RAPID COMMUNICATION

A Common Genetic Variation in the 3'-Untranslated Region of the Prothrombin Gene Is Associated With Elevated Plasma Prothrombin Levels and an Increase in Venous Thrombosis

By Swibertus R. Poort, Frits R. Rosendaal, Pieter H. Reitsma, and Rogier M. Bertina

We have examined the prothrombin gene as a candidate gene for venous thrombosis in selected patients with a documented familial history of venous thrombophilia. All the exons and the 5'- and 3'-UT region of the prothrombin gene were analyzed by polymerase chain reaction and direct sequencing in 28 probands. Except for known polymorphic sites, no deviations were found in the coding regions and the 5'-UT region. Only one nucleotide change (a G to A transition) at position 20210 was identified in the sequence of the 3'-UT region. Eighty percent of the patients had the 20210 AG genotype, as compared with 1% of a group of healthy controls (100 subjects). In a population-based case-control study, the 20210 A allele was identified as a common allele (allele frequency, 1.2%; 95% confidence interval, 0.5% to 1.8%), which increased the risk of venous thrombosis almost threefold (odds ratio, 2.8; 95% confidence interval, 1.4 to 5.6). The risk of thrombosis increased for all ages and both sexes. An association was found between the presence of the 20210 A allele and elevated prothrombin levels. Most individuals (87%) with the 20210 A allele are in the highest quartile of plasma prothrombin levels (>1.15 U/mL). Elevated prothrombin itself also was found to be a risk factor for venous thrombosis.

© 1996 by The American Society of Hematology.

DEEP-VEIN THROMBOSIS is a common disease, with an annual incidence in the general population of approximately 1 per 1,000.1 Risk factors include both hereditary and acquired conditions.2 Generally, a tendency toward venous thrombosis could arise from hyperactive coagulation pathways, hypoactive anticoagulant mechanisms, or hypoactive fibrinolysis.3 Mutations in genes that encode proteins in these pathways play an important role in the predisposition to venous thrombosis.4 Variant alleles of the genes encoding protein C,3 protein S,4 antithrombin,7 and fibrinogen8 have been shown to be relatively strong, but uncommon risk factors for thrombosis.9 Genetic analysis of these genes showed a large heterogeneity of mutations.9,8 More recently, a poor anticoagulant response of plasma to activated protein C (APC)10,11 due to the presence of a mutant factor V molecule12 (factor V Leiden) was discovered and is as yet the most common hereditary risk factor for thrombosis known. Recently, some support was obtained for the hypothesis that the clustering of thrombosis in families is due to epistatic effects.2 Studies in selected families with venous thrombosis indicated that the presence of mutations in two genes may increase the penetrance of the thrombotic disease.13-15 The discovery of genetic risk factors for thrombosis came after the identification of families in whom the thrombophilia segregated with an abnormal result in a plasma test (protein C, protein S, antithrombin, and APC resistance).10,11,16-21 However, despite the ever growing insight into the processes of coagulation and fibrinolysis, the underlying cause of many inherited thrombotic events remains unsolved. New technologies for genetic analysis of thrombophilic families offer the opportunity to use a direct genetic strategy for identification of other genetic defects involved in inheritable thrombophilia.12,22

We investigated the prothrombin gene as a candidate gene for venous thrombosis. Prothrombin is the precursor of the serine protease thrombin, a key enzyme in the processes of hemostasis and thrombosis, that exhibits procoagulant, anticoagulant, and antifibrinolytic activities.23-25 Prothrombin is encoded by a 21-kb-long gene26 localized on chromosome 11, position 11p11-q12.27 The prothrombin gene is organized in 14 exons, separated by 13 introns with the 5' upstream untranslated (UT) region and the 3'-UT region,26 which may play regulatory roles in gene expression.

The aim of the present study was to perform an analysis of the prothrombin genes of selected subjects with a history of venous thrombophilia using polymerase chain reaction (PCR) and direct sequencing of the coding regions and their flanking splice junctions and the 5'- and 3'-UT regions. One genetic variation in the 3'-UT region of the prothrombin gene, a G to A transition at nucleotide position 20210, was found in 18% of selected patients with a personal and family history of venous thrombosis; in 6.2% of unselected consecutive patients with a first, objectively confirmed episode of deep-vein thrombosis; and in 2.3% of healthy control subjects. Carriers of the 20210 A allele have higher plasma prothrombin levels than controls with the normal 20210 GG genotype and have a 2.8-fold increased risk of venous thrombosis.

