UNIVERSIDAD DE MURCIA

FACULTAD DE MEDICINA

Identification of new elements involved in antithrombin regulation.

Identificación de nuevos elementos implicados en la regulación de antitrombina.

Dª. Mª Eugenia de la Morena Barrio

2013

UNIVERSIDAD
DE MURCIA
DEPARTAMENTO
DE MEDICINA INTERNA
Facultad de Medicina

D. Javier Corral de la Calle y **D. Vicente Vicente García**, Profesores Doctores de la Universidad de Murcia del Área de Hematología en el Departamento de Medicina Interna de la Universidad de Mucia,

CERTIFICAN:

Que la tesis Doctoral titulada "**Identificación de nuevos elementos implicados en la regulación de antitrombina**", realizada por Dª. **Mª Eugenia de la Morena Barrio** bajo su inmediata dirección y supervisión, y que se presenta para la obtención del grado de Doctor Internacional por la Universidad de Murcia.

En Murcia, a 8 de Enero de 2013

Fdo: Dr. Javier Corral de la Calle

Fdo: Dr. Vicente Vicente García

UNIVERSIDAD
DE MURCIA
DEPARTAMENTO
DE MEDICINA INTERNA
Facultad de Medicina

Los trabajos de investigación llevados a cabo en la presente memoria de Tesis Doctoral han sido financiados por los siguientes proyectos:

- "Investigación genómica, proteómica y terapéutica del sistema hemostático en la enfermedad tromboembólica venosa y arterial", proyecto SAF 2006-06212 financiado por el Ministerio de Educación y Ciencia.
- "Estudio genético y estructural de la antitrombina. Desarrollo de nuevas drogas anticoagulantes", proyecto otorgado por la Fundación Mutua Madrileña.
- "Anomalías de glicosilación en antitrombina. Implicaciones patológicas y terapéuticas", proyecto PI12/00657 financiado por el Instituto de Salud Carlos III.

El firmante de esta memoria ha disfrutado de una Beca Nacional Predoctoral de Formación e Investigación en la Salud (FI09/00190) del Ministerio de Ciencia e Innovación.

En Murcia, a 8 de Enero de 2013

Fdo: Mª Eugenia de la Morena Barrio

Agradecimientos

Sois muchos los que formáis parte de esta tesis y a los que os debo también mucho, por eso estas líneas me resultan especialmente difíciles aunque por otra parte son las que más ilusión me hacen. A todos os doy las gracias.

En primer lugar, me gustaría expresar mi más sincero agradecimiento al Dr. Vicente Vicente García, co-director de esta tesis y jefe del grupo, por darme la oportunidad de iniciarme en el mundo de la investigación abriéndome las puertas de su grupo. Es mi deber agradecerle la confianza que ha depositado en mí y sus constantes esfuerzos por ofrecerme una formación científica de calidad. Quiero agradecerle sus consejos, preocupaciones, disponibilidad, buen hacer, su calidad profesional y capacidad de trabajo y todo ello transmitiendo siempre serenidad, tacto, interés y sentido del humor.

En segundo lugar, con profundo agradecimiento me dirijo al Dr. Javier Corral, director y principal motor de este trabajo. Espero que acepte mi reconocimiento no solo a la dirección de esta tesis, sino también a la orientación y formación científica que le debo. Su capacidad de trabajo, exprimiendo cada segundo de cada minuto, su espíritu crítico, su entusiasmo, inquietud y motivación junto con su gran bondad humana, mostrando interés más allá de lo profesional, con sus constantes consejos y preocupaciones, son sus pilares fundamentales, por lo que cualquiera se sentiría afortunadísimo de tenerlo como director de tesis, jefe, colega o amigo. Solo espero haber asimilado algo de todo lo que me ha transmitido.

A Toñi Miñano le debo tanto… Gracias por esos ratos de intenso trabajo en el laboratorio donde he trabajado muy a gusto. Orden, trabajo, disciplina, organización, constancia, generosidad y un gran corazón. Siempre con una sonrisa en los labios y con palabras acertadas de ánimo. Le agradezco su ejemplo, su amistad y sus constantes desvelos casi maternales.

A las Dras. Anabel Antón e Irene Martínez quisiera agradecerles de forma especial lo mucho que he recibido de ellas. A Anabel le debo la formación genómica de mis primeros años. Le agradezco su afectuoso acogimiento y cariño durante esos años, además de sus grandes dotes como pedagoga. A Irene mi formación más proteómica. Entre tantas cosas siempre me ha transmitido orden y rigurosidad en el trabajo y el no conformarme con la mediocridad. Les agradezco a las dos sus siempre buenos consejos, sugerencias y cariño.

Al Dr. Constantino Martínez me gustaría agradecerle toda la formación en biología molecular que me ha aportado y que ha sido fundamental en el desarrollo de algunos apartados de la tesis. Además, le agradezco su gran disponibilidad y su constante alegría creando un ambiente en el grupo muy agradable.

Al Dr. Pepe Rivera su espíritu crítico, su interés, atención y disponibilidad. Agradecerle los conocimientos sobre hemostasia primaria (la plaqueta) que siempre nos transmite.

A la Dra. Rocío González-Conejero su orden y trabajo acompañando siempre una sonrisa y una gran serenidad.

A los Dres. Paqui Ferrer, Marisa Lozano, Mª José Candela, Francisco Ayala y Gloria Soler por sus siempre acertadas sugerencias y espíritu crítico. A Marisa agradecerle especialmente su valiosa ayuda con el citómetro y la formación en hematología que me ha aportado.

A los Dres. José Anotonio Guerrero, Leyre, Adriana, Pepe Navarro, Virginia y Raúl les agradezco su ejemplo y trayectoria. Han ido abriéndome camino y me han ayudado siempre que lo he necesitado. Han sido mis puntos de referencia que he intentado seguir hasta estos días. A José G me gustaría agradecerle además su tiempo dedicado a la revisión del inglés del manuscrito, sus correcciones, observaciones y sugerencias.

A Isabel y Sonia les agradezco los muchos y muy buenos momentos que hemos compartido como becarias “precarias”. El día a día ha sido mucho más llevadero a su lado, compartiendo y aprendiendo juntas.

A las nuevas incorporaciones Ginés, Salam, Eva, Ana y Sergio por ir renovando y rejuveneciendo al grupo aportando nuevas ideas, visiones y modos de hacer.

A todo el maravilloso equipo técnico. A Lorena por todo lo que me ayudó los primeros años, su cariño y buenos momentos que hemos compartido. A Jose Padilla le agradezco todo su gran trabajo de secuenciación, su orden y organización. A Nuria le agradezco la formación en biología celular y molecular. Siempre ha sido una gran ayuda con los cultivos, las purificaciones de material genético, etc. A Natalia todo lo que me ha aportado y ayudado en el laboratorio. A “su Rodri” le agradezcosu ayuda con la impresión de la tesis. Y a Tina y Noelia sus favores y atenciones con las técnicas genómicas.

Me gustaría mostrar también mi gratitud al resto de miembros del servicio de Hematología y Oncología del Hospital Morales Meseguer y del Hospital Reina Sofía, por su gran preparación y masa crítica que generan en las sesiones.

También agradezco a los enfermeros del Centro Regional de Hemodonación su disponibilidad en la extracción de muestras de sangre.

En definitiva, mi agradecimiento a todos y cada uno de los que formáis parte de este grupo por todo lo que transmite: profesionalidad, calidad humana, espíritu crítico, disciplina, orden, saber hacer, exigencia, flexibilidad, rigurosidad, convivencia y por supuesto, su gran potencial culinario. Su excelencia profesional combinada con su excelencia humana han diluido los momentos duros, difíciles o de intenso trabajo, en momentos amenos donde he disfrutado muchísimo.

Además, es mi deber agradecer al personal técnico y sanitario de los hospitales y grupos con los que hemos colaborado, su interés y esfuerzo en el envío de muestras biológicas de pacientes. Concretamente dar las gracias al Dr. Francisco España (Hospital La Fe, Valencia), a la Dra. Carmen de Cos (Hospital Universitario Puerta del Mar, Cádiz), al Dr. Jose Manuel Soria (Institut d'Investigació Sant Pau, Barcelona), al Dr. Søren (Department of Haematology, Aalborg Hospital, Dinamarca), a la Dra. Teresa Sevivas (Centro Hospitalar de Coimbra, Portugal), al Dr. Jaeken (Universitair Ziekenhuis Gasthuisberg, Leuven, Belgium.) y al Dr. Dirk Lefeber (Institute for genetic and metabolic disease, Holanda), así como a todos los pacientes que han participado en este estudio. A todos ellos mi más profundo respeto y reconocimiento.

A los profesores Jaak Jaeken, Kathleen Freson, Marc Hoylaerts y Michela de Michele agradecerles además los meses de estancia en su departamento (Center for Molecular and Vascular Biology, KU Leuven, Leuven, Bélgica) y la interesante colaboración que hemos abierto desde entonces. Especialmente agradecer a la Dra. Michela y al Dr. Marc su cariño y esfuerzos por hacer productiva mi estancia allí.

A Vanessa Ferreiro, presidenta de la Asociación de enfermos con trastornos de glicosilación congénita, mostrarle mi admiración por la labor que desempeña y mi gratitud por hacer posible la colaboración con el grupo del profesor Jaeken.

Al Dr. Francisco España y a la Dra. Silvia Navarro quisiera agradecerles también los días que pasé en su departamento en el Hospital La Fe de Valencia. Les agradezco su ayuda con la técnica de generación de trombina.

A los profesores Marc Hoylaerts y Søren Risom Kristensen me gustaría expresarles mi reconocimiento por su tiempo dedicado en la valoración de la tesis para aspirar a la mención de Doctorado Internacional.

Y lo más importante: mi familia y amigos. Creo que no hace falta que me extienda mucho aquí. Bien saben ellos lo mucho que agradezco lo que han sido y son para mí, estos años y siempre.

Por último, debo agradecer al Instituto de Salud Carlos III la beca que he disfrutado estos cuatro años y que me ha hecho posible llegar hasta aquí.

A mi Abuelo (in memoriam)

A mis padres

Abbreviations

2D	Two dimension electrophoresis
α1AT	Alpha 1 antitrypsin
AB	Aminobenzamide
ACN	Acetonitrile
AMI	Acute Myocardial Infarction
AT	Antithrombin
bp	Nucleotide base pairs
CDG	Congenital Disorder of Glycosylation
CDT	Carbohydrate-Deficient Transferrin
CS-DS	Chondroitin-dermatan sulfate
kDa	Kilo Dalton
DHB	Dihydroxybenzoic acid
DVT	Deep Venous Thrombosis
ELISA	Enzyme Linked Immunosorbent Assay
FXa	Activated Blood Coagulation Factor X
FPLC	Fast Protein Liquid Chromatography
GAGs	Glycosaminoglycans
GAIT	Genetic Analysis of Idiopathic Thrombophilia
GPIbα	Glycoprotein Ib alpha
GPIIIa	Gycoprotein IIIa
GU	Glucose Units
GWAS	Genome Wide Association Study
HEK-EBNA	Human Embryonic Kidney cells expressing the Epstein Barr Nuclear Antigen 1
HepG2	Human hepatocellular liver carcinoma cell line
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High Performance Liquid Chromatography
HS	Heparan Sulfate
INR	International Normalized Ratio
LLO	Lipid-linked oligosaccharide
MALDI-TOF	Matrix Assited Laser Desorption Ionization-Time-of-flight

MPI	Mannose-6-phosphate isomerase
MS	Mass Spectrometry
PAI-1	Plasminogen activator inhibitor 1
PBS	Phosphate Buffer Saline
PCR	Polymerase chain reaction
pI	Isoelectric point
PLC-PRF-5	Human hepatome cell line
PMM2	Phosphomannomutase 2
PVDF	Polyvinylidene Fluoride
qRT-PCR	quantitative Real Time PCR
Q-TOF	Quadrupole-Time-of-Flight
ROS	Reactive oxygen species
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel
SERPIN	Serine protease inhibitor
siRNA	small interfering RNA
SNP	Single Nucleotide Polymorphism
TFA	Trifluoroacetic acid
vWF	von Willebrand Factor

General index

Page

OBJECTIVE 3

Index of tables and figures

Page

INTRODUCTION

MATERIAL AND METHODS

DISCUSSION

ABSTRACT

Antithrombin is a key anticoagulant serpin. Any factor affecting, even moderately, the levels or function of this molecule will unbalance a finely regulated haemostatic system, increasing the risk of thrombotic disorders, as widely demonstrated with congenital antithrombin deficiency. Mutations in the coding regions of the gene encoding antithrombin, *SERPINC1*, have been identified in a high proportion of patients with antithrombin deficiency. However, there are few studies evaluating other elements playing a role in such a relevant molecule. In this thesis, we have addressed this issue through the study of an interesting cohort of 29 patients with moderate or severe deficiency of antithrombin without mutation in the coding region of the *SERPINC1* gene. The search also involved the analysis of 307 healthy controls, a family study (GAIT) and large case control studies involving 2,980 patients and 3,996 controls from different countries and diverse thrombotic diseases, as well as patients with congenital disorder of glycosylation, which is a rare disorder. To achieve our objectives we have used extensive molecular methodology ranging, from simple genotyping of point mutations or polymorphisms to complex genome wide association studies, as well as sequencing of two genes (*SERPINC1* and *PMM2*) and quantitative RT-PCR. We also used cellular biology methods, such as reporter systems, silencing assays, and immunofluorescence staining. Finally, biochemical assays were also used, from western blot analysis of multiple proteins (antithrombin, α1-antitrypsin, transferrin, prothrombin and different platelet glycoproteins), to proteomic and glycomic studies. We also used HPLC and FPLC methodology.

