Functional Polymorphism in the Matrix Metalloproteinase-9 Promoter as a Potential Risk Factor for Intracranial Aneurysm

David G. Peters, PhD; Amin Kassam, MD; Pamela L. St. Jean, PhD; Howard Yonas, MD; Robert E. Ferrell, PhD

Background and Purpose—There is convincing evidence that susceptibility to intracranial aneurysms (ICAs) has a genetic component. However, few studies have sought to identify functional variation in specific candidate genes that may predispose individuals to develop an ICA.

Methods—ICA cases and controls were genotyped for a simple length polymorphism in the promoter of matrix metalloproteinase-9 (MMP-9) to test for association between variation in the promoter and the occurrence of ICA. Alternative alleles were cloned into an in vitro reporter vector, transfected into human HT1080 fibroblasts, and assayed for promoter activity by β-gal and luciferase assays. Electrophoretic gel shift assays were used to assess nuclear factor binding.

Results—A length polymorphism in the promoter of MMP-9 was nonrandomly associated with the occurrence of ICA in a case-control study. This polymorphism was shown, by direct sequencing of 36 individuals, to be the only sequence variation within a 736-base pair region proximal to the transcriptional start site of the gene. Variation in the length of this repetitive element was shown to modulate promoter activity in an in vitro reporter assay, with the highest promoter activity being observed in constructs bearing the longest [(CA)23] element. Electrophoretic mobility shift assays were used to show that the (CA) element is bound by a sequence-specific DNA-binding protein.

Conclusions—Genetic variation in the promoter of the MMP-9 gene results in variation in its expression at the level of transcription. This may result in subtle differences in MMP-9 activity within the circle of Willis, leading to increased susceptibility to ICA formation. (Stroke. 1999;30:2612-2616.)

Key Words: biological markers ■ cerebral aneurysm ■ genetics ■ polymorphism

Matrix metalloproteinase-9 (MMP-9; gelatinase B, type 4 collagenase) is a member of the MMP gene family, which encodes a family of zinc-dependent enzymes with proteolytic activity against connective tissue proteins, including collagens, elastin, and proteoglycans. MMP-9 plays an essential role in development and tissue remodeling.1 MMP-9 overexpression has been reported in abdominal aortic and cerebral aneurysms,2–5 and its overexpression contributes to the formation of aneurysms through the degradation of type 4 collagen, proteoglycan core protein, and elastin, which are resistant to degradation by some other MMPs.6 MMP-9 is regulated primarily at the level of transcription in response to such regulatory molecules as tumor necrosis factor-α, interleukin-1, platelet-derived growth factor, and epidermal growth factor.7,8

There is significant evidence that susceptibility to cerebral aneurysms has an important genetic component,9,10 with an imbalance in the local expression of MMP-9 and tissue inhibitors of metalloproteinases being a contributing factor.4 We have previously described a simple sequence repeat polymorphism in the 5′-untranslated region of the human MMP-9 gene11 within the regulatory domain. The present study was undertaken to test the hypothesis that polymorphic variation in this simple sequence repeat has a significant influence on MMP-9 expression.

Materials and Methods

Study Subjects

This research was approved by the Institutional Review Board of the University of Pittsburgh, and all participants gave written informed consent. Peripheral blood specimens and demographic, medical, and family histories were obtained from 76 sequentially ascertained, unrelated patients undergoing repair of a cerebral aneurysm at Presbyterian University Hospital, Pittsburgh, Pa. Eight (10%) of these patients reported a history of symptomatic intracranial aneurysm (ICA) among their first- and second-degree relatives, with no obvious pattern of occurrence. Among these cases, 70% were female.
and 30% were male, with a mean age of 54.9 and 46.6 years, respectively \((P<0.005)\). All participants were of mixed west
European white ancestry, and none had a personal or family history of
connective tissue disorders or polycystic kidney disease. These cases
were compared with an unselected sample of participants in observa-
tional epidemiological studies selected without reference to dis-
ease. The controls were drawn from the same population of mixed
European ancestry as the cases (western Pennsylvania) and had a
similar age and sex distribution. The controls were uncharacterized
with respect to vascular disease.
High-molecular-weight genomic DNA for genotyping and se-
quence analysis was prepared by standard methods.\(^9\) The MMP-9
CA-repeat polymorphism was genotyped as previously described.\(^1\)

**Data Analysis**

Allele frequencies were estimated by gene counting. Allele frequen-
cies in cases and controls were compared by standard \(x^2\) analysis.
Means and SDs of luciferase assays were computed by standard
methods, and significance of differences was assessed by ANOVA F
test.