MATERIALS AND METHODS

Subjects. In a previous study, we collected detailed information on the occurrence of venous thrombo-embolic events in the families of 113 probands with a personal and family history of venous thrombophilia.28 From these, we randomly selected 28 families using the following criteria: (1) apart from the proband, there should be at least two symptomatic (preferentially first degree) relatives; and (2) in probands and symptomatic relatives, deficiencies of protein C, protein S, antithrombin, or plasminogen or dysfibrinogenemia were excluded. On average, each proband had 2.4 (range, 1 to 6) symptomatic first degree relatives and 1.8 (range, 0 to 5) symptomatic

From www.bloodjournal.org by guest on October 31, 2017. For personal use only.
second degree relatives. All probands and family members gave
their informed consent for the study of unexplained familial thrombophilia. It was recently established by DNA analysis that, in this
panel of probands, the frequency of the factor V Leiden mutation,\(^\text{12}\) which is associated with a poor anticoagulant response to APC,\(^\text{10}\) is 40%.

The second group of patients came from a population-based case-
control study on venous thrombosis, the Leiden Thrombophilia Study (LETS).\(^\text{11}\) Briefly, patients were selected from the computer files of the Anticoagulation Clinics in Leiden, Amsterdam, and Rot-
tterdam. In the Netherlands, Anticoagulation Clinics monitor couma-
рин treatment in virtually all patients with venous thrombosis in a
defined geographic area.\(^\text{26,28}\) Included are 474 unselected and consec-
uative outliers younger than 70 years of age who were referred
for anticoagulant treatment because of a first, objectively diagnosed
episod of deep-vein thrombosis. The median time between the oc-
currence of the deep-vein thrombosis and blood collection was 19
months (range, 6 to 68 months). Ninety-one percent of the eligible
patients were willing to take part in the study. The thrombotic pa-
tients were asked to find their own healthy control subject according
to predefined criteria.\(^\text{11}\) This resulted in 474 population control sub-
jects matched for age and sex. The mean age for patients and controls
was 47 years (range, 16 to 70 years for patients; range, 16 to 73
years for controls) and the male/female ratio among patients and
controls alike was 3/4.\(^\text{11}\)

Blood collection and laboratory analysis. Blood was collected
in tubes containing 0.106 mmol/L trisodium citrate. Plasmas
were prepared by centrifugation for 10 minutes at 2,000g at room tempera-
ture and stored at -70°C in 1.5-mL aliquots. High molecular weight
DNA was extracted from the white blood cell fraction using standard
methods.

Prothrombin activity was measured with a chromogenic method
using S-2238 as substrate and Echis carinatus venom as activator.\(^\text{21}\) Prothrombin antigen was determined using a Laurell electroimmu-
noassay.\(^\text{31}\) Protein C activity was measured with Coame protein C
(Chromogenix, Möln达尔, Sweden). An amidolytic heparin cofactor
assay (Chromogenix) was used for antithrombin activity measure-
ments. Total protein S antigen was determined by polyclonal en-
zyme-linked immunosorbert assay (ELISA).\(^\text{32}\) The results are ex-
pressed in units per milliliter, in which 1 U refers to the activity or
antigen present in 1 mL of pooled normal plasma.

For the identification of a genetic abnormality (or abnormalities)
in the prothrombin gene in DNA from 28 probands with a family
history of deep venous thrombosis, we used the PCR followed by
direct sequencing.\(^\text{23,25}\) We compared these sequences with those of 5
healthy control individuals. Genomic DNA was specifically ampli-
fied for the 14 exons with their flanking regions and for the
5'- and 3'-UT regions of the prothrombin gene using PCR.\(^\text{24}\) The primers
used in the PCR were derived from the sequence of the gene\(^\text{26}\)
and are identical to those used in a previous study.\(^\text{23}\) The fragments
obtained by PCR were purified on 1% ultralow melting temperature
agarose gel. The segment of the gel containing the amplified frag-
ment was excised and sequenced with the appropriate primers using
the dideoxynucleotide chain termination method.\(^\text{27}\) Sequencing reac-
tions were electrophoresed on 40-cm-long 8% polyacrylamide gels.
The gels were dried on Whatman 3 mm paper (Whatman, Maidstone,
UK) and exposed to an x-ray film. Genetic abnormalities identified
by sequencing were confirmed by restriction enzyme digestion of
amplified gene fragments. When the abnormality did not create or
abolish a restriction site, such a site was created by introducing a
nucleotide substitution with a mutant oligonucleotide during ampli-
fication.\(^\text{36}\) The mutant oligonucleotide was designed with a nucleo-
tide substitution close to the 3' end, such that the combination of the
nucleotide substitution and the genetic abnormality created a
new restriction enzyme cleavage site. Sequence variations in the
prothrombin gene known as neutral polymorphic sites were identified
on the basis of previous published data,\(^\text{26,37}\) but are beyond the scope
of this study.