Our study has identified the first mutation affecting the promoter of *SERPINC1* gene that is associated with antithrombin deficiency. This result has allowed us not only to identify new regulatory regions of this gene, but also sustains the analysis of this region in molecular studies of patients with severe or moderate antithrombin deficiency. We also confirmed the low genetic variability of *SERPINC1*, particularly in the coding region, probably explained by the conformational and functional sensitivity of this serpin. Moreover, we found a minor role of *SERPINC1* genetic variability on the levels of this anticoagulant in healthy population. This result, together with the high heritability of this trait, sustains the existence of other genetic factors modulating antithrombin levels. Actually, our GWAS and validation studies together with further functional results have identified the first modulating gene of antithrombin. Impaired *LARGE* expression caused reduction of antithrombin, potentially by the intracellular role of the heparan-like molecule generated by the xylosyltransferase and

glucuronyltransferase activities of this molecule, although the mechanism is still unknown. Our study has also identified global glycosylation abnormalities interfering with antithrombin levels. Thus, we identified the first PMM2-CDG patient diagnosed on his severe thrombotic history and deficiency of antithrombin. We also identified a new disorder, called CDG-like, which shares the same biochemical pattern as PMM2-CDG but only one clinical phenotype (thrombosis) and one out of the two required *PMM2* mutation. Interestingly, these CDG-like disorders seems to have a relatively high prevalence in patients with thrombosis. These results also suggest disorders of glycosylation as strong thrombophilic disorders and encourage the analysis of this disorder among patients with thrombosis and antithrombin deficiency but no mutations in the *SERPINC1* gene.

Our study also opens new attractive questions that must be answered: additional key regulatory regions of the *SERPINC1* gene must exist and they should be identified; other antithrombin modulating genes must be explored and genetic modifications affecting these genes should be studied in a context of a potential role in thrombosis; the real incidence of CDG-like disorders on thrombotic diseases must be clarified; and the mechanism involved in the abnormal haemostatic system and risk of thrombosis of PMM2-CDG and CDG-like patients should be characterized.

INTRODUCTION

1. Haemostatic system

During thousands of years of evolution, higher organisms with a closed and highly pressured circulatory system have developed elements and mechanisms able to produce a quick, simple and immediate response to the endothelial damage to minimize blood loss. Most of these processes will not require protein synthesis to avoid the otherwise fatal consequences of bleeding. The result is a finely regulated haemostatic system that includes a network of interactions of vasoconstriction, platelet aggregation and coagulation reactions. Finally, fibrinolysis slowly dissolves clots to maintain organ perfusion. Physiological inhibitors in plasma maintain the balance between coagulation, anticoagulation and fibrinolysis by localizing coagulation activity at the site of injury and preventing systemic coagulation. In this introduction, we will briefly describe only procoagulant reactions and their control, focusing on the main objective of our project: antithrombin.

Exposure of tissue factor after vascular damage triggers serial proteolytic reactions in cascade in which a zymogen (inactive enzyme precursor presented in plasma at a very high concentration) of a serine protease and its co-factor are activated. These active proteases then catalyze the next reaction in the cascade, ultimately resulting in thrombin formation, the last procoagulant serine protease of the coagulation cascade. Thrombin generates a clot of fibrin fibres by proteolyzing huge amounts of plasma fibrinogen. Thrombin also activates FXIII, which by its transglutaminase activity cross-links fibrin fibres, strengthening the clot. Finally, thrombin is also a potent platelet agonist. All these data sustain the key role of thrombin in pro-coagulant and pro-aggregant reactions. As shown in figure 1, other serine proteases are also crucial elements of the coagulation cascade

This strong procoagulant potential has to be precisely controlled in space and time: procoagulant reactions must be restricted to the injured areas where & when the damage is present. Otherwise, uncontrolled thrombin will lead to thrombosis by an obstructive clotting disorder. For this reason, the fibrinolytic system, which is also based on serial serine protease activation, slowly breaks down the clot formed (Figure 1). Finally, there are other key proteins involved in the control of haemostasis: a myriad of inhibitors of haemostatic proteases. According to their final function, these proteins can be classified in a) inhibitors of fibrinolytic factors and b) natural anticoagulants [1].

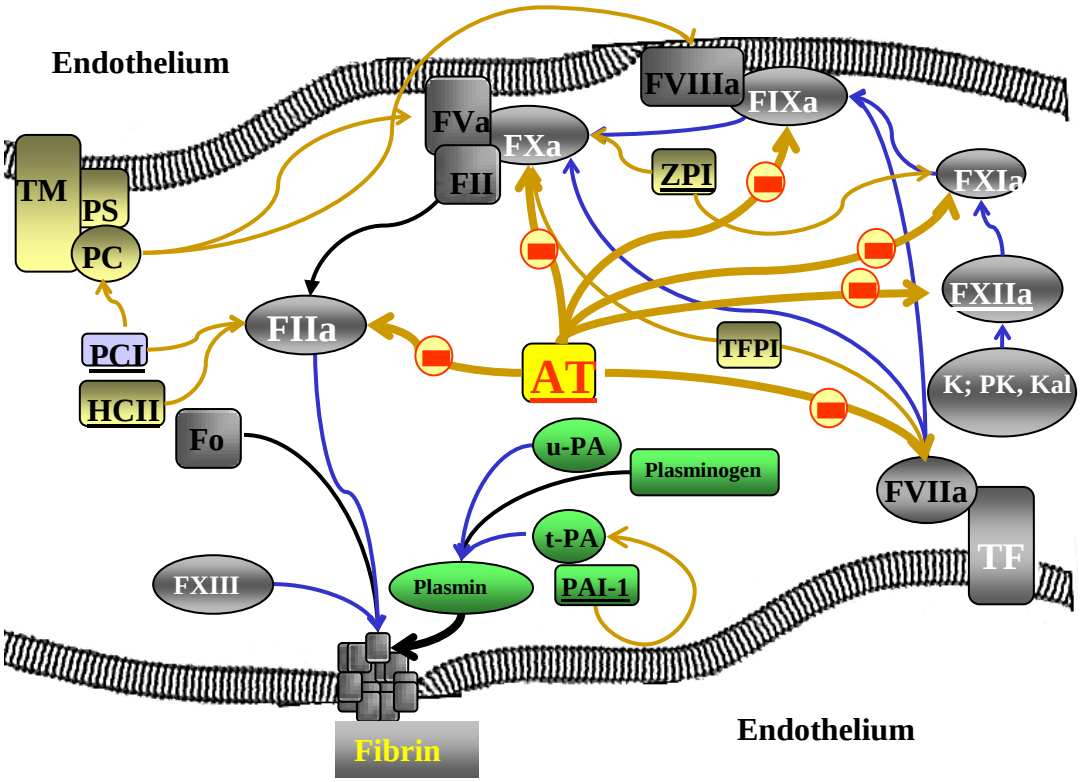


Figure 1. Clot cascade (grey), fibrinolysis (green) and anticoagulant molecules (yellow). Haemostatic serpins are underlined. TM: thrombomodulin; PS: protein S; PC: protein C; PCI: protein C inhibitor; HCII: heparin cofactor II; Fo: fibrinogen; ZPI: protein Z-dependent protease inhibitor; TFPI: Tissue factor pathway inhibitor; TF: tissue factor; K: kininogen, PK: prekalikrein; uPA: urokinase-type Plasminogen Activator; tPA: tissue plasminogen activator; PAI-I: plasminogen activator inhibitor; AT: antithrombin.

2. Thrombophilic factors

The haemostatic system necessitates a regulated equilibrium between strong procoagulant elements triggered by vascular damage and even stronger anticoagulant and fibrinolytic factors that maintain blood flowing. Accordingly, even minor disturbances of potentially any of the elements of the haemostatic system can have pathological consequences leading to an impair coagulation (hemorrhage), a hypercoagulation (thrombosis), or paradoxically to both of them (disseminated intravascular coagulation). The medical and socio-economical relevance of thrombotic disorders (the first cause of morbid-mortality in the world) have encouraged the search for mechanisms or factors increasing the risk to develop thromboembolic diseases.

The first description of a prothrombotic genetic defect arises in 1965, when the Norwegian haematologist O. Egeberg studied a family with a great incidence of venous thrombosis. Patients shared a common defect in the haemostatic system: deficiency of

antithrombin [2]. A rare mutation affecting the gene encoding antithrombin explained both the deficiency of this key endogenous anticoagulant and the high risk of thrombosis of carriers due to the role of this anticoagulant. In the 80's, the identification of other families with thrombosis and deficiency of other anticoagulant proteins (protein C or protein S), suggested that at least some cases of venous thrombosis might be considered as a monogenic disease. However, the penetrance of these defects is variable. Moreover, the incidence of all these defects collectively is low, even among thrombophilic patients (5-10%) (Table 1). These data, together with the fact that more than 60% of the variation in susceptibility to venous thrombosis is attributable to genetic factors [3], strongly suggested the presence of additional prothrombotic genetic risk factors.

The search for genetic factors involved in thrombosis changed from rare mutations associated with high risk to common polymorphisms following the attractive hypothesis of "common genetic variants are involved in common diseases". Certainly such interesting hypothesis was validated in patients that had suffered thrombosis by the group of R. Bertina, who in 1994 and 1996 identified two functional polymorphisms, the first impairing the inactivation of FV (FV Leiden Arg506Gln) and the second associated with moderately high levels of prothrombin (PT 20210G>A), which significantly increased the risk of thrombosis [4] (Table 1). These reports clearly demonstrated that venous thrombosis is a polygenic disorder and sustained the search for new prothrombotic polymorphisms, particularly since heritable risk factors for venous thrombosis can be identified only in 30–50% of affected patients [4]. This search was originally restricted to candidate genes, mainly from the haemostatic system, using case-control studies. The simplicity of such studies facilitated the sudden increase of thrombophilic studies (a pubmed's search using "polymorphisms and thrombosis" renders more than 2,000 entries). However, the final conclusion of these studies is quite frustrating since most of them found no significant association. Moreover, studies finding positive associations are usually restricted to some specific group of patients or populations, or have not been validated by other groups. Independently of this controversy, the risk of thrombosis associated with any polymorphism is always moderate (OR<2). A recent meta-analysis including 126,525 cases and 184,068 controls from 173 case-control studies that had evaluated 28 polymorphisms from 21 genes found that only few polymorphisms mildly associated with a thrombotic risk [5]. With the development of new genotyping and sequencing technologies, the design of experiments to identify novel prothrombotic polymorphisms changed, leading firstly to massive analysis of hundreds of polymorphisms located on

candidate genes, many of them not restricted to the haemostatic system, to genome wide association studies (GWAS) that genotyped up to 551,141 SNPs covering the whole genome in thousands of cases and controls [6,7]. Unfortunately, these expensive studies only confirmed classical associations and suggest that it is unlikely that new common polymorphisms could play a key role in venous thrombosis. In contrast, recent excellent studies identified new mutations with strong thrombophilic consequences affecting the levels or function of different haemostatic proteins [8,9]. New thrombophilic defects must be identified by a deep study of cases with idiopathic thrombosis.

Table 1. Frequent thrombophilic defects.

