Role of RNA splicing in mediating lineage-specific expression of the von Willebrand factor gene in the endothelium

Lei Yuan¹,², Lauren Janes¹,², David Beeler¹,², Katherine C. Spokes¹,², Joshua Smith¹,², Dan Li¹,³, Shou-Ching Jaminet¹,³, Peter Oettgen¹,², William C. Aird¹,²

From ¹The Center for Vascular Biology Research and Division of Molecular and Vascular Medicine, Beth Israel Deaconess Medical Center, Boston, MA; ²Division of Molecular and Vascular Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA; ³Department of Pathology, Beth Israel Deaconess Medical Center, Boston, MA

Address all Correspondence to: William C. Aird, M.D., Beth Israel Deaconess Medical Center, Molecular and Vascular Medicine; RN-236, 99 Brookline Ave., Boston MA

Short Title (Running Head): RNA splicing and vWF expression
KEY POINTS

- RNA splicing of the first intron of the von Willebrand factor gene is essential for expression in the endothelium
- RNA splicing may play a role in mediating endothelial cell heterogeneity
ABSTRACT

We previously demonstrated that the first intron of the human von Willebrand factor (vWF) is required for gene expression in the endothelium of transgenic mice. Based on this finding, we hypothesized that RNA splicing plays a role in mediating vWF expression in the vasculature. To address this question, we employed transient transfection assays in human endothelial cells and megakaryocytes with intron-containing and intronless human vWF promoter-luciferase constructs. Next, we generated knockin mice in which LacZ was targeted to the endogenous mouse vWF locus in the absence or presence of the native first intron or heterologous introns from the human beta-globin, mouse DSCR-1 or hagfish coagulation factor X genes. In both the in vitro assays and the knockin mice, the loss of the first intron of vWF resulted in a significant reduction of reporter gene expression in endothelial cells, but not megakaryocytes. This effect was rescued to varying degrees by the introduction of a heterologous intron. Intron-mediated enhancement of expression was mediated at a post-transcriptional level. Together, these findings implicate a role for intronic splicing in mediating lineage-specific expression of vWF in the endothelium.
INTRODUCTION

Von Willebrand factor (vWF) is a plasma protein that mediates platelet hemostatic function and stabilizes coagulation factor VIII. vWF plays a particularly important role in platelet adhesion and aggregation at sites of high shear rates (reviewed in 1). Expression of vWF is restricted to endothelial cells and platelets. vWF is differentially expressed in the vasculature. For example, vWF protein and mRNA levels are higher in veins, compared with arteries 2, 3, and in venules compared with arterioles 4. Within the microvasculature, vWF is expressed at particularly low levels in the liver. Expression of vWF also varies between neighboring endothelial cells 5. The elucidation of the mechanisms underlying vWF expression may provide important insights into the molecular basis of vascular diversity.

The structural organization of the mouse, human, and bovine vWF promoter-proximal region is closely related 6-8. The first exon (+1 to +246 in human vWF) encodes 5′ untranslated sequence. The second exon contains the ATG translational start site. Exons 1 and 2 are separated by an intron (+247 to +1475 in human vWF). To study mechanisms of vWF gene regulation, we have previously employed a plug-in-socket approach in which a single copy of a vWF promoter-LacZ cassette is targeted to the Hprt locus of mice. Using this strategy, we showed that a region of the human vWF gene (termed vWF2, between -2182 and the end of the first intron) directed expression in the vasculature of the brain, heart and skeletal muscle, but not in other organs such as aorta, lung, liver, spleen and kidney 9. Nor was expression observed in megakaryocytes. In
contrast, an intronless version of vWF2 (between -2182 and +246) coupled to LacZ failed to express in the endothelium of heart and skeletal muscle. Expression was rescued by replacement of the first intron of vWF with the second intron of the human β-globin gene. These findings suggested one of two possibilities. First, the β-globin intron contains cis-regulatory elements sufficient for rescuing expression. Alternatively, splicing of the first intron of vWF is necessary for expression of LacZ when coupled to the vWF promoter in the Hprt locus.

The goal of this study was to determine the role of RNA splicing in mediating expression of vWF. Using transient transfections, we show that splicing is required for vWF promoter activity in endothelial cells, but not megakaryocytes. Similarly, in mice in which LacZ has been knocked into the endogenous vWF locus, the loss of the first intron abrogated reporter gene expression specifically in endothelial cells, while the introduction of heterologous introns reversed this effect. Together, these findings implicate a role for intronic splicing in mediating lineage-specific expression of vWF in the endothelium.
METHODS

**Plasmid constructions.** Generation of vWF and Tie2 promoter-luciferase constructs used in transient transfection assays and of the gene targeting constructs is detailed in supplemental Methods.