Genetic analysis of the FV Leiden mutation (1691 G → A) was
performed as previously described.\(^\text{12}\)

Statistical analysis. Odds ratios (ORs) were calculated as a mea-
sure of relative risk in the standard unmatched fashion. A 95%
confidence interval (CI) was constructed according to Woolf.\(^\text{38}\) Gen-
erally, the OR estimates the risk of thrombosis when a risk factor
is present relative to the reference category.

For risk factor analysis concerning plasma prothrombin values,
48 patients using oral anticoagulant therapy were excluded from the
LETS group. To assess a dose-response relation, we stratified
the prothrombin values of both patients and controls into quartiles and
and calculated the ORs for the three higher levels relative to the lowest
reference level. Adjustment for current oral contraceptive use (yes/
no), body mass index (in kilograms per square meter), menopause
(yes/no), smoking (yes/no), age, and sex was performed by uncondi-
tional logistic regression. Effect modification was assessed by strat-
ified analysis and logistic regression with interaction terms.

Materials. Deoxynucleotides, dideoxynucleotides, and bovine
serum albumin were purchased from Pharmacia (Uppsala, Sweden).
(α-S)-dATP (>1,000 Ci/mmol) was obtained from Amersham Inter-
national (Amersham, UK). Klenow DNA polymerase was from
Boehringer Mannheim (Mannheim, Germany). Taq-DNA polymer-
ae (Amplitala) was purchased from Perkin Elmer-Cetus (Norwalk,
CT). The chromogenic substrate S-2238 was obtained from Chromo-
ogenix (Mölündal, Sweden). The Echis carinatus venom was obtained
from Sigma (Sigma Chemical, St Louis). Restriction enzymes were
obtained from New England Biolabs (Beverly, MA). Oligonucleo-
tides were synthesized on a Cyclone DNA synthesizer (Millipore,
Bedford, CT). All other chemicals were of analytical grade from
Merck (Darmstadt, Germany).

RESULTS

Our strategy for the identification of sequence variations
in the prothrombin gene was to amplify and sequence the
exons and their splice junctions and the 5'- and 3'-UT re-
gions of the gene. These regions contain the most likely sites
for mutations or polymorphisms that would affect transcrip-
tion or translation or the stability of the translated product.
The PCR products amplified from genomic DNA of the 28
probands and 5 healthy controls were sequenced as reported
previously.\(^\text{23}\) Except for sequence variations known as neu-
tral polymorphisms,\(^\text{26,37}\) no nucleotide change was found in
the 14 exons and the 5'-UT region of the prothrombin gene.
Only one heterozygous nucleotide transition (G to A) at
position 20210, the last nucleotide of the 3'-UT region,\(^\text{26}\)
was found in DNA of 5 of the 28 probands (18%), but not in
DNA of the 5 healthy control individuals (Fig 1). The
presence of this sequence variation was confirmed with re-
striction enzyme analysis using one mutagenic primer (Fig 2).
In an extended analysis of 100 healthy subjects, the heto-
zygous state (20210 AG) was detected in 1%. Homozy-
gous (20210 AA) carriers were absent.

Figure 3 shows the pedigree of the family of one of the
five thrombophilic patients carrying the variant prothrombin
allele. Both the parents of the proband are heterozygous
for the 20210 A allele, whereas her sister is homozygous.
Interestingly, both the variant prothrombin allele and the
factor V Leiden allele segregate in this pedigree. All individ-
uals that experienced a thrombotic event (or events) (II, 1;
The high frequency (18%) of the 20210 A allele among patients with thrombophilia is probably affected by selection. To study the relevance of the 20210 A allele in the population, we undertook the analysis of a population-based patient-control study (LETS). The prevalence of carriers of the 20210 A allele among healthy control subjects in the LETS was 2.3%, which corresponds to an allele frequency of 1.2% (95% CI, 0.5% to 1.8%). Table 1 shows a higher prevalence of the 20210 AG genotype among patients (6.2%) than among control subjects (2.3%). Homozygous AA carriers were not found (expected prevalence, 0.014%). The relative risk for thrombosis associated with the 20210 A allele was 2.8 (95% CI, 1.4 to 5.6). This association persisted when controlling for age, sex, current pill use, body mass index, menopause, and smoking. The 20210 A allele was associated with an increased risk for thrombosis both in men and women. We also found that the 20210 A allele increased the risk for all age groups.