Genetic Defect	Haemostatic consequence	Prevalence Patients	Prevalence Controls	Relative Risk (OR)
Antithrombin deficiency	Reduced anticoagulant capacity (50%)	0.5-2%	0.02%	10-50
Protein C deficiency	Reduced anticoagulant capacity (50%)	3-5%	0.3%	10
Protein S deficiency	Reduced anticoagulant capacity (50%)	2%	0.03-0.13%	8-10
Factor V Leiden	Activated Protein C resistance	10-20%	4%	3-5
Prothrombin 20210G>A	High prothrombin levels	6-10%	3%	2.-3
ABO blood group No-O	High FVIII&WF levels?	71%	50%	2

3. Serpins

As stated before, anticoagulation is crucial for maintaining blood flowing. If procoagulant reactions are strong and efficient, anticoagulant mechanisms should be superior. The primary objective of these proteins is the control of proteases unspecifically generated in order to avoid spontaneous thrombosis. They should also avoid excessive or mislocalized thrombin formation under vascular damage. This strong anticoagulant activity is not too diverse. Although there are many proteins with capacity to inhibit procoagulant proteases, there are three main natural anticoagulants: the tissue factor pathway inhibitor (TFPI), the protein C, and particularly antithrombin. With minor relevance, other anticoagulants are the protein Z-dependent protease inhibitor (ZPI), the heparin cofactor II

(HCII), and the protein C inhibitor (PCI) (Figure 1). Other elements such as protein S, thrombomodulin, endothelial receptor of protein C or different glycosaminoglycans will also play a role in different anticoagulant pathways (protein C and antithrombin).

Interestingly, four of these anticoagulants (antithrombin, heparin cofactor II, the protein Z-dependent protease inhibitor and protein C inhibitor) belong to the superfamily of serpins. Actually, the features of these proteins make them ideal for the haemostatic system, as we will discuss soon. This also explains the presence of additional haemostatic serpins, this time controlling fibrinolytic reactions: α2-antiplasmin and plasminogen inhibitor 1 (PAI-1).

The superfamily of serpins (SERine Protease Inhibitors) spans more than 1,500 proteins [10] due to the so close molecular, structural and functional similarities among three proteins: human antithrombin, human α1-antitrypsin and chicken ovoalbumin [11]. The genes coding these proteins are divided into subgroups called clades (16 clades A-P) based on various aspects of similarity [10]. The systematic name of each serpin is, SERPINXy where X is the clade and y is the number within the clade [10]. Thus, the appropriate names for the haemostatic serpins are the following: Protein C Inhibitor (SERPINA5), Protein Z-dependent protease inhibitor (SERPINA10), heparin cofactor II (SERPIND1), antithrombin (SERPINC1), α2-antiplasmin (SERPINF2) and plasminogen inhibitor 1 or PAI-1 (SERPINE1).

Most of the serpins have serine protease inhibitor activity (although certain of them have caspase activity or do not inhibit enzymes [12]) and this function makes them central in the control of a big amount of proteolytic cascades inside or outside the cell. Consequently, they play important roles in different processes such as inflammation, apoptosis, complement, and as it can be expected coagulation and fibrinolysis [13-15]. Their sophisticate structure of serpins has played a central role in the understanding of their function and biology. Most of the serpins comprise 400-500 residues, a molecular weight of 37-70 kDa and an amino acid homology of 35% [16]. They are folded in a very conserved tertiary structure with 3 β-sheets (A-C) and 8 or 9 α-helixes (A-I) with a flexible and mobile region in the molecule pole responsible for the interaction with the target protease, which is known as Reactive Centre Loop (RCL) (Figure 2).

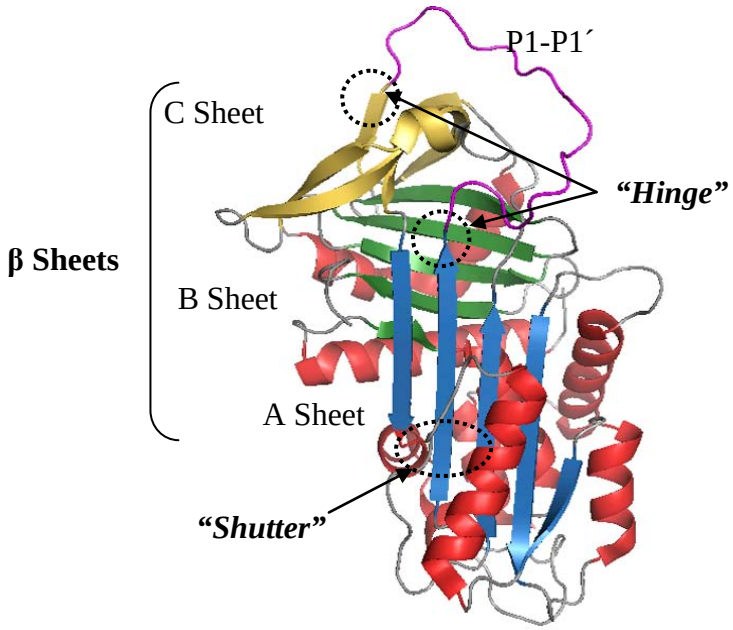


Figure 2. Native structure of α1-antitrypsin, the archetype structure of serpins (PDB: 1QLP). The Reactive Centre Loop (RCL) is drawn in magenta, sheets A in blue, B in green, C in yellow and α helix (A-I) in red. Dashed circles highlight the hinge position in the RCL and the shutter, both are areas very conserved.

Like other protease inhibitors, the RCL of serpins acts as a protease-specific bait. However, unlike other families of serine protease inhibitors, which function in a simple way by extending a substrate-like RCL that passively blocks the active centre of the target protease [17], the RCL cleavage of serpins by target proteases results in inhibition by an unique and extraordinary efficient suicide mechanism that has been compared with an mouse-trap (Figure 3). This mechanism is possible because serpins do not fold to their thermodynamically most favourable conformation. The serpin's native state is metastable, a kinetically trapped folding intermediate that upon RCL cleavage by proteases releases the energy for a significant conformational transition resulting in the most stable cleaved serpin [18]. The cleaved RCL moves to fill the middle strand of the main β-sheet of the molecule (sA), and in doing so carries with it the entrapped protease. The protease is flung a distance of 70 Å with an accompanying plucking and distortion of its structure that ensures its destruction [19]. Moreover, covalent protease-serpin complexes are rapidly cleared from circulation.

The conformational flexibility of serpins also makes them vulnerable to even minor genetic or environmental changes, which may affect a correct folding and/or a correct inhibitory function. These features, together with their relevance in important biological regulatory processes make serpins prone to mutation-related diseases [20]. Mutations that

alter even a single amino acid in critical regions of these proteins can disrupt their function. Abnormal serpins often form clumps (aggregates) that can build up to toxic levels within cells. These protein aggregates may also cause a shortage of the inhibitor in areas where it is needed to control chemical reactions. Disorders caused by aggregates of abnormal serpins are called serpinopathies [20,21]. Mutations in *SERPIN* genes have been identified as the cause of multiple disorders, including a lung disease called emphysema, hereditary angioedema, a type of familial dementia, and abnormal blood clotting (thrombosis) [10,22].

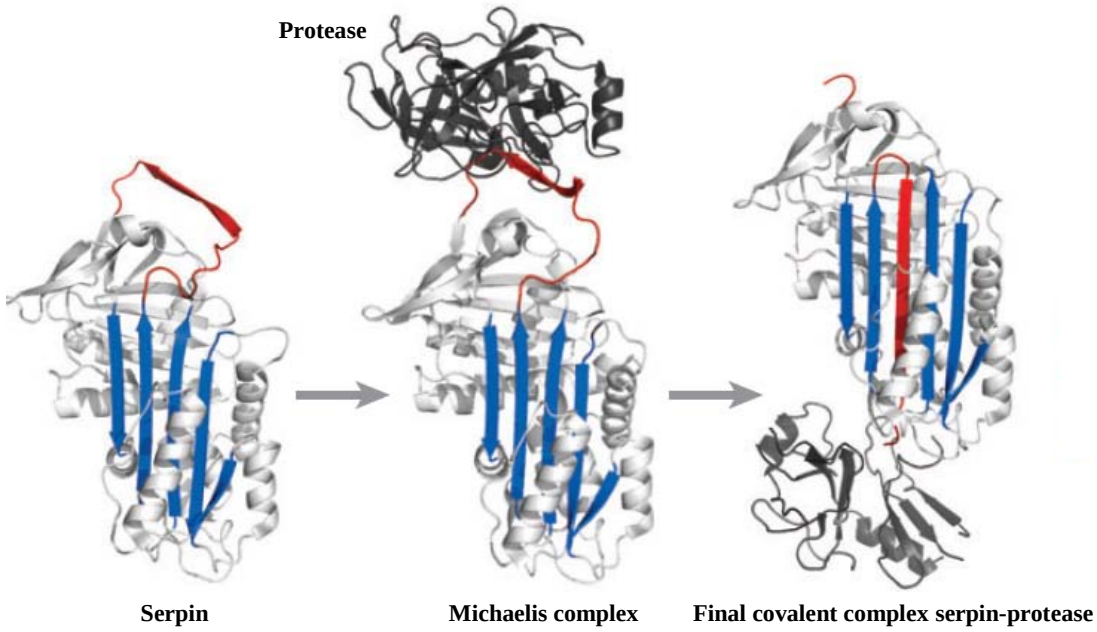


Figure 3. Inhibitory mechanism of Serpins.

4. Antithrombin

The first evidence of the existence of antithrombin can be found in 1982, when A. Schmidt suggests the presence of a plasma cytoglobulin that prevented coagulation. The name antithrombin was originally given by P. Moratowitz in 1905, while the first clinical assay to determine the plasma levels of antithrombin was developed in 1963 by A. Hensen y E.A. Loeliger. The first proof of the key role of antithrombin in haemostasis and thrombosis was the association with venous thrombosis of a heterozygous deficiency observed by O. Egeberg in 1965 [2].

Antithrombin is therefore an old well-characterised molecule. However, there are still many fascinating points concerning this molecule.

4.1 Structure and biological function

Antithrombin is synthesized in the liver and released in plasma reaching an average concentration of 150µg/ml. The molecule, with a plasma half-life of approximately 3 days, is degraded by liver cells [23].

Antithrombin contains 432 residues accounting for a molecular mass of 58 kDa. During its synthesis within the hepatocyte, three disulfide bridges are formed, a 32-residues signal peptide is released by unknown proteases, and up to four glycosylation sites are linked by the *N*-glycosylation machinery (Figure 4).

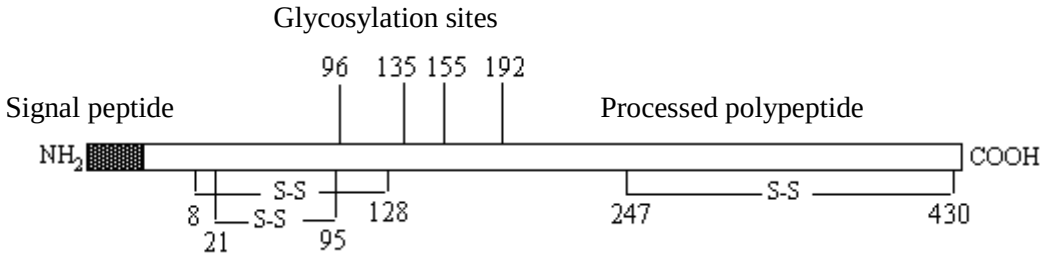


Figure 4. Localization of disulfide bridges and *N*-glycosylation sites in antithrombin.

The *N*-glycan present in antithrombin is biantenary and monosialylated, without core-fucosylation, which may reduce the heparin affinity [24]. The exact composition and linkages are shown in figure 5.

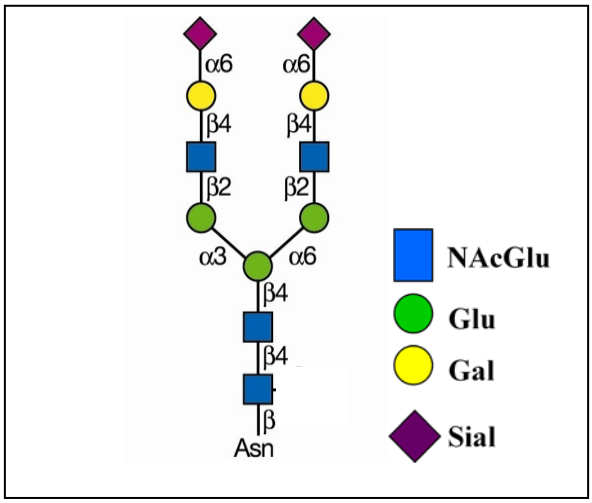


Figure 5. Composition of the *N*-glycans of antithrombin. NACGlu: N-Acetyl glucosamine; Glu: glucose; Gal: galactose; Sial: Sialic acid.