**Generation and analysis of *Hprt* and *vWF*-locus-targeted mice.** The generation of *Hprt*-vWF2-βGIVS2-lacZ mice was previously described 10. The generation of additional mice is detailed in supplemental Methods (please see Supplemental Fig. 1). Whole mounts and/or tissue sections from targeted mice were assayed by LacZ staining or vWF immunostaining, as detailed in supplemental Methods. All animal studies were approved by the Animal Care and Use Committee at the Beth Israel Deaconess Medical Center.

**Cell culture and transfections.** Human umbilical vein endothelial cells (HUVEC) were purchased from Lonza and cultured in endothelial cell medium supplemented with EGM-2-MV bullet kit (Lonza). Mouse endothelial cells were harvested from the hearts of mice as detailed in supplemental Methods.

**Transient transfection assays.** Transient transfections were carried out as described in supplemental Methods.

**ChIP Assay.** Chromatin immunoprecipitation (ChIP) assays were preformed using a chromatin immunoprecipitation assay kit (Upstate Biotechnology) according to the manufacturer’s instructions and as detailed in supplemental Methods.
Quantitative reverse transcriptase polymerizes chain reaction. Quantitative real-time transcriptase polymerase chain reaction (qPCR) was carried out as previously described 11, 12. Primer sequences are shown in supplemental Methods.

Nuclear-cytoplasmic fractionation and protein degradation assays. Subcellular fractionation and protein degradation assays were carried out as described in supplemental Methods.

Statistical analyses. Data are expressed as mean plus or minus SD. The statistical significance of differences of the means was determined by 1-way analysis of variance and multiple comparisons by Tukey-Kramer multiple range test.
RESULTS

The promoter proximal intron-exon structure is conserved across vertebrate species

In human, mouse and bovine vWF genes, the first exon has been shown to encode 5′ untranslated sequences and the second exon begins with the ATG translational start site. Exons 1 (+1 to +246 in human vWF) and 2 are separated by an intron of approximately 1200 bp (+247 to +1475 in human vWF). The exon/intron boundaries are highly conserved between mouse and human, and the splice donor and acceptor sites conform with the GT and AT rule. There is increasing evidence that proximal promoter introns, including 5’UTR introns, play an important role in mediating gene expression. Thus, we asked whether the proximal exon-intron arrangement in humans and mice is conserved across other vertebrate species. Indeed, as shown in Fig. 1, the same scheme (5’UTR followed by first intron followed by second exon beginning with ATG initiation codon) is found in chimpanzee, canine, and chicken vWF genes. In the case of the rat vWF gene, the first coding exon begins with ATG. A recent annotation for the zebrafish vWF gene assigned the translational start site to the beginning of exon 1 (of a total of 51 exons). To determine if the zebrafish gene contains an overlooked 5’ UTR-containing first exon (hence a previously unrecognized first intron), we carried out 5’ RACE using total RNA from zebrafish embryos. We found an additional 98 nucleotides of mRNA preceding the translational start and separated from it in the genome by 2866 nucleotides constituting a plausible intron. Thus, the promoter proximal intron-exon is highly conserved in vertebrates.
Intronic splicing enhances vWF promoter-luciferase reporter gene expression in cultured endothelial cells, but not megakaryocytes

Various lengths of the human vWF promoter with and without introns were coupled to firefly luciferase cDNA in pGL3 and the resulting plasmids were transfected into HUVEC. As an internal control for transfection efficiency, cells were co-transfected with a construct containing the CMV enhancer/promoter coupled to Renilla luciferase. vWF2-luc contains 2182 nucleotides upstream of the transcriptional start site, the first exon (+1-+246) and the first intron (+247-+1475) of the human vWF gene coupled to luciferase cDNA. vWFΔInt-luc is identical to vWF2-luc except that the first intron has been removed. As shown in Fig. 2A, luciferase activity was significantly lower in cells transfected with the vWFΔInt-luc construct compared with that of cells transfected with vWF2-luc. The need for the first intron of vWF could reflect a requirement for RNA splicing or it may indicate that the intron contains gene-specific transcriptional elements. To differentiate between these two possibilities, a plasmid was constructed in which the vWF intron was replaced with the second intron of the human β-globin gene (vWFΔIntGIn-luc). vWFΔIntGIn-luc activity was comparable to that of vWF2-luc. Thus, the enhancing effect of intronic sequence is not specific to the vWF intron. By contrast, when either the first intron from the vWF gene or the second intron of the human β-globin gene was placed upstream of the vWF promoter (in either direction), expression was no longer rescued (Fig. 2A shows Int-vWFΔInt, RInt-vWFΔInt, GInt-vWFΔInt). These latter findings argue against the presence of position- and orientation-independent
intronic enhancers and support a role for RNA splicing in mediating luciferase activity. We next wished to determine if the effect of splicing was restricted to the second intron of the human β-globin gene. To that end, constructs were generated in which the first intron of vWF was replaced with either the second intron from mouse Down syndrome critical region 1 (DSCR-1) or the sixth intron from factor X in hagfish. As shown in Fig. 2A, the DSCR-1 intron (vWFΔIntDIn-luc) had minimal effect on expression, whereas the hagfish intron (vWFΔIntHIn-luc) resulted in partial rescue of expression. Therefore, different introns have different effects on luciferase activity. To test whether the intron requirement is dependent on the promoter, we generated DNA plasmids in which a 2100-bp fragment of the human Tie2 promoter was placed upstream of luciferase in the absence or presence of an intervening vWF intron. The intron-containing and intronless Tie2 promoter constructs (Tie2P-vWFIn and Tie2P, respectively) were transiently transfected into HUVEC. Luciferase activity of these two constructs was identical (Fig. 2A). Thus, the permissive effect of splicing on gene expression is specific to the vWF promoter. Since the vWF gene is expressed in both endothelial cells and megakaryocytes, we assayed for an effect of RNA splicing in transiently transfected MEG-01 cells. Interestingly, there was no difference in luciferase activity in cells transfected with vWF2-luc or vWFΔInt-luc (Fig. 2A, inset). These findings suggest that splicing stimulates vWF expression in endothelial cells, but not megakaryocytes.