**DNA Sequencing**

CLG4B sequence was amplified from 50 ng of genomic DNA with
taq DNA polymerase (BRL) in 20 mmol/L Tris-HCl (pH 8.4),
500 mmol/L KCl, 1.5 mmol/L MgCl\(_2\). Amplicons were purified to
remove unincorporated primer and dNTPs with Microcon 100 spin
columns (Amicon) and subjected to cycle sequencing on an ABI
9600 (PE Applied Biosystems) with dye-labeled terminators. Se-
quences were purified on Centricon columns (Princeton Separations)
and analyzed on an ABI 377 automated sequencer (PE Applied
Biosystems).

**Subcloning Polymerase Chain Reaction Products of CLG4B Promoter Sequence**

Polymerase chain reactions (PCRs) were performed with the high-
fidelity thermostable DNA polymerase \(pfu\) (Stratagene) and 50 ng of
genomic DNA in 100 mmol/L KCl 100 mmol/L \((NH_4)_2SO_4,
200 mmol/L Tris-Cl (pH 8.75), 20 mmol/L MgSO\(_4\). 1% Triton
X-100, 1000 \(\mu g/mL\) BSA. Primers were designed to incorporate
either a \(Sma-I\) or \(Kpn-I\) restriction recognition sequence at either end
of the resulting PCR product and were identical to P1 and P2 (see
below). The resulting 736-bp PCR products were purified on 1% agarose
gels and subjected to restriction digestion. After an addi-
tional round of gel purification, fragments were ligated with \(Sma-I:\)
\(Kpn-I\) double-digested pGL3B reporter vector (Promega) and trans-
formed into chemically competent XL1 Blue Escherichia coli
(Stratagene). Recombinant plasmid DNA was purified by miniprep
(Qiagen) and confirmed by sequencing. Large-scale preparations
(Qiagen) of recombinant plasmid DNA were used for transfection of
HT1080 (ATCC) human fibroblasts.

**Transfection of Human HT1080 Fibroblasts**

Cells were plated at a density of 0.75 cells/mL in 2 mL of DMEM
containing 10% fetal bovine serum and penicillin/streptomycin (Life
Technologies) and grown overnight until \(\sim 70%\) confluent. Trans-
fections were performed in either duplicate or triplicate with Lipo-
fectamine (Life Technologies) and grown overnight until \(\sim 70%\) confluent. Trans-
fections were performed in either duplicate or triplicate with Lipo-
fectamine (Life Technologies) according to the manufacturer’s
instructions with a 5-hour incubation in the presence of the Lipo-
fectamine. For each plate, 2 \(\mu g\) of reporter plasmid and 0.5 \(\mu g\) of a
cytomegalovirus (CMV)-driven \(\beta\)-gal positive control vector
(pCMV, In Vitrogen) were used. Cell extracts were assayed for \(\beta\)-gal
and luciferase activity by use of commercially available kits (Promega).
\(\beta\)-Gal levels were measured by spectrophotometry, and luciferase
activity was determined by use of a luminometer (Monolight).

**Electrophoretic Mobility Shift Assay**

CLG4B nucleotide numbers are assigned based on the sequence of
Huhtala\(^4\) and assume a microsatellite length of 21 repeats. Double-
stranded probes for electrophoretic mobility shift assays were con-
built with a short common primer to fill in the second strand of a
full-length single-stranded oligonucleotide. The common primer was
radio-labeled at its 5’ end with polynucleotide kinase (Boehringer
Mannheim) in the presence of \(a\)^32\(_P\)-dCTP and purified by use of
G25 spin columns (Boehringer Mannheim). Oligonucleotides were
annealed by boiling for 3 minutes and cooling slowly to room
temperature. Extension reactions were performed with the Klenow
fragment of DNA polymerase I (Life Technologies). The (CA)
probe element contains 23 CA repeats with 44 bp of flanking sequence corresponding to nucleotides \(-148\) to \(-63\).\(^1\) The “Re-
place” probe is identical except that the CA motif is replaced by a
sequence of the same length derived from the polylinker of the
pGL3Basic vector (nucleotides 31 to 66 [Promega]). Oligonucleotide
sequences are as follows:

- (CA) element oligonucleotide: TCTCATGCTGTGTCAGGCA-
  CACACACAACACACACAACACACACACACACACACACACACACACACAC-
  ACACCTGACCCCTGGTACTGTTG-
  CACACACACACACACACACACACACACACACACACACAC-

- Replace oligonucleotide: TCTCATGCTGTGTCAGGCA-GATCCAGGATCA-
  TAAAGGCTTGGGACTTCGGGACTGTTG-TG-

Common primer: TCTCATGCTGTG.