The increased risk (2.8) associated with the 20210 A allele was not the result of overrepresentation of other risk factors, such as APC resistance (factor V Leiden); a deficiency of protein C, protein S, or antithrombin; or the presence of lupus anticoagulants. After excluding all these subjects (n = 141), we found an unmatched OR for thrombosis of 2.7 (95% CI, 1.3 to 5.6).

Individuals with the normal 20210 GG genotype had a mean prothrombin level of 1.05 U/mL (n = 860; SD, 0.15; range, 0.55 to 1.56), whereas individuals with the 20210 AG genotype had a significantly higher mean prothrombin level of 1.32 U/mL (n = 40; SD, 0.18; range, 0.95 to 1.78; P < .001). There was no notable difference in prothrombin levels between patients and control subjects within each of the two genotypic groups. The mean levels (in units per milliliter) of protein C, total protein S, and antithrombin did not differ between the 20210 GG and 20210 AG genotype (1.03, 1.04, and 0.99 vs 1.02, 1.03, and 0.98, respectively).

To assess to what extent an increased prothrombin level in itself is a risk factor for venous thrombosis, we stratified the prothrombin levels of patients and control subjects into quartiles (Table 2). The OR increased with increasing prothrombin levels: subjects with a prothrombin level of greater than 1.15 U/mL had a 2.1-fold higher risk than those in the reference category (<0.95 U/mL). The high-risk stratum of greater than 1.15 U prothrombin/mL comprised no less than 31% of the patients and 20% of the control subjects.

Table 3 shows the distribution of the 20210 genotypes...
Fig 3. Pedigree of a family in whom both the 20210 G/A sequence variation in the prothrombin gene and the FV Leiden mutation (1691 G/A) are segregating. Thrombotic symptoms are indicated by a dotted upper left quartile of the symbols; heterozygosity for the FV Leiden mutation is indicated by a hatched upper right quartile of the symbols; the presence of the 20210 AG genotype in the prothrombin gene is indicated by a solid lower right quartile of the symbols; and the presence of the 20210 AA genotype in the prothrombin gene is indicated by a solid lower left and right quartile of the symbols. Individuals tested for the 20210 G/A sequence variation in the prothrombin gene and the FV Leiden mutation are indicated by a dot to the left of the symbol. The arrow denotes the proposita; individuals indicated by a slash through the symbol are deceased.

over the different categories of prothrombin activity. Both in patients and control subjects, around 87% of the individuals with the 20210 AG genotype were in the highest category of prothrombin activity (>1.15 U/mL), whereas less notable differences were observed in individuals with the normal 20210 GG genotype.

DISCUSSION

Our study shows that a novel sequence variation in the prothrombin gene (nt 20210 *A) is a moderate risk factor for venous thrombosis (OR, 2.8; 95% CI, 1.4 to 5.6). The further observations that the 20210 A allele is associated with elevated prothrombin levels, that carriers of this allele have significantly higher prothrombin levels than noncarriers, and that elevated plasma prothrombin itself is also a risk factor for thrombosis suggest that the 20210 A allele acts through the elevated prothrombin levels.

In the LETS, the 20210 A allele was found in 6.3% of consecutive unselected patients with a first episode of deep vein thrombosis, indicating that the 20210 A allele is a relatively common risk factor for venous thrombosis. As expected, a much higher prevalence of 20210 A carriers was found in a group of selected patients with familial venous thrombosis (18%). In 60% of the 20210 A carriers, the 20210 A allele was the only genetic abnormality found, whereas in 40%, the FV Leiden mutation (R506Q) was also present. The prevalence of carriers of the 20210 A allele among

Table 1. Frequencies and Thrombotic Risk for the 20210 G/A Genotypes in the Prothrombin Gene

<table>
<thead>
<tr>
<th>Genotype (nt 20210)</th>
<th>No. of Patients (%)</th>
<th>No. of Controls (%)</th>
<th>OR_{uni}*</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GG</td>
<td>442 (93.8)</td>
<td>463 (97.7)</td>
<td>1.0†</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>29 (6.2)</td>
<td>11 (2.3)</td>
<td>2.8</td>
<td>1.4-5.6</td>
</tr>
<tr>
<td>AA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Adjustment for age and sex, current pill use (yes/no), body mass index, in menopause (yes/no) and smoking (yes/no) did not affect these results.
† Reference category.