RNA splicing has been shown to influence gene expression at multiple steps, including at the levels of transcription, translation and protein degradation. To investigate the mechanisms underlying reduced luciferase activity in endothelial cells transfected with
intronless constructs, we measured luciferase mRNA levels using real-time PCR. These experiments revealed equivalent luciferase mRNA levels in cells transfected with vWF2-luc and vWFΔInt-luc (Fig. 2B), suggesting that the enhancing effect of RNA splicing on vWF-luciferase activity is mediated at a posttranscriptional level. Subcellular fractionation revealed comparable levels of cytoplasmic luciferase mRNA in vWF2-luc- and vWFΔInt-luc-transfected endothelial cells (Fig. 2B). These data argue against differential nuclear export of spliced and unspliced mRNAs. Finally, in protein degradation assays, we found that luciferase protein stability was comparable in the two transfectants (Fig. 2C). Thus, the increased luciferase activity of vWF2-luc (compared with vWFΔInt-luc) is not explained by a reduction in protein degradation.

The first intron of the mouse vWF gene is required for β-galactosidase activity in the endothelium of vWF-LacZ knockin mice

We previously demonstrated that expression of a human vWF promoter spanning the region between -2182 and +246 (namely vWF2) directed LacZ expression in blood vessels of the heart and skeletal muscle of Hprt-targeted mice in the presence, but not absence of the native first intron or the second intron from the human β-globin gene. To investigate whether rescue by the β-globin intron was mediated by the presence of enhancer elements, we targeted the Hprt locus of mice with a construct in which the intron was placed upstream of the vWF promoter. At this 5’ position, the β-globin intronic sequence failed to rescue expression in heart and skeletal muscle (Supplemental
These findings argue against a role for a position-independent enhancer within the β-globin intron.