**Preparation of Nuclear Extracts**

HT1080 and bovine aortic endothelial cells (BAECs) were cultured
to 70% confluence and crude nuclear extracts prepared by a
modification of published methods.\(^1\) Extracts were measured for
protein content using the Bradford method (Bio-Rad) and frozen at
\(-80^\circ C\).

**Electrophoretic Mobility Shift Assays**

Binding reactions were performed in binding buffer (25 mmol/L
HEPES pH 7.9, 500 mmol/L KCl, 10 mmol/L DDT, 10 mg/mL
BSA) containing 15% glycerol, 2 to 5 pmol of probe, and 2.0 \(\mu g\)
of poly (dl-dC)(dl-dC) (Pharmacia). Either 1, 3, or 5 \(\mu g\) of crude
nuclear extract was added to this, to a final volume of 20 \(\mu L\).
Reactions were incubated on ice for 30 minutes, run on 4% nondenaturating polyacrylamide gels, and electrophoresed for 2 to 3
hours at 4°C and 10 \(\mu A/cm\). After electrophoresis, gels were trans-
ferred to Whatman 3 MM paper, dried under vacuum at 80°C, and
subjected to autoradiography.

**Results**

**MMP-9 Genotyping and Sequencing**

MMP-9 allele frequencies in cases and controls are shown in
Table. There is a significant difference in the allele
frequency between cases and controls, with an excess of
(CA)\(_{23}\) alleles among cases.
These cases were unrelated, sequentially ascertained subjects who were undergoing operative aneurysm repair. Approximately 10% of these cases report a history of an ICA in a first-degree relative, but the frequency of the (CA)$_{23}$ allele was not different between those with and those without a family history of ICA. The small number of cases with a positive family history ($n=8$) precludes the detection of small differences in allele frequency. The frequency of (CA)$_{23}$ homozygous individuals did not differ significantly between cases and controls. To test the hypothesis that the microsatellite variation in the regulatory region of MMP-9 is responsible for the association with ICA, we sequenced a 740-bp region of the MMP-9 promoter from 18 unrelated aneurysm cases and 18 control individuals. The only DNA sequence variation in this region was found to be within the (CA)$_n$ sequence that was used for the genotyping experiment. A schematic of the MMP-9 promoter is given in Figure 1, with the CA repeat and key promoter features shown. We reasoned that such placement of a variable-length repetitive sequence between both proximal and distal regulatory elements that may bind a variety of transcription factors$^{14}$ may have a subtle, yet significant effect on transcription efficiency. To test this hypothesis, we tested multiple alleles of the CA element for transcriptional activity.

Transfection Analysis of Constructs Bearing Different Length (CA) Elements

To determine whether the variable length of (CA)$_n$ repeat is able to modulate transcription of the CLG4B gene, a 736-bp region of the CLG4B promoter encompassing the CA repeat sequence was amplified by PCR from individuals homozygous for microsatellite length and cloned into the luciferase reporter pGL3Basic (Promega). Constructs were assayed for their ability to drive expression of the reporter in human fibroblasts (HT1080). To ensure transfection of equal amounts of each construct, plasmid concentration was determined by quantification of UV-stained linearized vector with the Stratagene Eagle Eye System and also by spectrophotometry. Luciferase values were normalized to $\beta$-gal activity generated by cotransfection of a CMV-driven internal control plasmid (In Vitrogen). We found consistent differences over 5 independent sets of experiments ($F_{1,16}=8.78; P=0.001$; see Materials and Methods) between MMP-9 constructs bearing inserts with different length microsatellite repeats (Figure 2). Specifically, promoters containing 14, 21, and 22 CA repeats gave $\approx 0.59$-, 0.36-, and 0.61-fold (respectively) the activity of the promoter containing 23 CA repeats, which consistently resulted in the highest luciferase activity. These specific length repeats were chosen because they are the 4 most common alleles found in normal white populations$^{11}$ (see Table). Although the (CA)$_n$ sequence is able to modulate transcriptional activity in this in vitro system, it is unclear whether this effect is due to a direct interaction between the transcriptional machinery and the CA repeat or is merely due to a conformational change in the 3D structure of the active promoter caused by variation in its length.

Electrophoretic Mobility Shift Assays

To further investigate whether the CA repeat has direct involvement in modulating transcriptional activity, we investigated site-specific DNA binding activity in HT1080 cells and BAECs using a 90-bp DNA sequence encompassing the CA repeat (see Figure 1). As a control probe, we used a sequence of identical length in which the (CA) element was replaced by a “benign” sequence derived from the pGL3Basic polylinker, which has no ability to drive luciferase expression in HT1080 cells (not shown). Sequence-specific DNA binding involving the CA repeat sequence (complex E and possibly complex B in Figure 3) was observed in both cell types. Complex E appears at the lowest concentration of protein and is eliminated (or reduced) at higher concentrations, whereupon complex B seems to become more prominent. This suggests a concentration-dependent cooperative binding effect. Other complexes (A, C, and D) appear to be shared between probes and may therefore involve the putative activator protein-1 (AP-1) binding site immediately downstream of the (CA) element (Figure 1).

