ\textregistered} (b) embryos, showing the identical staining pattern but with much higher intensity in \text{Tg(Alk1-LacZ)} c and d. Strong LacZ-positive capillary-like vessels in the developing neural tube; the section of E11.5 embryo (c) and a dorsal view of E12.5 neural tube (d). e and f, Sections of X-gal–stained E11.5 \text{Tg(Alk1-LacZ)} embryos counterstained with nuclear fast red (NFR) (e) or immunostained with anti-smooth muscle \text{α}-actin antibodies (f). Note that the X-gal–positive cells were found in a single layer of endothelium in the artery (a, arrowheads), but not in the vein (v), and were not concomitant with VSMCs (f). g and h, Blood vessels on the thoracic wall of newborn pups before (g) and after (h) X-gal staining, showing that only arteries and connecting small arteries were X-gal positive (h). i, Whole-mount X-gal–stained mesenteric vessels of E18.5 fetuses. Only mesenteric artery (a) but not vein (v) showed the transgene expression. j and k, X-gal–stained aorta and intercostal blood vessels before (j) and after (k) clearing with the organic solvents. Similar sizes of intercostal arteries (a) and veins (v) were running side-by-side (j), whereas only arteries and dorsal aorta (ao) were X-gal positive (j and k). X-gal–positive staining in the rib skeleton (‘) is from the endogenous \text{β}-gal activity. l, Intense X-gal staining in the lungs and aorta (ao) of E15.5 \text{Tg(Alk1-LacZ)} embryos. m, NFR counterstained section of E15.5 lungs showed X-gal–positive arteries (a) and capillaries (arrowheads). Note that bronchial epithelial cells were X-gal negative. Scale bars in c, e, f, and m=50 \mu m.}
\end{figure}
in intron-2 may contain the enhancer elements for the artery-specific gene expression. This study also suggests that the expression pattern observed in the \(Alk1^{+/+}\) mice was not a consequence of \(Alk1\) haplinsufficiency. The variance of expression patterns between \(Tg(Alk1-\text{LacZ})\) and \(Alk1^{+/+}\) mice in the skin wounds may be due to either the strain differences, or a lack of additional \(cis\)-acting elements involved in the \(Alk1\) expression during wound-induced angiogenesis.

The presented 9.2 kb regulatory fragment, as the first regulatory elements specific for arterial endothelial cells, would provide invaluable information regarding regulatory mechanisms of \(Alk1\), as well as other artery-specific genes.

**Figure 3.** Postnatal transgene expressions in \(Tg(Alk1-\text{LacZ})\) mice. a through d, X-gal–stained blood vessels in the brain (a and c) and the iris (b and d). Intense transgene expression in newborn brain (a) and 2-week-old iris (b) were greatly diminished in the adult stage (c and d). Only speckled staining in small branches (arrowheads) was observed (c and d). e, Adult lung section counterstained with NFR. Alveolar capillaries expressed the transgene in the adult stage. f, Intense staining in the arteries of newborn skin. g through j, X-gal staining in the skin during days 3 (g), 5 (h), 8 (i), and 10 (j) after wounding. LacZ expression was induced during wound healing (arrowheads), but in a limited extent. Number and intensity of X-gal–positive vessels peaked at day 5 after wounding (h) and diminished completely by day 10 (j). Asterisks (*) in g through j indicate the wounds.

Furthermore, this regulatory fragment can be used to activate or silence a gene in developing arteries to study its function in arteriogenesis and remodeling of arterial vessels.

**Materials and Methods**

Transgenic strains were generated at the University of Florida or the National Cancer Institute. All animal procedures performed were reviewed and approved by the University of Florida Institutional Animal Care and Use Committee.

**Generation and Characterization of Transgenic Mice**

The first two constructs, pXh4.5-SIBN and pBam9-SIB, were generated from Xhol 4.5 kb and BamH1 9.2 kb fragments of \(Alk1\), which contain 2.7 kb and 8 kb of the promoter regions, respectively, as well as exon 1 through a 5’ region of intron 2 (Figure 1b). For the third construct, pXh4.5-in2-SIB, the 3’ end of the Xhol 4.5 kb fragment was extended to 3’ region to include the rest of intron 2, exon 3, and a 5’ region of intron 3, for a total length of 9.2 kb. Each \(Alk1\) genomic fragment was connected to either SIBN or SIB cassette, which contained SV40 splicing donor/acceptor signals, internal ribosomal entry sequence, \(\beta\)-galactosidase cDNA, poly A signal, and with/without a neomycin resistant gene cassette driven by the PGK promoter. The microinjections of the DNA constructs into male pronuclei of fertilized eggs from the FVB strain were performed using established procedures. The founder lines were screened by genomic Southern blot using the \(Alk1\) genomic fragment as probe. Several founder lines per construct were further examined for their transgene expression by whole-mount X-gal staining of F1, F2, or F3 embryos at E9.5 or E10.5. Consequently, four pXh4.5-SIBN, one pBam9-SIB, and two pXh4.5-in2-SIB mouse lines were established.