A limitation of the Hprt locus targeting studies is that a defined length of the human vWF promoter (-2182 to +246, with or without an intron) is inserted into a foreign genomic locus (the Hprt locus) of mice. In contrast to Hprt-targeted vWF2, which drives reporter gene expression in blood vessels of the brain, heart and skeletal muscle, the endogenous vWF gene is expressed in many other vascular beds. Moreover, in vWF2-LacZ mice, expression of LacZ in the capillaries of heart and skeletal muscle is more widespread than that of vWF. Thus, while the Hprt locus targeting studies provide a valuable tool for dissecting mechanisms of vascular bed-specific gene expression, the extent to which the results may be extrapolated to the endogenous vWF gene is unclear. To circumvent this problem, we targeted LacZ to the endogenous vWF locus. The ATG initiation codon of LacZ was introduced immediately downstream of the first intron (in place of the ATG start site of vWF) (Fig. 3). Heterozygous mice (vWFLacZ+/+) were analyzed for LacZ expression. In real-time PCR assays, LacZ mRNA closely paralleled that of the endogenous vWF gene, with highest levels expressed in the lung, followed by the heart and skeletal muscle (Supplemental Fig. III). X-Gal staining of whole-mount organs and tissue sections revealed LacZ staining in multiple vascular beds, including the brain, lung, heart, liver, kidney and spleen (Fig. 4 and Supplemental Fig. IV). The pattern was identical to that observed with the endogenous vWF protein, with the exception that LacZ staining in the glomeruli was less intense compared with that of the endogenous protein.
These findings suggest that the LacZ knockin mouse is a suitable surrogate for tracking expression of the endogenous vWF gene. We next generated a LacZ knockin mouse in which the first intron was removed. In this case, the ATG start site of LacZ was placed immediately downstream of the first exon of vWF (Fig. 4). The resulting mice were bred to obtain heterozygotes (vWFΔIntLacZ/+). Whole-mount organs and tissue sections from these animals did not reveal any detectable reporter gene activity in the vasculature, with the exception of the brain where faint LacZ staining was observed in occasional superficial blood vessels (Fig. 4 and Supplemental Fig. V). In contrast, β-galactosidase activity in bone marrow megakaryocytes was unaffected by the loss of the first intron. However, in real-time PCR assays, LacZ mRNA expression in the various organs was identical to that of the intron-containing vWF^LacZ/+ mouse (Figs. 5A-5B). To validate these findings, we used magnetic beads to isolate endothelial cells from various organs (heart, lung, brain and diaphragm) of the vWF^LacZ/+ and vWFΔIntLacZ/+ mice and assayed for LacZ mRNA levels within 90 min of tissue harvest. Consistent with the findings in whole organs, LacZ transcript levels were comparable in endothelial cells from both lines of mice (Fig. 5C). In other experiments, endothelial cells from the heart were cultured for one-two weeks to obtain enough material for ChIP and subcellular fractionation assays. A limitation of primary mouse endothelial cell cultures is that vWF (and in the case of the knockin mice, LacZ) mRNA expression is rapidly downregulated (not shown). Nonetheless, Pol II ChIP revealed equal levels of association with the proximal vWF promoter, providing further support that gene transcription is unaffected by the loss of the first intron (Fig. 5D). Moreover,
subcellular fractionation assays showed comparable levels of LacZ mRNA in the cytoplasm of endothelial cells from vWF^{LacZ+/+} and vWFΔInt^{LacZ+/+} mice (Fig. 5E). These findings suggest that the loss of β-galactosidase activity in the absence of the first intron is mediated by a post-transcriptional mechanism that does not involve a defect in nuclear export of mRNA.

**Loss of endothelial expression in the intronless vWF-LacZ knockin mice is rescued by the inclusion of heterologous introns**

We next asked whether the introduction of heterologous introns into the endogenous vWF locus could rescue expression of LacZ. To that end, we chose the same three introns that were tested in transient transfection assays, namely the second intron of the human β-globin gene, the second intron of the mouse DSCR-1 gene and the sixth intron of the hagfish factor X gene. Each of these introns was inserted between the end of the first exon of the mouse vWF gene and the ATG translational start site of LacZ. The resulting animals were bred to heterozygosity and analyzed for LacZ expression. Consistent with the *in vitro* transfection data, introns from the β-globin and hagfish genes rescued expression in endothelial cells (Fig. 6 and Supplemental Fig. VI). In contrast to the cell culture findings, the DSCR-1 intron also rescued expression. While all three introns were able to promote LacZ expression in the endothelium, they exhibited quantitative and qualitative differences in their effects. For example, mice carrying the β-globin intron (vWFGIn^{LacZ+/+}) demonstrated the most complete rescue, while mice carrying the DSCR-1 intron (vWFDIn^{LacZ+/+}) demonstrated the least. Endothelial cells of the atria in
vWFGIn$^{LacZ/+}$ mice had much stronger staining compared with vWF$^{LacZ/+}$ mice or mice carrying the hagfish and DSCR-1 introns (Supplemental Fig. VII). In contrast, mice with the hagfish intron (vWFHIn$^{LacZ/+}$) expressed at much lower levels in arteries compared with the other lines of mice (Fig. 7 and Supplemental Fig. VIII). Collectively, the data support the conclusion that RNA splicing plays a role in mediating vWF expression in the endothelium.
DISCUSSION

Most genes in higher eukaryotes are interrupted by at least one intron. Introns must be precisely removed, or spliced from the initial pre-mRNA before nuclear export and translation. The primary role of introns is to generate increased proteomic diversity by alternative splicing. However, introns are also required for optimal gene expression in many systems, including yeast, plants, insects, mammalian tissue culture cells and transgenic mice. In some cases, intron-dependent gene expression is mediated by transcriptional control elements. For example, previous studies in transgenic mice have implicated a role for intronic enhancers in mediating the expression of vascular endothelial growth factor receptor 2, Tie2, GATA2 and endoglin. In other cases, splicing has been shown to directly stimulate the expression of genes. Our findings are the first to implicate a role for RNA splicing in mediating expression of an endothelial-restricted gene.