The presence of a serine instead of threonine at position 137 reduces the efficiency of *N*-glycosylation at the consensus sequence of Asn135 leading to the generation of two

main glycoforms of antithrombin in plasma [25]: α-antithrombin, with four *N*-glycans, represents up to 90% of total plasma antithrombin, and β-antithrombin lacking the *N*-glycan at Asn135 (Figure 6). The different site-occupancy affects its size, pI, and more structurally relevant its heparin affinity of the molecule, which is crucial for the activation of this serpin [26]. The absence of the *N*-glycan at Asn135 increases 6-fold the affinity for heparin [27]. Its higher heparin affinity sustains that despite of being the minority antithrombin glycoform in plasma (10%), β-antithrombin could be the most relevant thrombin regulator, particularly at the endothelial surface [28,29] and it could also explain its increased clearance rate, which contributes to the smaller amount of β-antithrombin in plasma [30]. Based on the clearance rate of the different antithrombin glycoforms, their distribution between blood, non-circulating vessel surface and subendothelial compartments, it has been proposed that α-antithrombin accounts for most of the anticoagulant function in flowing blood, and that β-antithrombin plays a relatively greater role at the vessel wall [31-35].

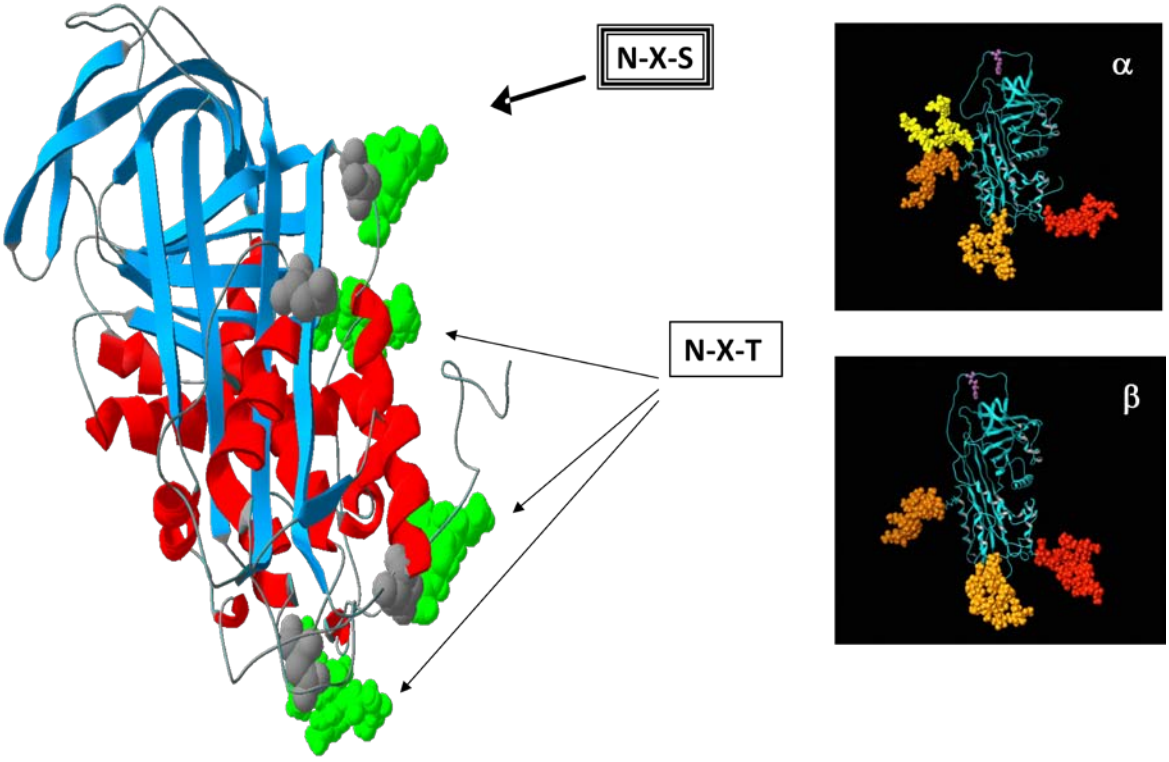


Figure 6. Main plasma glycoforms of antithrombin. The number and structural localization of *N*-glycans are shown.

This molecule shares the highly conserved structure of all serpins with three β-sheets (A–C) and nine α-helices (A–I) [36,37] and the efficient and strong mechanism of inhibition described for serpins, which for antithrombin is triggered after the proteolytic cleavage of the bound between Arg-393 and Ser-394 residues (P1-P1') by the target proteases [38]. However, in contrast to other serpins, the RCL of native antithrombin remains partially inserted into the top of the central A β-sheet (Figure 7) [39]. This native conformation together with the saline bridge between P1 residue (Arg393) and Glu237 [40] justify the relative inaccessibility of the target protease to the P1 residue and so the low anticoagulant activity of the native form of antithrombin [41]. Actually, the full anticoagulant activity of antithrombin requires the activation of this molecule triggered by heparin and other glycosaminoglycans (Figure 7). *In vivo*, physiologically relevant forms of heparin include heparan sulphate, found lining the endothelial layer and heparin released from the granules of mast cells associated with the endothelium.

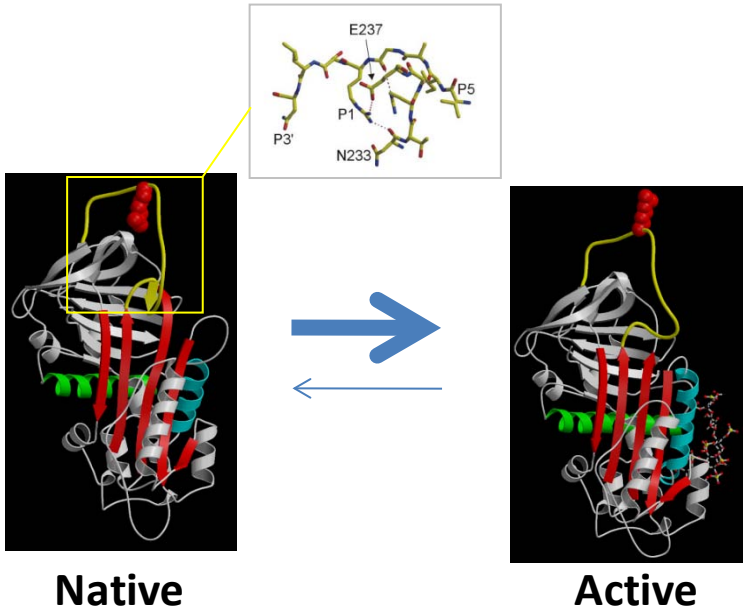


Figure 7. Native and active forms of antithrombin. The interactions of the P1 residue (Arg393) in the native conformation are also shown.

The binding of heparinoid molecules to antithrombin occurs at the heparin binding site of the serpin, which is centred on the D helix (Figure 2) [42]. An initial weak complex with a specific pentasaccharide sequence in heparin is formed [43]. This causes complex sequencial conformational changes in the molecule, where the D helix is extended closing the central A-β and releasing the RCL. The activation of antithrombin culminates with the

final exposure of the P1 lateral chain adopting a more favourable conformation for binding the target serine protease (Figure 8) [42,44,45].

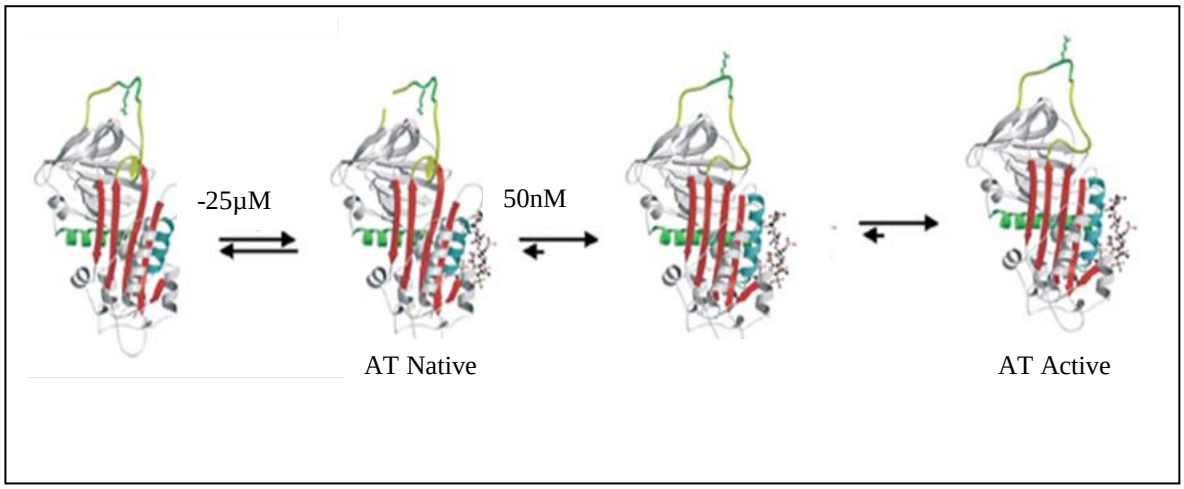


Figure 8. Conformational changes on antithrombin induced by binding of heparin.

The conformational changes brought about by binding the pentasaccharide sequence accelerates antithrombin's interaction with a central coagulation enzyme, factor Xa (FXa), by 100-fold [43]. Therefore, a synthetic form of this pentasacharide sequence is currently used as an effective therapeutic anticoagulant an effective anticoagulant [46]. Moreover, the exposition of this new exosites for FXa and FIXa contributes to increase the anticoagulant activity of the active form [47].

Full acceleration of the inhibition of serine proteases, particularly that of another major target, thrombin [43] but also FVIIa [48], requires the presence of heparin chains which are longer than 26 residues [43] to facilitate the formation of a ternary complex by serpin, protease and heparin (Figure 9) [45,49]. This complex accelerates the interaction with serine proteases by up to 17,000-fold [43,45,49].

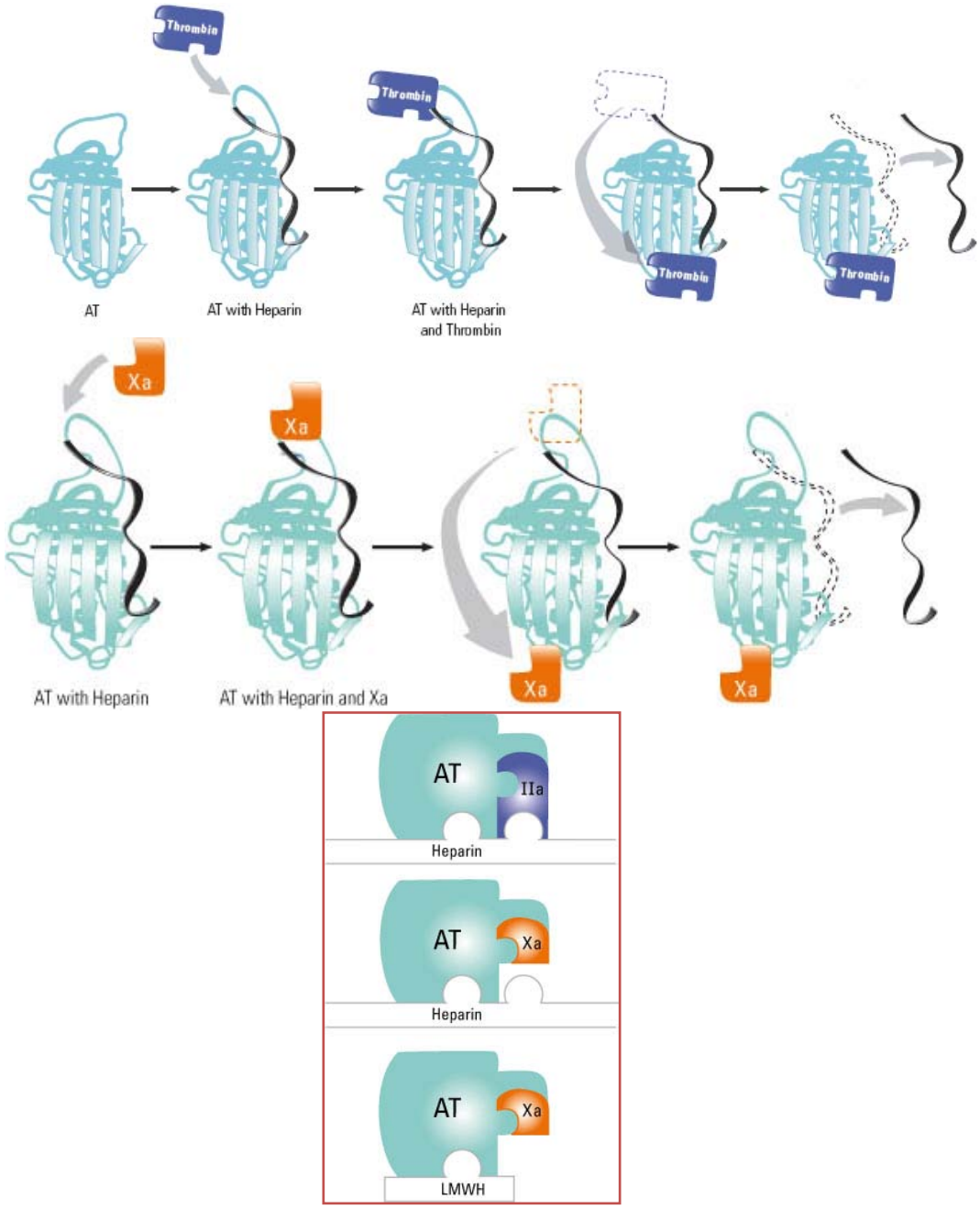


Figure 9. Inhibition of FXa and thrombin by antithrombin. Relevance of the length of heparin. AT: Antithrombin. LMWH: Low molecular weight heparin.