Table 2. Thrombosis Risk for Plasma Prothrombin Levels

<table>
<thead>
<tr>
<th>Prothrombin Activity (U/mL)</th>
<th>No. of Patients (n = 426) (%)</th>
<th>No. of Controls (n = 474) (%)</th>
<th>Total No. (n = 900) (%)</th>
<th>OR†</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.95</td>
<td>85 (20)</td>
<td>134 (28)</td>
<td>219 (24)</td>
<td>1.0†</td>
<td></td>
</tr>
<tr>
<td>0.95-1.04</td>
<td>107 (25)</td>
<td>125 (26)</td>
<td>232 (26)</td>
<td>1.3</td>
<td>0.9-2.0</td>
</tr>
<tr>
<td>1.05-1.15</td>
<td>102 (24)</td>
<td>118 (25)</td>
<td>220 (24)</td>
<td>1.4</td>
<td>0.9-2.0</td>
</tr>
<tr>
<td>&gt;1.15</td>
<td>132 (31)</td>
<td>97 (20)</td>
<td>229 (25)</td>
<td>2.1</td>
<td>1.5-3.1</td>
</tr>
</tbody>
</table>

*Patients on oral anticoagulant treatment are excluded (n = 48).
† Test for trend, P < .001.
‡ Reference category.
controls was about 2.3%, corresponding to an allele frequency of 1.2% (95% CI, 0.5% to 1.8%). This is about eightfold higher than for protein C deficiency (0.3%) but about twofold less frequent than the so far most common genetic risk factor for venous thrombosis, the factor V Leiden mutation, which is associated with APC resistance (3% to 5%).

The 20210 A allele was not only found to be a risk factor for thrombosis but also to be associated with elevated prothrombin levels. Interestingly, elevated prothrombin levels were also a risk factor for thrombosis (Table 2). Thus, the prothrombin level may be considered as an effector, suggesting also that other factors than the 20210 A allele can be responsible for high prothrombin levels. How elevated prothrombin levels may stimulate the formation of venous thrombi is still unclear. They may lead to an imbalance between the procoagulant, anticoagulant, and fibrinolytic system. For instance, when higher concentrations of prothrombin would lead to increased rates of thrombin generation, this might result in excessive growth of fibrin clots.

This study does not show the mechanism(s) by which the 20210 A allele of the prothrombin gene may contribute to higher prothrombin levels. The association found for these two variables (Table 3) and the location of 20210 G to A transition in the 3′-UTR region of the prothrombin gene may indicate a relatively higher translation efficiency or higher stability of the transcribed mRNA. The G/A sequence variation is located at the last position of the 3′-UTR at or near the cleavage site in the mRNA precursor to which poly A is added. Three conserved sequences in mRNA precursors, located in the vicinity of this site, are required for cleavage and polyadenylation: the AUA AAA sequence, the nucleotide to which poly A is added, and the region downstream of this nucleotide. Generally, the nucleotide to which poly A is added is an A, mostly preceded by a C. As a consequence of the G to A transition at position 20210, a CA dinucleotide (instead of GA) has been introduced at or near the cleavage and polyadenylation site. However, in vitro experiments so far do not support a hypothesis in which this nucleotide substitution will result in an increased efficiency of the 3′ end formation. Alternatively, it cannot be excluded that the 20210 A allele is in linkage disequilibrium with another sequence variation (that escaped our analysis) that is responsible for the elevated prothrombin levels.

Finally, our approach of sequencing a candidate gene for thrombosis in a panel of probands from families with documented thrombophilia, followed by estimating the risk associated with any observed sequence variation in a population based patient-control study, proved to be useful. This approach seems suitable for unraveling more unknown genetic defects in other candidate genes for inherited thrombophilia.

ACKNOWLEDGMENT
We thank T. Visser and H. de Ronde for skillful technical assistance; A. van Beek, W. Noteboom, and Y. Bauman-Souverein for secretarial and administrative support; and T. Koster and R. Lenssen for collecting blood samples of patients (and their families) and control subjects. We also express our thanks to all patients and control subjects who participated in the Leiden Thrombophilia Study.

REFERENCES
PROTHROMBIN GENE VARIATION AND THROMBOSIS

44. Bimstiel M, Busslinger M, Strub K: Transcription termination and 3' processing: The end is in site! Cell 41:349, 1985
A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis

SR Poort, FR Rosendaal, PH Reitsma and RM Bertina