The life cycle of eukaryotic mRNA includes a number of steps that include transcription, capping, splicing, polyadenylation, nuclear export, and cytoplasmic translation and degradation. Each of these steps is carried out by tightly coupled multicomponent machines. Introns have been shown to influence gene expression at multiple steps along this pathway. For example, splicing has been reported to enhance mRNA transcription, promote 3’ end processing/polyadenylation and reduce mRNA decay. Our finding that luciferase and LacZ mRNA levels were unchanged in the absence of an intron argues against a role for any one of these mechanisms. Other studies have
implicated a role for mRNA splicing in mediating increased nuclear mRNA export\textsuperscript{16, 23, 41, 42}. However, our experiments in cultured cells failed to reveal an effect of splicing on the nuclear-cytoplasmic distribution of mRNA. Finally, there is increasing evidence that mRNA splicing enhances mRNA translational yield (i.e., the number of protein molecules per mRNA molecule)\textsuperscript{13, 15, 43–45}. Our finding that intronless vWF promoter-luciferase and vWF promoter-LacZ constructs produced normal levels of mRNA but decreased protein levels in transient transfections and LacZ knockin mice, respectively, indicate that the enhancing effects of splicing in the endothelium are mediated at this level. Increased translational yield may be explained by increased translation of spliced messages or increased stability of encoded polypeptide\textsuperscript{21}. We found that the luciferase protein was equally stable whether it was produced from an intronless or intron-containing vWF promoter construct, suggesting that the effects of splicing are secondary to increased translational efficiency.

The splicing reaction results in the deposition of proteins, termed the exon junction complex (EJC), at a conserved position, 20–24 nt upstream of exon-exon junctions\textsuperscript{46}. These proteins are involved in diverse processes including nuclear export, nonsense-mediated mRNA decay, and translational efficiency (reviewed in\textsuperscript{47}). Previous studies suggest that in some cases EJC deposition contributes to translational enhancement by increasing polysome association of the spliced mRNA\textsuperscript{48}. It has been proposed that EJC may promote stable formation of a translationally active mRNP through EJC-dependent recruitment of translation initiation factors, including RNPS1, Y14, Magoh, SR proteins and members of the hnRNP family of proteins (reviewed in\textsuperscript{47}). Alternatively, it has been
suggested that splicing may enable mRNA to escape translational masking of mRNA by proteins such as FRGY2 [Bouvet, 1994 #8654]. Whether or not these mechanisms are involved in mediating the effect of mRNA splicing on translational efficiency of vWF remains an open question.

The cDNA sequence of the gene is an important determinant of intron dependence. For example, in a previous study, the second intron from the rabbit β-globin gene was shown to increase expression of a β-globin cDNA fragment 400-fold, the mouse dihydrofolate reductase gene 10- to 20-fold, and the thymidine kinase gene only 2-fold. Another study investigated 10 different human and non-human genes for their intron dependence in vertebrate cells. Although all genes tested were expressed more efficiently in the presence of an intron, there were important quantitative differences: enhancement of human β-globin was 35-fold, whereas with 9 other cDNAs, it was 2- to 8-fold. Here, we have shown that the deletion of the first intron of vWF in vWFΔIntLacZ/ mice results in a virtual loss of detectable β-galactosidase activity in endothelial cells and that expression is rescued by heterologous introns. Thus, vWF expression in the endothelium appears to be completely dependent on RNA splicing.

Several other factors have been shown to influence the magnitude of intron-dependent effects on gene expression. For example, introns in the 5’ region of genes, as exemplified by the first intron of vWF, have a greater enhancing effect than those located in the coding sequence or 3’UTR. Also, the requirement for intronic splicing depends on the identity of the promoter. For example, introns were shown to be important for
expression from immunoglobulin or \( \beta \)-globin promoters, but not from a heat shock promoter\(^{49} \). Consistent with these data, we demonstrated that the first intron of vWF was unable to augment expression of the Tie2-luciferase construct in transient transfections. The reason for promoter specificity with respect to intron dependence is unclear. Perhaps, intron-insensitive promoters are more heavily transcribed, and thus do not require intronic splicing. Alternatively, it has been proposed that components involved in RNA processing or export may tether to the promoter region\(^{49} \).

Finally, in addition to exon sequence context and promoter identity, the effect of mRNA splicing on gene expression is dependent on the sequence of the intron\(^ {13, 20, 28} \). For example, in transgenic mice carrying the human histone H4 promoter couple to chloramphenicol acetyltransferase, a hybrid intron had a much more pronounced effect on expression compared with the SV40 small-t intron\(^ {20} \). In our study, the native vWF first intron and each of the heterologous introns had different effects on the magnitude of LacZ expression in the vasculature, with the \( \beta \)-globin intron exhibiting the most complete rescue. The mechanism underlying intron-specific effects may include inefficient splicing of some introns due to weak splice sites. Indeed, one previous study demonstrated that the introduction of heterologous introns in dihydrofolate reductase minigenes resulted in less efficient splicing \textit{in vitro} compared with the native intron\(^ {21} \). Alternatively, the difference may reflect the presence of intron-specific regulatory elements that affect others steps in the mRNA life cycle.
An interesting observation was that the various heterologous introns resulted in qualitatively different patterns of rescue. Most striking was the increased β-galactosidase activity in the endothelium of the atria of the vWFGInLacZ/+ mice and the reduced expression of LacZ in the arteries of the vWFHInLacZ/+ mice. These data indicate that the identity of the intron determines efficiency of splicing (and/or some other component of the life cycle of vWF mRNA). Our findings have two important implications. First, they suggest that the native first intron contributes to the spatial regulation of vWF. Second, they raise the interesting possibility that introns play a role in mediating differential expression of genes in the vasculature.