The control of the activity of antithrombin by glycosaminoglycans has additional effects, to delay and to localize this strong anticoagulant capacity. Otherwise, the high levels of antithrombin in plasma might impair the procoagulant response necessary for a correct haemostasis under vascular damage [50] (Figure 10).

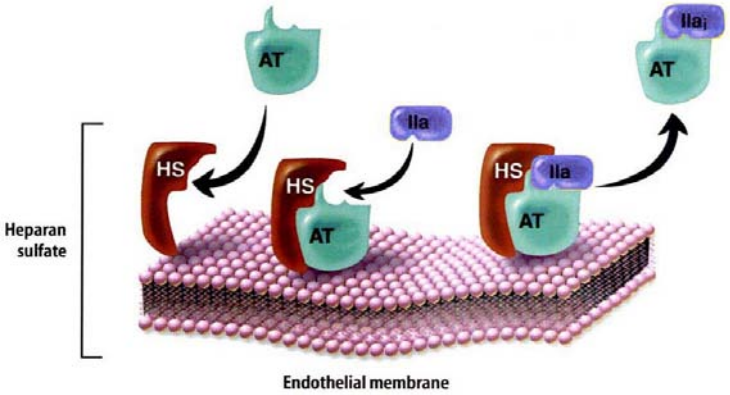


Figure 10. Localization of the anticoagulant activity of antithrombin (AT) to the endothelial surface by heparan sulfates (HS). IIa: Thrombin.

Therefore, to properly fulfil its inhibitory mechanism, antithrombin needs an extraordinary structural flexibility. In addition to the conformational changes shared with other serpins, antithrombin also requires an additional conformational change to fully activate it. Accordingly, this flexibility also makes antithrombin more susceptible than other serpins to minor structural changes that could potentially result in a missfolded conformation or inactive forms. Thus, in the normal native state, antithrombin contains a 5-strand β–sheet (β–sheet A) and a surface-exposed RCL (bridging the C-terminus of strand 5A to the N-terminus of strand 1C). This state is, however, not the most stable. An astounding increase in thermodynamic stability (best estimate – 32 kcal mol^{-1}) [51] can be achieved through the incorporation of the RCL into β-sheet A, triggered through extension of strand 1C, to form the so-called latent conformation, which is not obviously inhibitory (Figure 11). The identification of up to 5% of plasma antithrombin in a latent conformation has suggested that this transformation could be a consequence of the senescence of the native molecule [52,53].

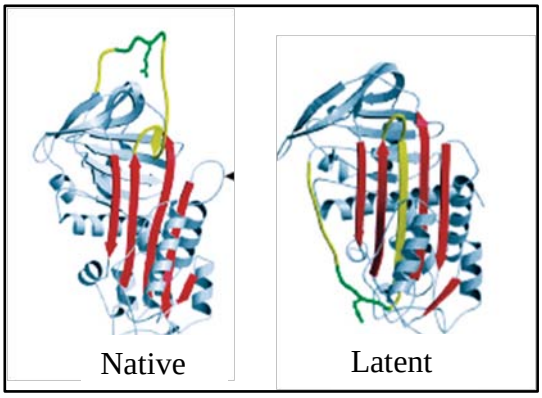


Figure 11. Native and latent conformations of antithrombin (PDB: 1T1FA and PDB: 2BEH, respectively). The RCL is displayed in yellow, while the central β-sheet A is in red.

Finally, besides its key role in coagulation, antithrombin has also been described to be important in other biological functions. A relaxed conformation seems to have anti-angiogenic activity [54]. Moreover, a potent anti-inflammatory activity has also been tightly associated with its heparin-binding domain [55]. Finally, our group has recently found evidences suggesting a potential anti-apoptotic role of antithrombin [56].

4.2 Genetics

SERPINC1, the gene encoding antithrombin (GenBank X68793.1, OMIM# 107300) is located at chromosome 1q23-25.1 and comprises 7 exons, encompassing 13.5 kb of genomic DNA (Figure 12) [57,58] (start in 172.139.562 to 172.153.096 of pter; access number GC01M172139: http://genecards.weizmann.ac.il/geneloc-bin/display_map.pl?chr_nr=1&range_type=gc_id&gc_id=GC01M172139&contig=). It is remarkable that only a few missense polymorphisms have been described, in concordance with the structural and functional relevance of this molecule. So far, 267 polymorphisms have been identified in *SERPINC1*, but its functional effects are unknown for most of them (http://www.genecards.org/cgi-bin/carddisp.pl?gene=SERPINC1&snp=267&search=serpinc1&rf=/home/genecards/current/website/carddisp.pl#snp).

The promoter region of this gene is largely unknown. The promoter lacks TATA element and GC-rich regions [58-60] and few studies, most of them using deletional analysis and reporter or DNase I footprint assays, have identified few regions potentially involved in the transcriptional regulation of this gene [60-63].

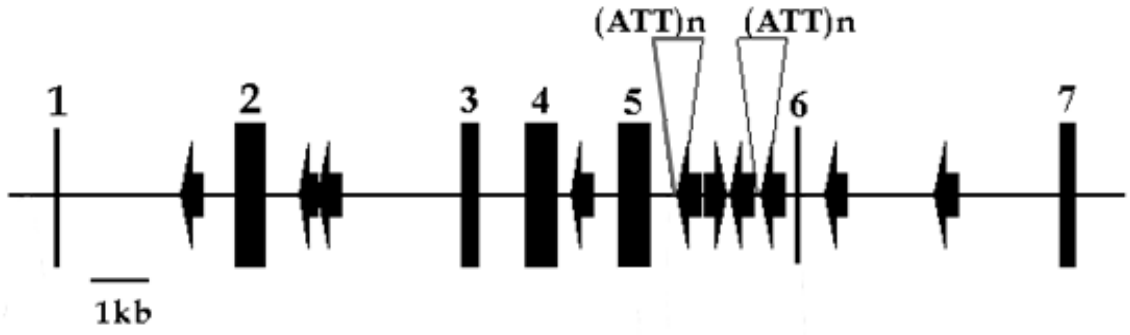


Figure 12. *SERPINC1* gen. Exons are marked as black boxes and Alu sequences with arrows.

4.3 Interindividual variability of antithrombin

Significant reduction of anticoagulant activity of antithrombin (values around 50%) deeply increases thrombotic risk. However, as for many other haemostatic factors, antithrombin levels have a wide range (70-120%) in general population. Factors such as sex, body mass index, oral contraceptives or race, may influence antithrombin levels. Furthermore, genetic factors must also play a role in determining antithrombin levels as revealed by the high heritability of this trait (h= 0.486) [64]. Polymorphisms on the gene encoding antithrombin (*SERPINC1*) could be involved in the interindividual variability of antithrombin. Indeed, our group has described that a single nucleotide polymorphism (SNP), rs2227589, located in intron 1 only explains up to 7.2% of the interindividual variation of antithrombin [65]. Therefore, the search for new genetic factors associated with the interindividual variability of antithrombin, with potential relevance on the risk of thrombosis, must continue by evaluating other *SERPINC1* polymorphisms, and other genes that indirectly may influence plasma levels or the anticoagulant function of this key haemostatic molecule. In fact, our group has previously confirmed the high conformational and functional sensitivity of the molecule to different posttranslational modifications and environmental factors [66-69].

4.4 Antithrombin deficiency and venous thrombosis

The extraordinary inhibitory mechanism of antithrombin and the large number of procoagulant serine proteases able to be inhibited by this serpin (Figure 1) explain the key haemostatic role of antithrombin. Therefore, even moderate reductions of antithrombin activity increase the risk of venous thrombosis, as recently suggested [70]. As firstly

described by Egeberg, heterozygous antithrombin deficiency significantly increases the risk of venous thrombosis (OR: 10-50) and the complete absence of this molecule causes embryonic lethality [71]. Moreover, the rate of thrombosis among carriers is significantly high (increasing with age up to 60% in 60 year-old carriers) (Figure 13) as well as the risk of recurrence [72].

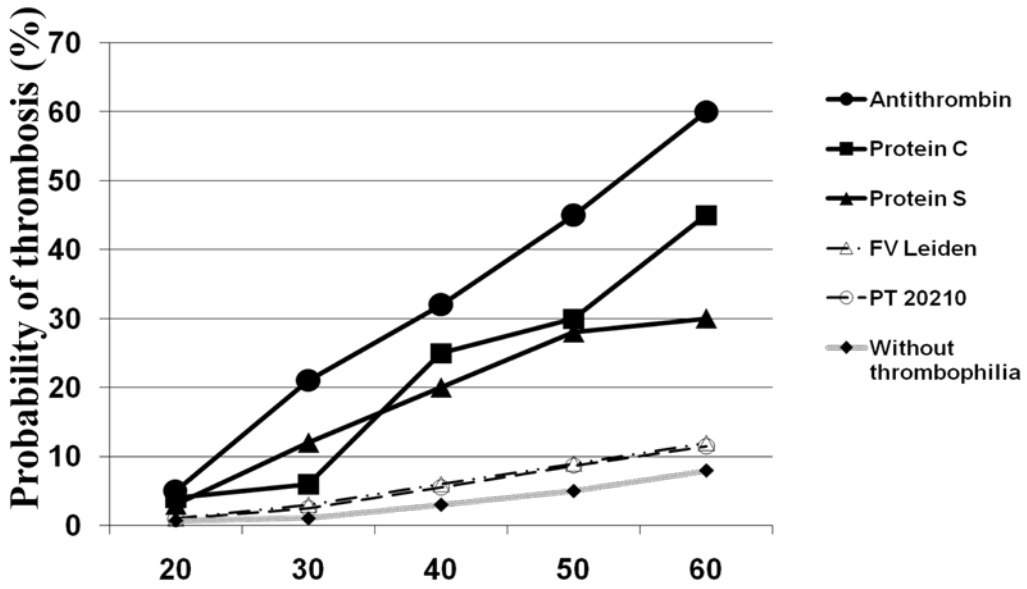


Figure 13. Probability of thrombosis among carriers of different thrombophilic defects according to age.

From a molecular point of view, more than 250 different mutations have been identified in subjects with antithrombin deficiency (Bayston T, Lane D. Imperial College of London. Antithrombin mutation database. http://www1.imperial.ac.uk/departmentofmedicine/divisions/experimentalmedicine/haematology/coag/antithrombin/ OMIM. http://omim.org/entry/107300 ; HGMD. http://www.hgmd.cf.ac.uk/ac/gene.php?gene=SERPINC1). Two large and recent cohorts have identified 69 new mutations [73,74]. Almost all antithrombin deficiencies are caused by heterozygous mutations that for most cases are point mutations, although large gene deletions have also been identified in *SERPINC1*. We have just described a relatively high frequency of compound heterozygous carriers with antithrombin deficiency [75], all of them carrying antithrombin Cambridge II, a common variant with a mild functional effect [76]. Only few homozygous mutations have been described, all mild functional variants [77].