The probe sequences used in the above experiment encompass the adjacent downstream putative AP-1 binding site. To confirm that the observed sequence-specific DNA-binding
Discussion

We report that variation within the regulatory region of the MMP-9 gene may be associated with ICA. A microsatellite marker used for typing MMP-9 variation in our sample and control populations was found to be the only polymorphic site within a 736-bp region upstream of the MMP-9 coding sequence. Differences in the length of this repeat sequence vary in the human population between 13 and 23 repeats. The (CA)_{23} repeat was nonrandomly associated with the occurrence of aneurysm in our case-control study. The repetitive element is situated in the MMP-9 promoter in a region previously shown to be essential for transactivation. Not only did variation in (CA)n element length modulate transcriptional activity in vitro, but the highest activity was seen when identical probes were used in which the (CA) element was replaced by a benign sequence of equal length and therefore represented CA-specific DNA binding protein. Electrophoretic mobility shift assay probes containing the (CA) element and adjacent downstream putative AP-1 site were bound by a sequence-specific factor, as were probes lacking the AP-1 element. Such interactions were not seen when identical probes were used in which the (CA) element was replaced by a benign sequence of equal length and therefore represented CA-specific DNA binding factors. These interactions do not depend on the presence of the AP-1 site. These data suggest that (CA) element length has a subtle, yet significant effect on MMP-9 promoter activity. This effect may be due in part to variation in the spacing between the proximal and distal regulatory elements and may be mediated by a (CA) element sequence-specific DNA-binding protein.

There has been limited and inconclusive experimentation to identify an underlying biochemical imbalance that may be significant in ICA development. This effort has largely focused on factors that are either constituents of the extracellular matrix or are involved in its homeostasis. Such interest focused on factors that are either constituents of the extracellular matrix or are involved in its homeostasis. Such interest is not surprising given that ICAs are the result of an acute lesion at sites subject to considerable hemodynamic insult, which implies fundamental weakness due to structural defects or dysregulation of normal turnover in the arterial wall. A number of epidemiological risk factors are consistent with respect to the susceptibility to and pathogenesis of a number of complex multigenic diseases. The relatively low frequency of ICA patients.

Although the differences in length of the MMP-9 CA repeat alleles are small, there are a number of reasons why they may be significant. For example, we have shown that the (CA) element may serve as a binding site for a specific regulatory protein. This concept is not without precedence, because it has been suggested that the estrogen receptor may be involved in the formation of DNA conformational structures, and a GC-binding transcription factor may regulate SV-40 promoter activity in a model system. Variation in the length of the repeat within the promoter element may have a dramatic effect on the dynamics of DNA-protein interactions and lead to altered MMP-9 expression. In addition, there is considerable evidence that transcriptional regulation involves a complex 3-dimensional interaction between transcription factors at numerous sites along the DNA. It is possible that small alterations in these structures may have a significant impact on gene expression. Of particular significance with respect to MMP-9 regulation is the discovery that the transcription factor AP-1 is able to interact with other components of the transcriptional machinery, including GATA-2. It is intriguing that the CA repeat found in the MMP-9 promoter sits between putative binding sites for these factors and an adjacent AP-1 site (Figure 1). The fact that simple dinucleotide repeats can form unusual DNA-like conformations supports these notions. Subtle alterations in conformations of protein-protein interactions may thus modulate transcriptional activity and moderate levels of MMP-9 mRNA.

Although the association between trinucleotide repeat length and human disease is well characterized, this is the first reported example of which we are aware of the modulation of promoter function by a dinucleotide repeat within the context of normal human variation. Because microsatellite sequences are commonly found in human gene promoters, their ability to modulate transcription may be significant with respect to the susceptibility to and pathogenesis of a number of complex multigenic diseases. The relatively low frequency of the (CA)_{23} allele in the general population means that this variation may represent only one of multiple genetic and environmental factors that contribute to the risk of ICA. These findings now require confirmation in other ICA populations.

Acknowledgment

This work was supported in part by NIH grant No. HL44682.
References


Functional Polymorphism in the Matrix Metalloproteinase-9 Promoter as a Potential Risk Factor for Intracranial Aneurysm

David G. Peters, Amin Kassam, Pamela L. St. Jean, Howard Yonas and Robert E. Ferrell

Stroke. 1999;30:2612-2616
doi: 10.1161/01.STR.30.12.2612

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/30/12/2612

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/