Our previous studies in transgenic mice and Hprt-targeted mice suggest that the transcriptional mechanisms underlying vWF expression differ between endothelial cells and megakaryocytes. Specifically, each of the human and mouse vWF promoter fragments tested to date has been shown to direct expression in endothelial cells, but not bone marrow. Thus, DNA sequences outside these regions are likely to be required for expression in megakaryocytes. The results of the present study suggest that the role of RNA splicing in mediating vWF expression also differs between endothelial cells and megakaryocytes. Type 1 von Willebrand disease (vWD) is found in persons who have partial quantitative deficiency of vWF. A subgroup of individuals with Type I vWD have reduced plasma vWF, but normal platelet vWF:Ag and VWF:Rco (platelet-normal). It has been assumed, though never proven, that the endothelium of these individuals also has normal vWF:Ag and that the very low plasma levels in these is secondary to defective release of vWF into the plasma and/or increased clearance of vWF.
from the plasma⁵⁴,⁵⁵. However, our findings raise the interesting possibility that a
splicing defect could result in a vWD phenotype in which vWF expression is
preferentially lost in endothelial cells.
ACKNOWLEDGEMENTS

We thank Paula Frankel for providing us with zebrafish RNA. We thank Nadia Jahroudi for her helpful input. This work was supported by National Institutes of Health grant HL076540.
AUTHORSHIP CONTRIBUTIONS AND DISCLOSURE OF CONFLICTS OF INTEREST

Contribution: L.Y., L.J., D.B., K.C.S., J.S., D.L., S.J. designed and performed experiments and analyzed the data; P.O. designed experiments and analyzed the data; and W.C.A. designed and performed experiments, analyzed the data, and wrote the manuscript.

The authors have no relevant financial conflict of interest to disclose.
REFERENCES


FIGURE LEGENDS

Figure 1. Promoter proximal intron-exon structure of the vWF gene across vertebrate species. Shown in order are the first exon (5’UTR), first intron, second exon (which begins with the ATG translation initiation codon), second intron, third exon and part of the third intron in multiple species. In the case of the rat gene, there is no information about the genomic sequence upstream of the translation start site. Thus, the presence of an upstream first exon containing the 5’UTR (box with dashed outline) and a first intron (dashed line) can only be inferred from their conserved arrangement in higher and lower species.

Figure 2. Effect of RNA splicing on vWF promoter activity in cultured primary human endothelial cells and MEG-01 cells. A, HUVEC were transiently transfected with a pGL3 construct containing the human vWF promoter spanning the region between -2182 and +246 (which corresponds to the end of the first exon) without an intron (vWFΔInt), or linked 5’ of the first intron from the human vWF gene (vWF2), the second intron from the human β-globin gene (vWFΔIntGIn), the second intron from the mouse DSCR-1 gene (vWFΔIntDIn), or the sixth intron from the hagfish factor X gene (vWFΔIntHIn). Alternatively, cells were transfected with a pGL3 construct in which the first intron of the vWF gene was placed upstream of the -2182 and +246 vWF promoter in the correct orientation (Int-vWFΔInt) or in the reverse direction (RInt-vWFΔInt), with a a pGL3 construct in which the second intron from the human β-globin gene was placed upstream of the -2182 and +246 vWF promoter in the correct orientation (GInt-
vWFΔInt), or with a pGL3 construct containing a 2100-bp fragment of the mouse Tie2 promoter linked in the absence or presence of the first intron of human vWF (Tie2P, Tie2P-vWFIn, respectively). Inset, MEG-01 cells were transiently transfected with a pGL3 construct containing the human vWF promoter spanning the region between -2182 and +246 (which corresponds to the end of the first exon) in the absence or presence of the first intron from the human vWF gene. The results show the mean ± SD of luciferase light units (relative to vWF2) obtained in triplicate from at least 3 independent experiments. *P < .05. n.s., non-significant. B, HUVEC were transfected with vWF2-luc or vWFΔInt-luc constructs. After 24 h incubation, total or cytoplasmic (cyto) extract fractions were prepared. RNA was extracted and luciferase mRNA levels were measured by qPCR. The results show the mean ± SD of luciferase mRNA levels (relative to vWF2) obtained in triplicate from 3 independent experiments. n.s., non-significant. C, HUVEC were transfected with vWF2-luc or vWFΔInt-luc plasmids. After 24 h incubation, cells were treated with cyclohexamide (100 μg/ml) and collected at indicated time points to measure the luciferase activity. Data are presented as percentage relative to 0 h. N=3 independent experiments.