**Immunohistochemistry and X-Gal Staining**

Mouse embryos and 3-month-old adult tissues were stained with X-gal and photographed before and after clearing as described. Immunohistochemistry was performed in paraffin-embedded whole-mount X-gal–stained samples were stained with anti-smooth muscle \(\alpha\)-actin (clone: 1A4;
To identify potential transcriptional factor binding sites (TFBSs) in binding sites with rVISTA 2.0, analysis of potential transcriptional factor

Figure 4. Potential transcriptional factor binding sites identified in the mouse Alk1 pXh4.5-in2 fragment by the rVISTA analysis. Top, 16 TFBS matrices, which represent TFBSs for 12 different transcription factors. Some transcription factors have multiple TFBS matrices. For example, SP1 has 4 different TFBS matrices (SP1, SP1_Q2, SP1_Q4, and SP1_Q6). Positions of each conserved TFBS are indicated as a bar within the matrix. Bottom, Mouse Alk1 gene structure (the top line), and the sequence homology between mAlk1 and hALK1 sequences. Color legends in the boxed area indicate the nature of the highly homologous regions. Repetitive sequences are indicated by green boxes on the bottom line. Only the transcriptional factors known to have active roles in regulating EC-specific genes are shown.

Sigma) antibodies using the Vectastain staining kit (Vector Laboratory).

Wound Healing Study

Three full-thickness excisional wounds, 4 mm in diameter each, were created on the back skin of 3-month-old mice and analyzed as previously described.12

Sequence Comparison Analyses

Mouse and human ALK1 genomic DNA sequences were obtained from UCSC Genome browser (http://genome.ucsc.edu). To visualize human and mouse ALK1 genomic sequence homology, dot-plot sequence comparison was performed by using the Dot Matcher program available at European Molecular Biology Open Software Suite (http://bioinfo.pbi.nrc.ca:8090/EMBOSS/; Figure 1c), and DNA sequence homology alignment was conducted using Blast 2 Sequences program (http://www.ncbi.nlm.nih.gov/gorf/bl2.html). The Blast 2 Sequences were also used to search for common regulatory elements between Alk1 and endothelium- or artery-specific genes. The following genes were included: Tek (Tie-2),27 Nos3 (eNOS),36 and Cdhl5 (VE-cadherin),32 for endothelium-specific genes; and Ephb2 (Ephrin-B2),33 Bmx,34 Nrp (Neuropilin-1),35 Dll4,36 Notch1,37 Notch3,37 and Notch4,37 for artery-specific genes. Full-length gene sequences as well as 10-kb 5′ regulatory regions from each gene were used as queries, and comparisons against 9.2-kb pXh4.5-in2 sequence were performed with default settings.

Analysis of Potential Transcriptional Factor Binding Sites With rVISTA 2.0

To identify potential transcriptional factor binding sites (TFBSs) in the pXh4.5-in2 sequence, the rVISTA 2.0 program provided by Lawrence Livermore National Laboratory was utilized (http://rvista.dcode.org/). The program matches TFBS consensus matrix from TRANSFAC 7.3 database (http://www.biobase.de) to query sequences, and test whether matched consensus matrices are conserved between two query sequences, based on the specific criteria (for details, see Loots et al).51 Some transcriptional factors (TFs) have multiple consensus matrices with different names in the database. Therefore the number of potential TFBS consensus matrices is usually greater than the number of potential TFs. In addition, the number of TFBSs is also usually greater than the number of TFBS consensus matrices because each TFBS consensus matrix can be matched to multiple locations in the sequence queries. For the Alk1 regulatory element analysis, the same set of sequences used for the dot-plot analysis was inputted to the rVISTA 2.0 program, and potential TFBSs that were conserved between mouse and human sequences were obtained. The program located 471 conserved potential TFBSs in the 9.2-kb mouse sequence, where each TFBS was matched to one of the 112 TFBS consensus matrices. Among these 112 matrices, 16 TFBS consensus matrices for the 12 TFs were known to regulate EC-specific genes (Figure 4; Table).

Acknowledgments

This work was supported by the NIH (HL64024) and the American Heart Association (SDG 9930031N) to S.P.O., as well as by an American Heart Association postdoctoral fellowship to T.S.

References


Isolation of a Regulatory Region of Activin Receptor-Like Kinase 1 Gene Sufficient for Arterial Endothelium-Specific Expression
Tsugio Seki, Kwon-Ho Hong, Jihye Yun, Seong-Jin Kim and S. Paul Oh

_Circ Res._ 2004;94:e72-e77; originally published online April 1, 2004;
doi: 10.1161/01.RES.0000127048.81744.31

_Circulation Research_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2004 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/94/8/e72