Antithrombin deficiencies are normally classified attending to the presence of the mutant protein in plasma. Thus, type I deficiency states are characterized by a complete

loss of the mutant antithrombin protein in plasma, resulting in the patient having 50% of the normal levels and 50% of anticoagulant activity. This kind of antithrombin deficiency is easily explained by non sense, frameshift, or splicing mutations, as well as large deletions of the *SERPINC1* gene. Interestingly, up to 26 missense mutations associate with type I deficiency. Some of them have been explained by its effect on RNA stability. Interestingly, other aminoacid changes, particularly those affecting mobile regions (mainly the hinges of the RCL or the region involved in the shutter-like opening of the main β-sheet), cause oligomer formation and intracellular retention, which has revealed a new pathological mechanism causing antithrombin deficiency [78]. Thus, some cases of thrombosis may be included within the group of conformational diseases [79]. Indeed, antithrombin has been used to suggest a new and revolutionary mechanism of serpin polymerization by domain swap [80]. Type I mutants are strongly correlated with a strong predisposition towards venous thrombosis. In contrast, type II mutations are usually the consequence of missense single amino acid changes that result in a protein which is synthesized and usually secreted to the plasma with a normal or reduced rate, but they can be detected at least by antigenic methods as the molecule is functionally defective. According to the mutation-induced effect, three varieties of type II deficiencies have been defined:

1. Heparin binding defect. These mutations impair the affinity of antithrombin for heparin, resulting in a molecule that has normal inhibitory activity in the absence of heparin, but reduced activity in its presence. Obviously, most mutations causing this type of antithrombin deficiency affect directly the heparin biding domain. However, other mechanisms might also cause heparin binding defects. In this regard, our group has recently described an abnormal fucosylation of a variant of antithrombin that impairs heparin affinity, resulting in a type II deficiency with low heparin affinity [24], pointing out the relevance that glycosylation may have in the function of this molecule.
2. Reactive site defect. This group includes mutations normally localized in the RCL impairing the activity of the molecule independently of heparin binding.
3. Pleiotropic defect. This group includes a large variety of mutations with multiple effects on the activation by heparin and the inhibitory function of the protein.

The thrombotic risk associated with type II deficiencies is quite heterogeneous. Thus, some cases have severe clinical consequences, while others have mild consequences as

they have been identified in combination with other prothrombotic risk factors, such as factor V Leiden or in homozygous or hemizygous states. Recent studies from our group suggest different mechanisms to explain such a clinical heterogeneity. A post-translational mosaicism may explain the mild thrombotic role of heterozygous mutations as well as the survival of homozygous carriers of heparin binding defects, as the mutation does not abolish the activation induced by heparin in the β-glycoform [77]. In contrast, other mutations, in addition to a loss-of-function, also have a prothrombotic gain-of-function. Thus, mutations affecting the P1 residue of the RCL (Arg393), particularly antithrombin London (p.Arg393del), have higher heparin affinity that impairs activation of the wild-type molecule with low concentrations of this glycosaminoglycan [81]. We have recently described another gain-of-function for conformational mutations. The growing polymers of conformationally unstable antithrombin mutants may incorporate wild-type monomers under moderate stress conditions [82].

Accordingly, antithrombin assays are included among thrombophilic tests aiming to identify congenital antithrombin deficiency. The challenge in managing these patients is preventing potentially life-threatening thrombosis, while minimizing the equally significant risk of haemorrhage associated with long-term anticoagulation. This is achieved in the first instance by identifying high-risk episodes such as surgery, immobility and pregnancy for which prophylactic anticoagulation can be used in the short term. Prophylaxis for such periods is best provided by the use of low molecular weight heparin (LMWH) with substitution or addition of antithrombin concentrate in particularly high-risk circumstances. In the case of pregnancy, antithrombin concentrate is often used around the time of birth when LMWH may increase the risk of post-partum hemorrhage. As patients with congenital antithrombin deficiency get older, their thrombotic risk gradually increases and for many patients long-term oral anticoagulant therapy becomes unavoidable because of recurrent episodes of venous thromboembolism. The use of heparins might not be recommended in patients carrying type II deficiency with heparin binding defect, particularly in homozygous state. Concentrates of antithrombin for replacement therapy have been classically obtained from human plasma, and there is now a recombinant antithrombin with β-like properties [83].

OBJECTIVES

Hypothesis

There might be unknown elements modifying antithrombin levels with potential thrombotic implications as the current known elements do not explain themselves either this wide variability or the high heritability of this trait in the population and a significant proportion of patients with antithrombin deficiency without mutations in the *SERPINC1* gene.

The main objective of this thesis was to identify new genetic elements playing a role on the levels or function of plasma antithrombin and to determine possible physiopathological implications.

Specific objectives:

1. To identify genetic alterations in *SERPINC1* gene associated with interindividual variability of antithrombin and/or thrombotic risk.
2. To find out new antithrombin modifier genes and the potential molecular mechanisms underlying.
3. To assess the importance of a correct *N*-glycosylation on antithrombin levels and on the control of normal haemostasis.

MATERIAL AND METHODS

1. Patients and healthy subjects

This study included the following cohorts of subjects:

1) Three hundred and seven Spanish Caucasian healthy blood donors (138 male/169 female), with an average age of 43 years.
2) One hundred and eighteen subjects with antithrombin deficiency from 22 Spanish hospitals, 1 Portuguese hospital and 1 Israeli hospital. Two different groups were selected from this cohort.
    a. Twenty-nine unrelated thrombophilic patients with moderate antithrombin deficiency (70-90% anti-FXa activity of the reference value). In this group, sequencing analysis of *SERPINC1* revealed a relatively common point mutation in exon 7 responsible for the Cambridge II variant (p.Ala384Ser) [76] in four patients. Thus, our study was mainly focused to identify new factors that might be involved in the reduced antithrombin levels of the remaining 25 patients (named as P1-P25).
    b. Eighty nine unrelated and consecutive subjects with verified antithrombin deficiency (30-70%). Molecular analysis identified a genetic defect in the *SERPINC1* gene in 85 cases that explained the associated deficiency. Therefore, we addressed our interest to identify the mechanism underlying antithrombin deficiency of the remaining 4 cases that had neither gross deletion, nor mutation in the coding or flanking regions of the *SERPINC1* gene. In this rare but scientifically-fascinating group of subjects, we differentiated two subgroups based on familial history
        i. Two subjects (P26 and P27) that reported no relatives with antithrombin deficiency.
        ii. Two subjects that had up to three relatives from two generations with type I antithrombin deficiency (F1 and F2), sustaining a congenital defect.

3) Two patients with congenital disorder of glycosylation. Prof. J Jaeken kindly supplied two young Belgian patients with congenital disorder of glycosylation (CDG) type Ia, also named PMM2-CDG. The first patient carried two mutations in *PMM2* gene responsible for the amino acidic changes p.Phe119Leu and

p.Arg141His. The second patient also carried two mutations in *PMM2* leading to p.Pro113Leu and p.Thr237Arg substitutions, respectively.

Informed consent was obtained from all donors, patients or their legal representatives.

2. Blood sampling

Blood samples were obtained by venopuncture collection into 1:10 volume of trisodium citrate. Platelet poor plasma fractions were obtained (within 5 min after blood collection) by centrifugation at 4 ºC for 20 min at 2200x*g* and stored at -70 ºC. Genomic DNA was purified by the salting out procedure and stored at -20 ºC.

All subjects gave their informed consent to enter the study, which was approved by the ethics committee of the Centro Regional de Hemodonación (Murcia, Spain) and performed according to the declaration of Helsinki, as amended in Edinburgh in 2000.

3. Antithrombin levels and other thrombophilic tests

Routine thrombophilic tests including protein C, protein S, antiphospholipid antibodies, FV Leiden, and prothrombin G20210A were performed, as described elsewhere [84]. Plasma FXa-inhibiting activity was measured using a chromogenic method in presence of heparin (HemosIL TH, Instrumentation Laboratory) as previously reported [68]. Antithrombin antigen levels in plasma were determined by rocket immunoelectrophoresis and/or a home-made ELISA. For both parameters, values were expressed as a percentage of the result observed in a pool of citrated plasma from 100 healthy subjects.

4. Amplification and sequencing of *SERPINC1*

Polymerase chain reactions (PCR) covering the whole *SERPINC1* gene and the potential promoter region were carried out using Expand Long Template Polymerase (Roche, Spain) and the oligonucleotide sets described in Table 2. PCR products were purified and sequenced with ABI Prism Big Dye Terminator v3.1 Cycle sequencing kit and resolved on a 3130xl Genetic Analyzer (Applied Biosystems, Spain). Comparison with the reference sequence (GenBank accession number NG_012462.1) was performed with SeqScape v2.5 software (Applied Biosystems, Spain).

Table 2. Oligonucleotides used for PCR amplification of the *SERPINC1* gene. E: Exon; I: Intron. F: forward primer; R: reverse primer.

Amplicon	Primer name	Oligonucleotide sequence (5´-3')	Size (bp)
5' region	AT_-1F	GGACCTTATTAACATCTGAC	743
	AT_-728R	CCAAGTTCACAGGGCGCATG	
	AT_-660F	GGAGTCCTTGATCACACAGCA	1088
	AT_IN1R	GTCTTTGACTGTAACTACCAG	
	AT_0,5F	GGCTACTGTTCACACTGCACATG	1067
	AT_1R	GAAAGCTCACCCCTCTTAC	
E1-E2	AT_1F	CTGTCCTCTGGAACCTCTGCGAGA	2880
	AT_2R	GGTTGAGGAATCATTGGACTTG	
I1	AT_19F	ATCCGGGAAGAGAGCAAATGC	1200
	AT_31R	CTCAAAACAGCAACAACAAAC	
E2-E3	AT-2F	TGCAGCCTAGCTTAACTTGGCATTT	3180
	AT_3aR	AGAGGAAGAACTCGGAGGTCAGG	
E3-E4	AT_3aF	AACTAGGCAGCCCACCAAACCC	1410
	AT_3bR	GAAGAGCAAGAGGAAGTCCCT	
E4-E5	AT_3bF	TTGAATAGCACAGGTGAGTAGGTT	1380
	AT_4R	AAGGGAGGAAACTCCTTCCTAG	
E5	AT_4F	TGTGTTCTTACTTTGTGATTCTCT	391
	AT_4R	AAGGGAGGAAACTCCTTCCTAG	
E5-E6	AT_4F	TGTGTTCTTACTTTGTGATTCTCT	2521
	AT_5R	CATGCATCTCCTTTCTGTACCC	
I5-E6	AT_95F	TGGAGAGGAATTTGAAAG	1800
	AT_5R	CATGCATCTCCTTTCTGTACCC	
I5	AT_95F	TGGAGAGGAATTTGAAAG	1080
	AT_106R	AACATCTTTCTTTCCAGTCTG	
E6-E7	AT_5F	TTCTCCCATCTCACAAAGAC	3660
	AT_6R	AGAGGTGCAAAGAATAAGAA	
I6	AT_130F	GGGAGACTAGGGTGTTGA	1200
	AT_142R	CAGATCTAGAGGGAAACACC	
I6-E7	AT_130F	GGGAGACTAGGGTGTTGA	1860
	AT_6R	AGAGGTGCAAAGAATAAGAA	
3' region	AT_130F	GGGAGACTAGGGTGTTGA	2860
	AT_159R	ACACACAAGTAATAACATCCAC	
	AT_3'F	TTATCTTCATGGGCAGAGTAGC	1000
	AT_159R	ACACACAAGTAATAACATCCAC	

5. Multiplex Ligation-dependent Probe Amplification (MLPA)

Multiplex Ligation-dependent Probe Amplification technique, a multiplex PCR method able to distinguish sequences differing in only one nucleotide, was used to detect abnormal copy numbers of up to 50 different genomic DNA or RNA sequences. To

specifically detect deletions or duplications in the *SERPINC1* gene we used the *SALSA*® *MLPA*® Kit P227 SerpinC1 (MRC-Holland).

6. Genotyping of *SERPINC1* genetic alterations

The genotype of rs5778785 was determined by high resolution melting (HRM) analysis following the Type-it HRM PCR protocol (Qiagen, Spain). The g.2085 T>C transition was genotyped by a HRM/razor probe technique (Figure 14).