**Figure. 3. Schematic of the targeting constructs used to generate mice.** E, exon; ATG, translation initiation codon.

**Figure 4. LacZ staining of whole-mount organs from vWF<sup>LacZ<sub>+</sub></sup> and vWFΔInt<sup>LacZ<sub>+</sub></sup> mice.** Hearts, chest walls, en face aortae, brains, and diaphragms and were harvested from 6- to 8-week-old male vWF<sup>LacZ<sub>+</sub></sup> mice (A, C, E, G, H) and vWFΔInt<sup>LacZ<sub>+</sub></sup> mice (B,
D, F, G, H) and processed in parallel for whole-mount staining with X-Gal. In G, the brain from the vWF\textsuperscript{LacZ/+] mouse is on the left, and the brain from the vWF\textsuperscript{ΔIntLacZ/+] mouse is on the right. In H, the diaphragm from the vWF\textsuperscript{LacZ/+] mouse is on the left, and the diaphragm from the vWF\textsuperscript{ΔIntLacZ/+] mouse is on the right. Whole-mount tissues were analyzed under a Nikon SMZ-U dissecting microscope, and microphotographs were collected using a Nikon Coolpix 8400 camera.

**Figure 5. Mechanisms underling enhancing effect of first intron on vWF\textsuperscript{LacZ/+] gene expression.** A, Organs from vWF\textsuperscript{LacZ/+] and vWF\textsuperscript{ΔIntLacZ/+] mice were assayed for LacZ mRNA expression by qPCR. The results show the mean ± SD of mRNA expression (relative to vWF\textsuperscript{LacZ/+] in brain) obtained in triplicate from 3 independent experiments. Br, brain; He, heart; Lu, lung; Li, liver; Ki, kidney; Di, diaphragm; SM, skeletal muscle; Ao, aorta; Sp, Spleen. B, Bone marrow from vWF\textsuperscript{LacZ/+] and vWF\textsuperscript{ΔIntLacZ/+] mice was assayed for LacZ mRNA expression by real-time PCR. The results show the mean ± SD of mRNA expression (relative to vWF\textsuperscript{LacZ/+] ) obtained in triplicate from 3 independent experiments. n.s., non-significant. C, Endothelial cells were harvested from different organs from vWF\textsuperscript{LacZ/+] and vWF\textsuperscript{ΔIntLacZ/+] mice using CD31-coated magnetic beads. RNA was isolated within 90 min and assayed for LacZ mRNA expression by qPCR. The results show the mean ± SD of mRNA expression obtained in triplicate from 3 independent experiments. D, ChIP assay was performed using double-sorted endothelial cells from the hearts of vWF\textsuperscript{LacZ/+] and vWF\textsuperscript{ΔIntLacZ/+] mice. DNA was sheared and the resulting DNA-protein complexes were immunoprecipitated in the absence or presence of antibodies to PolII or control IgG. Real-time PCR analysis was performed using the
precipitated DNA fragments and primers for the proximal promoter region. E, mRNA in
cytoplasmic and nuclear fractions or total mRNA was extracted from heart ECs of
vWF\textsuperscript{LacZ/+} and vWF\textsuperscript{ΔIntLacZ/+} mice. LacZ mRNA levels were measured by qPCR using
gene-specific primers. The results show the mean ± SD of mRNA expression (relative to
vWF\textsuperscript{LacZ/+}) obtained from 5 independent experiments.

**Figure 6.** LacZ staining of whole-mount organs and bone marrow from vWF-LacZ knockin mice with or without native or heterologous intron. Hearts (A, B, C, D, E),
diaphragms (F, G, H, I, J), chest walls (K, L, M, N, O), brains (P, Q, R, S, T), and en face
aortae (U, V, W, X, Y) were harvested from 6- to 8-week-old male vWF\textsuperscript{LacZ/+},
vWF\textsuperscript{ΔIntLacZ/+}, vWFG\textsuperscript{LacZ/+}, vWF\textsuperscript{ΔIntLacZ/+} and vWFH\textsuperscript{LacZ/+} mice and processed in
parallel for whole-mount staining with X-Gal. The chest wall shows the intercostal space
between two ribs (a superior rib on the left and an inferior rib on the right). Bone marrow
aspirates (Z, Aa, Ab, Ac, Ad) were stained for LacZ. Whole-mount tissues were analyzed
under a Nikon SMZ-U dissecting microscope, and microphotographs were collected
using a Nikon Coolpix 8400 camera. Scale bar, 200 μM.