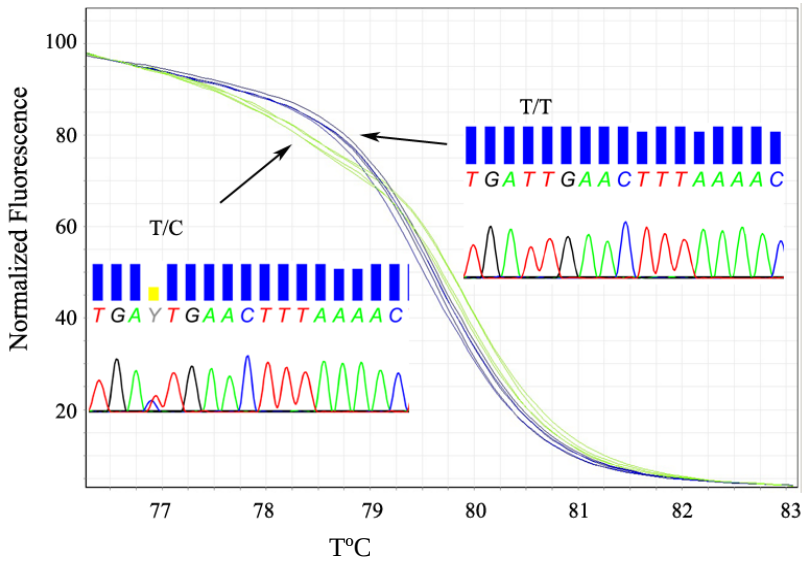


Figure 14. HRM-razor melting curves and electropherograms of wild type and heterozygous g.2085 T>C genotypes.

These two HRM analyses were carried out in Rotor-Gene Q 6000 (Qiagen, Spain) with the primers and probes shown in Table 3.

Nine *SERPINC1* polymorphisms were genotyped by PCR-allelic specific restriction assays using primers and restriction enzymes described in Table 3, as previously described [85]. For some polymorphisms, mutated primers were designed (Table 3).

rs2227589 was genotyped by a Taqman SNP Genotyping Assay (C__16180170_20, Applied-Biosystem), as previously described [65].

The g.2143 C>G mutation was studied by PCR-allelic specific restriction assay using: 5'-GGAGTCCTTGATCACACAGCA-3' forward and 5'-GTCTTTGACTGTAACTACCAG-3' reverse primers and Bsa JI restriction enzyme (New England BioLabs, Spain) following manufacturer's specifications.

For all genotyping methods, results were confirmed by sequencing.

Table 3. Genotyping by PCR-allele specific restriction assay and High Resolution Melting. Mutated nucleotides are underlined.

Polymorphism	Amplification/Probe Oligonucleotides	PCR Size (bp)	Method of genotyping
rs2227588	AT_0.5F: GGCTACTGTTCACACTGCACATG AT_-728R: CCAAGTTCACAGGGCGCATG	198	Mnl I (NEB)
rs2227590	AT_18F:CATTTCAAGTGCTCTCCCTCC AT_19R:GCATTTGCTCTCTTCCCCGAT	203	FokI (Takara)
rs2227596	AT_30F:ATCTTCAGCACACAGATATTG AT_33R:GCTGGGTGTGGTGGCGC	229	Tsp 45I (NEB)
rs2227603	AT_53F:TTGAGACCAGGGGCCCAG AT_57R:AATTGACCACTAGAATAGAAG	408	FokI (Takara)
rs941989	AT_60F:GCTTCCAAACCACACATGTTC AT_62R:CCATGCTCTCTCCTCAGGC	241	HinfI (Promega)
rs2227607	AT_71F:GACTGCTGGAGGGCTGAC AT_73R:GAACTGGCATCTTCAAAACACCC	186	NlaIII (NEB)
rs2759328	AT_112F:TCCAAACTGAATTCCCATCTGCG AT_5R: CATGCATCTCCTTTCTGTACCC	190	FauI (Fermentas)
rs677	AT_114F:TGAGTACACCTTCCCCAC AT_115R:CTGTATTATAGCAGGCCTGTGG	126	DdeI (NEB)
rs2227616	AT_130F:GGGAGACTAGGGTGTTGA AT_133R:CCGGGCACGGTGGCTCATGCCGC TAA	332	BstEII (NEB)
rs5778785	AT_95F:TGGAGAGGAATTTGAAAG AT_97R:GTGAGCCGAGATCGCACCGC	178	HRM
g.2085 T>C (rs78972925)	AT_13F:GAGGGAGTGTGGGCAAGAGA AT_13R:AGGGTTAAGCAAAGTGTAGAG Razor:ACCCGGACTAACTTGAAAT	107 Razor probe	HRM/Razor probe

NEB: New England BioLabs; HRM: High Resolution Melting

7. Prediction of transcription factor binding sites to the *SERPINC1* potential promoter

The search for sequences with potential transcriptional relevance affected by genetic variations identified in this study was performed by TESS and TFSEARCH programs, based

on TRANSFAC databases (http://www.cbil.upenn.edu/cgi-bin/tess/tess and http://www.cbrc.jp/research/db/TFSEARCH).

8. Cloning and reporter assay

The 5' non transcribed portion of *SERPINC1* (838 bp) from one individual heterozygous for the g.2143 C>G mutation was amplified using Expand High Fidelity DNA polymerase (Roche, Spain) to minimize nucleotides misincorporation with 5'-GGAGTCCTTGATCACACAGCA-3' (forward) and 5'-AATCTCGCAGAGGTTCCAGA-3' (reverse) primers. PCR products were ligated into a T/A cloning vector (pCR®2.1, Invitrogen, Spain). Clones with wild-type and mutant alleles were selected by PCR-allelic restriction assay with Bsa JI, and the inserts cloned into Kpn I and Sac I sites of luciferase pGL3 basic vector (Promega, Spain). Competent cells were transformed, plasmid DNA was purified using a commercial kit (Plasmid Midi Kit, Qiagen, Spain), and they were sequenced in both directions to faithfully confirm the wild-type and mutant sequences.

The ability of each sequence to promote transcription of luciferase gene was transiently tested in human cell lines PLC-PRF-5 (human hepatome cell line) and HepG2 (Human hepatocellular liver carcinoma cell line) both with antithrombin constitutive expression.

All plasmid DNA were quantified by Nanodrop (Thermo Scientific, Spain). Eighteen hours before transfection, cells were plated at 80,000 per well of a 24-well plated (3 replicates per clone for each cell line and in each experiment) and cultured in Dulbecco's Modified Eagle Medium (Invitrogen, Spain), at 37 ºC with 5% CO_2. Reporter plasmid (1,000 ng) along with 100 ng of Renilla luciferase control plasmid pRL-TK (Promega, Spain) were transfected using Lipofectamine 2,000 (Invitrogen, Spain) for PLC-PRF-5 cell line and GenJet reagent (SignaGen® Laboratories) for HepG2 cell line, according to the manufacturer's instructions. Luciferase assays were performed 48 h after transfection by using the Dual-GloTM Luciferase assay System (Promega, Spain). Promoter activity was normalized by dividing luciferase activity by Renilla luciferase activity for each transfected well. pGL3 control vector (Promega, Spain) was used as a transfection positive control.

9. Genome wide association study

We carried out a genome wide genotype-phenotype association study in the GAIT (Genetic Analysis of Idiopathic Thrombophilia) study, which included 352 individuals from

21 extended Spanish families [3]. Twelve of these families were selected on the basis of a proband with idiopathic thrombophilia, whereas the remaining 9 families were selected randomly. Average pedigree size was 19. Importantly, no family had congenital antithrombin deficiency.

A genome-wide set of 307,984 SNPs was typed in all of the participants using the Infinium® 317k Beadchip on the Illumina platform (San Diego, CA, USA). Genotype imputation was performed with Merlin [86] to avoid missing values and all genotypes were checked for Mendelian inconsistencies. In addition, any SNP with call rate<95%, MAF<0.025 or failing to fit Hardy-Weinberg proportions taking into account multiple testing ($p<5\times10^{-7}$) was removed from the study. In total, 24,547 SNPs failed to pass the data cleaning criteria, leaving a set of 283,437 SNPs for further analysis.

10. Validation study

A cohort of 307 Spanish Caucasian healthy blood donors (138 males/169 females) with a mean age of 43 years and from a different region than the GAIT study was selected to replicate significant associations identified in the GWAS. Genotyping was performed using TaqMan® probes specified in Table 4.

Table 4. TaqMan® probes used for genotyping in the validation study. The SNPs identified in the GWAS associated with anti-FXa activity and the genes affected are also shown**.** *rs713703 located in *LARGE* and strongly linked to rs762057, was not evaluated.

TaqMan® probe	SNP	Gene
c_15834925_20	rs2152192	*SAMD3*
c_2747618_20	rs1411771	*DISC1*
c_9839576_10	rs13193455	*LOC 154449*
c_1282266_10	rs11681944	*C1D*
c_1298059_10	rs6768189	*CACNA2D3*
c_ 30869706_10	rs10880942	*SLC38A1*
c_11634765_10	rs1860867	*COBL*
c_16207835_20	rs2356895	*LOC283553*
c_30230995_10	rs9896932	*CD7*

TaqMan® probe	SNP	Gene
c_2486441_10	rs762057	*LARGE*
c_1839936_10	rs713703	*LARGE*
c_790509_10	rs240082	*LARGE*

11. Case-control study

The SNP found to be significantly associated with antithrombin levels in the GWAS and verified in the replication study, rs762057, was genotyped using a TaqMan® probe (c_2486441_10) (Applied Biosystems) in a case-control study including 1018 patients with venous thrombosis and 1018 controls previously described [87]. Briefly, patients with venous thrombosis diagnosed appropriately were enrolled from the files of the anticoagulation clinics in four Spanish Hospitals. Patients with known malignant disorders were excluded. Controls were randomly selected among two sources: blood donors and traumatology and ophthalmology patients matched by age, gender, race, and geographical distribution with 1018 cases.

12. Statistical analyses

In all studies, deviation from Hardy-Weinberg equilibrium (HWE) was investigated using a standard χ^2 with 1 degree of freedom. Allele and genotype frequencies, haplotype analysis, association of haplotypes with anti-FXa activity and linkage disequilibrium analysis were calculated with the SNPstats software [88].

In the GAIT study, HWE was tested using parental data only. Association between SNPs and plasma anti-FXa activity was tested using a measured genotype association analysis assuming additive allele effects. This analysis was carried out using the variance-components methodology implemented in the SOLAR Version 4.0 software (Southwest Foundation for Biomedical Research, http://solar.sfbrgenetics.org/download.html) [89]. Analyses were adjusted for age, gender, and body mass index, and in females, for oral contraception as well.

Association between SNPs and plasma anti-FXa activity in the validation study was tested by a Student's *t*-test following a dominant model and adjusted for factors described to influence antithrombin levels (age, gender and the *SERPINC1* rs2227589 polymorphism)

[65,90]. Data are presented as mean ± standard deviation. This analysis was carried out using the Statistical Package for Social Science (SPSS version 15.0, USA).

In the case-control study, univariate and multivariate analysis were performed using the SPSS statistical package.

13. *LARGE* and *SERPINC1* gene expression

LARGE gene expression was assessed in mononuclear cells of 10 healthy subjects by qRT-PCR using Hs00893935_m1 TaqMan® Gene Expression Assay (Applied Biosystem) and beta-actin (Hs99999903_m1) as constitutive reference gene. The correlation between plasma anti-FXa activity and *LARGE* expression was determined by linear regression. These analyses were also done in HepG2 and Human Embryonic Kidney cells expressing the Epstein Barr Nuclear Antigen 1 (HEK-EBNA) cell lines transfected with *LARGE* gene silencers.

SERPINC1 gene expression in HepG2 and HEK-EBNA cell lines transfected with *LARGE* gene silencers was determined by qRT-PCR with SYBR® Green-Based Detection (Applied Biosystem) using Tubuline beta-2C chain as constitutive reference gene. Primers for amplification were: SERPINC1-F: TGTTCAGCCCTGAAAAGTCC; SERPINC1-R: GCTGCTTCACTGCCTTCTTC; TUBULINE-F: GGCCGGACAACTTCGTTT and TUBULINE-R: CCGCGCCTTCTGTGTAGT.