**Figure 7.** Comparison of LacZ staining in arteries from vWF-LacZ knockin mice
with or without native or heterologous intron. Organs were harvested from 6- to 8-
week-old male mice and processed in parallel for whole-mount staining with X-Gal.
Shown are basilar artery (thick arrow), middle cerebral artery (thin arrow) and internal
thoracic artery (asterisk) from vWF\textsuperscript{LacZ/+} mice (A, B, and C respectively), vWF\textsuperscript{ΔIntLacZ/+}
mice (D, E, and F respectively), vWFG\textsuperscript{LacZ/+} mice (G, H, and I, respectively),
vWFΔInLacZ+/ mice (J, K, and L, respectively) and vWFΔInLacZ+/ mice (M, N, and O, respectively). Note the reduced LacZ staining in basilar artery, middle cerebral artery, and internal thoracic artery from vWFΔInLacZ+/ mice. Whole-mount tissues were analyzed under a Nikon SMZ-U dissecting microscope, and microphotographs were collected using a Nikon Coolpix 8400 camera.
Figure 1

- **vWF**
  - vWF\(^{+/+}\)
  - vWF\(^{LacZ/+}\)
  - vWF\(^{ΔInLacZ/+}\)
  - vWFGIn\(^{LacZ/+}\)
  - vWFDIn\(^{LacZ/+}\)
  - vWFHIn\(^{LacZ/+}\)

- **Annotations**
  - Mouse VWF 1\(^{st}\) intron
  - Human β-globin 2\(^{nd}\) intron
  - Mouse DSCR-1 2\(^{nd}\) intron
  - Hagfish factor X 6\(^{th}\) intron
**Figure 2**

- **Homo sapiens**
  - Exon 1: 250 bp
  - Exon 2: 1224 bp
  - Exon 3: 1803 bp

- **Pan troglodytes (Chimpanzee)**
  - Exon 1: 269 bp
  - Exon 2: 1224 bp
  - Exon 3: 1809 bp

- **Canis lupus Familiaris (Dog)**
  - Exon 1: 104 bp
  - Exon 2: 1186 bp
  - Exon 3: 1469 bp

- **Bos taurus (Cattle)**
  - Exon 1: 233 bp
  - Exon 2: 1860 bp
  - Exon 3: 1492 bp

- **Rattus Norvegicus (Rat)**
  - Exon 1: ? bp
  - Exon 2: ? bp
  - Exon 3: 829 bp

- **Mus musculus (Mouse)**
  - Exon 1: 252 bp
  - Exon 2: 1253 bp
  - Exon 3: 842 bp

- **Gallus gallus (Chicken)**
  - Exon 1: 237 bp
  - Exon 2: 1219 bp
  - Exon 3: 1236 bp

- **Danio rerio (Zebrafish)**
  - Exon 1: 97 bp
  - Exon 2: 2866 bp
  - Exon 3: 1603 bp
Figure 3

A.

vWF2
vWFΔInt
vWFΔIntGln
Int-vWFIntΔInt
RInt-vWFΔInt
GInt-vWFΔInt
vWFΔIntDln
vWFΔIntHln
HInt-vWFΔInt
Tie2P
Tie2P-vWFInt

Relative units

Luciferase activity

B. Luciferase mRNA

Luciferase mRNA

C. Luciferase activity

Total RNA

Cyto-RNA

n.s.
**Figure 5**

A. LacZ

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

B. Relative units

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

C. LacZ mRNA

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

D. Mouse heart endos

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

E. LacZ mRNA

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

Cytoplasmic

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

Nuclear

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**

Total

- **vWFΔIntLacZ/+**
- **vWFΔIntLacZ/+**
Figure 6  

**vWF**\textsuperscript{LacZ/+}  
**vWFΔInt**\textsuperscript{LacZ/+}  
**vWFGln**\textsuperscript{LacZ/+}  
**vWFDln**\textsuperscript{LacZ/+}  
**vWFHln**\textsuperscript{LacZ/+}  

**Heart**

A  B  C  D  E

**Diaphragm**

F  G  H  I  J

**Chest wall**

K  L  M  N  O

**Brain**

P  Q  R  S  T

**Aorta**

U  V  W  X  Y

**Bone marrow**

Z  Aa  Ab  Ac  Ad
Figure 7

Basilar artery

vWF LacZ/+  

vWFΔInt lacZ/+  

vWF Gln lacZ/+  

vWF Dln lacZ/+  

vWF Hln lacZ/+  

Cerebral artery

vWF LacZ/+  

vWFΔInt lacZ/+  

vWF Gln lacZ/+  

vWF Dln lacZ/+  

vWF Hln lacZ/+  

Thoracic artery/vein

*
Role of RNA splicing in mediating lineage-specific expression of the von Willebrand factor gene in the endothelium

Lei Yuan, Lauren Janes, David Beeler, Katherine C. Spokes, Joshua Smith, Dan Li, Shou-Ching Jaminet, Peter Oettgen and William C. Aird