14. Purification of plasma antithrombin. Proteomic and glycomic analysis

Plasma predominant antithrombin glycoform (α) from two subjects selected from 10 healthy blood donors, with the highest and lowest *LARGE* expression values, as well as antithrombin minor glycoform (β, with 3 *N*-glycans) from a pool of 100 healthy blood donors were purified by Fast Protein Liquid Chromatography (FPLC) (ÄKTA Purifier from GE Healthcare) using heparin affinity chromatography on HiTrap Heparin columns (GE Healthcare), in 100 mM Tris-HCl and 10 mM citric acid, with a gradient from 0 to 0.25M NaCl and a step of 2M NaCl. Fractions with antithrombin were applied to a HiTrap Q column (GE Healthcare). Finally, proteins eluted were desalted through a dialysis tubing (Sigma Aldrich) and stored at -70 ºC, prior to analysis. The molecular mass and glucidic components of these molecules were determined by MALDI-TOF-MS (Matrix Assited Laser Desorption Ionization-Time-of-Flight) analysis and HILIC (Hydrophilic Interaction Liquid Chromatography) HPLC (High Performance Liquid Chromatography). Briefly, for

protein molecular weight determination a solution of 3,5-dimethoxy-4-hydroxycinnamic acid (10 g/L) in acetonitrile (ACN)/water/trifluoroacetic acid (TFA) (50:50:0.1 by vol.) was employed. Experiments were carried out on a Voyager-DE™ STR Biospectrometry workstation (Applied Biosystems), equipped with a N_2 laser (337 nm). Samples were measured both in the linear, providing information on the total number of different structures, and in the reflectron mode for identification of molecular formulas based on precise mass measurements. Recorded data were processed with Data Explorer™ Software (Applied Biosystems). The analysis of the *N*-glycans was performed by HILIC chromatography. Briefly, *N*-glycans were released with *N*-glycosidase F (Roche Diagnostics GmbH, Mannheim) following prior denaturing (5 min at 95 ºC in 150 mM sodium phosphate buffer, pH 7.4). Afterwards, samples were chilled on ice and digested with 0.6 U *N*-glycosidase F by incubation at 37 ºC, for 15 hours. Glycans were labeled as described [91] and subjected to chromatographic separation on an Agilent 1100 HPLC equipped with a fluorescence detector (1100 Agilent fluorescence module) using excitation and emission wavelengths of λ=330 nm and λ=420 nm, respectively. The following gradient conditions were employed on a ACQUITY UPLC™ BEH HILIC column (2.1 x 150 mm, 1.7 µm): solvent A was 10% 50 mM ammonium formate (pH 4.4) in 90% ACN, solvent B was 90% 50 mM ammonium formate (pH 4.4) in 10% ACN, and the flow rate was 15 µl/min. Following injection, samples were eluted by a linear gradient of 20-55% B over 100 min, followed by a linear gradient of 55-100% B over the next 5 min. The column was eluted using 100% B for 2 min, and subsequently re-equilibrated in 20% B before injection of the next sample. The system was calibrated in glucose units (GU) using a 2-AB-labelled dextran hydrolysate. The total running time was 125 min [91]. Mass spectrometric analyses of 2-aminobenzamide (2-AB)-labeled glycans were performed 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg/ml) in ACN: H_2O (50:50 v/v). Typically, spectra of sialylated *N*-glycans were acquired in linear mode with negative polarity, and in neutral *N*-glycans reflectron mode and positive polarity. External calibration of the spectrometer was performed using a mixture of 2-AB-labelled glucose oligomers in the positive-ion mode and 2-AB-derivatised fetuin N-glycans in the negative mode. Recorded data were processed with Data ExplorerTM Software (Applied Biosystems).

15. *LARGE* gene silencing and effect on *N*-glycoproteins

For these experiments we used two cell lines expressing antithrombin: HepG2 with constitutive antithrombin expression, and HEK-EBNA transiently transfected with pCEP4-AT plasmid (generously provided by Prof. JA Huntington) that expressed high levels of the beta glycoform of human antithrombin [92].

HepG2 and HEK-EBNA cell lines were grown to 60% confluence at 37 °C, 5% CO_2, in DMEM (Invitrogen) supplemented with 5% fetal bovine serum (Sigma-Aldrich). Then, they were transfected with 50 nM of specific *LARGE* siRNA: s17620 (Applied Biosystems) for 30 minutes in OptiMEM with siPORT™ (Applied Biosystem). Appropriate controls: transfections without siRNA, or with 50 nM of scramble siRNA (*Silencer*® Negative Control AM4611, Applied Biosystem) were used. After 12 hours, cells were washed with PBS and exchanged into CD-CHO medium (Invitrogen) supplemented with 4 mM L-glutamine (Invitrogen). Cells were grown at 37 ºC for 48 hours. Then, RNA was purified using TRIzol® Reagent (Invitrogen) following manufacturer instructions. We determined the silencing efficiency evaluating *LARGE* and *SERPINC1* expression by qRT-PCR, as indicated above. Additionally, conditioned medium was harvested and in case of HepG2 cell cultures, concentrated 5-fold using a CentriVap Concentrator (Labconco). The levels of secreted antithrombin, transferrin, prothrombin and α1-antitrypsin in conditioned medium were determined by western blotting, essentially as described elsewhere [93]. Briefly, electrophoresis was carried out using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) in 10% (w/v) polyacrylamide gels under reducing conditions. Proteins were transblotted onto a polyvinylidene difluoride membrane. Proteins were immunostained with specific rabbit [anti-human antithrombin (Sigma Aldrich) and anti-human α1-antitrypsin (Dako Diagnostics, Denmark)], goat [anti-human transferrin (Sigma Aldrich)], or sheep [anti-human prothrombin (Cerdalane laboratories, Canada)] polyclonal antibodies; followed by proper secondary IgG-horseradish peroxidase conjugates (GE Healthcare), and ECL detection (GE Healthcare). Antithrombin levels in the conditioned medium were also determined by ELISA, as previously described [93]. Additionally, anti-FXa activity of conditioned medium was measured by the chromogenic method described above. Finally, we also evaluated the intracellular content of antithrombin by western blotting and immunofluorescence, basically as previously described [93]. Briefly, cells were extensively washed with sterile PBS and then lysated with 50 µl of lysis buffer (10 mM TrisHCl, 0.5 mM DTT, 0.035% SDS, 1 mM EGTA, 50 mM sodium fluoride, 50 µM sodium

orthovanadate, 5 mM benzamidine and 20 mM phenylmethylsulphonyl fluoride) and stored at -70 ºC, prior to analysis. Intracellular antithrombin was evaluated by Western blotting, essentially as indicated above. For immunofluorescence analysis, cells were fixed with an equal volume of 4% paraformaldehyde in PBS buffer pH 7.4 (22 ºC, 20 min). After fixation, cells were washed with PBS, permeabilized with 0.1% Saponin, 0.2% Gelatin, 0.02% Azide (3x5 min). All subsequent incubations and washes contained 0.1% Saponin, 0.2% Gelatin, 0.02% Azide in PBS buffer. Anti-antithrombin antibody was used at 1:1,000 and incubated for 1 h at 22 ºC. Indirect immunofluorescence was carried out using the appropriate fluorescein conjugated goat anti-Rabbit IgG (Vector laboratories, USA) 1:1,000. Fluorescence was analyzed on a Confocal Microscope LEICA TCS-SP2 using its associated software (Leica Microsystems).

16. Electrophoretic evaluation

Electrophoretic analyses were performed by using SDS and native PAGE under reducing and non reducing conditions as previously reported [94]. Plasma samples were transblotted onto a PVDF membrane and immunostained with different antibodies: rabbit anti-human antithrombin polyclonal antibody (A9522, Sigma-Adrich), rabbit anti-human α1-antitrypsin polyclonal antibody (A0012, Sigma-Aldrich), goat anti-transferrin polyclonal antibody (T2027, Sigma-Aldrich), followed by donkey anti-rabbit IgG-horseradish peroxidase conjugate (NA9340V, GE Healthcare) or mouse monoclonal anti-Goat/Sheep IgG–Peroxidase (A9452, Sigma-Aldrich) with detection via an EC kit (Amersham Biosciences, Piscataway, NJ).

17. *N*-glycosidase F, neuraminidase and galactosidase plasma treatment

Plasma from probands and healthy donors (1µl) were treated with 1 U α2–3,6,8,9 neuraminidase (sialidase) (N 3786, Sigma-Aldrich, Saint Louis, USA) at 37 ºC for 24h in 50 mM sodium phosphate buffer, pH 6. Afterward, samples were treated with β-galactosidase (G 1288 Sigma-Aldrich, Saint Louis, USA) at 37 ºC for 24h in the same buffer.

Treatment with *N*-glycosidase F (Roche Diagnostics GmbH, Mannheim, Germany) required a previous denaturing step (5 min at 95 ºC in 150 mM sodium phosphate buffer, pH 7.4). Then, samples were chilled on ice and digested with 0.6 U *N*-glycosidase F by incubation at 37 ºC for 24h.

18. Purification and proteomic analysis of an antithrombin variant from a CDGIa-like patient

Antithrombin variant from the patient was purified by heparin affinity chromatography at pH 7.4 onto HiTrap Heparin columns (GE Healthcare, Barcelona, Spain), as previously described [24] using an ÄKTA Purifier (GE Healthcare, Barcelona, Spain) in 50 mM Tris-HCl, in a gradient from 0.15 to 2 M NaCl. Fractions with variant antithrombin were applied to a HiTrap Q column (1 mL) (GE Healthcare, Barcelona, Spain). Proteins eluted with 50 mM Tris-HCl pH 7.4, in a gradient from 0 to 1 M NaCl. Fractions containing variant antithrombin were pooled, desalted over 5 mL HiTrap Desalting columns (GE Healthcare, Barcelona, Spain) and stored at -70ºC. Protein purity was checked by 8% SDS-PAGE.

A solution of 3,5-dimethoxy-4-hydroxycinnamic acid (10 g/L) in acetonitrile (ACN)/water/trifluoroacetic acid (TFA) (50:50:0.1 by vol.) was chosen for protein analysis. Experiments were carried out on a Voyager-DE™ STR Biospectrometry workstation (Applied Biosystems, Madrid, Spain), equipped with a N_2 laser (337 nm). Recorded data were processed with Data Explorer™ Software (Applied Biosystems, Madrid, Spain).

19. Identification and quantification of transferrin glycoforms

Transferrin in serum was completely saturated with iron by mixing 100 µL of serum with 20 µL of FeNTA (final concentration, 1.67 mM), a well-known transferrin iron donor. Each transferrin molecule can bind a maximum of two iron ions, but typically, serum transferrin is only partially (~30%) saturated. Because the FeNTA complex gives almost instant iron saturation of transferrin, no additional incubation time was necessary at this stage. Thereafter, the lipoproteins in the serum sample were precipitated by addition of 20 µL of the dextran sulfate–$CaCl_2$ solution (10% dextran sulfate, 1 mM $CaCl_2$). After gentle mixing, the samples were left in the cold (4 °C) for 30–60 min and then centrifuged at 3500*g* at 4 °C for 5 min. Of the clear supernatant, 130 µL were withdrawn and concentrated at 3500*g* at 4 °C in an YM-10 Microcon filter (Millipore, Darmstadt, Germany). The concentrated sample was then diluted with 260 µL of water and transferred to glass HPLC vials.

We injected 100 µL of the treated serum samples into the HPLC system, which consisted of an Agilent 1100 Series Liquid Chromatography, equipped with a quaternary pump and degasser, thermostated autosampler and column compartment, multiple wavelength detector, and ChemStation software (Agilent Technologies, Madrid, Spain).

Separation of the transferrin glycoforms was performed on a SOURCE® 15Q PE 4.6/100 anion-exchange chromatography column (Amersham Biosciences) at 22 °C, by linear salt gradient elution at a flow rate of 1.0 mL/min. Buffer A consisted of 10 mM Bis-Tris, adjusted to pH 7.0 with 2 M HCl; buffer B was the same buffer containing 0.5 M NaCl, pH 6.2; and buffer C was 10 mM Bis-Tris, pH 6.2. After every sample injected, the column was regenerated and cleaned by washing with 2.0 M NaCl (buffer D) and finally reequilibrated with the starting buffer [95]. All buffers were degassed and filtered through a 0.45 µm filter before use. Total run time under these conditions was 40 min. Quantification of the transferrin glycoforms relied on the selective absorbance of the iron–transferrin complex at 470 nm referred as the area [95].

Q-TOF (Quadrupole Time-Of-Flight) analysis of transferrin forms was done by Dr. D Lefeber (Netherlands) following a method developed in his laboratory (Monique Van Scherpenzeel; Gerry Steenbergen; Dirk Lefeber N-glycan profiling of human serum transferrin using microfluidic chip-Q-TOF with integrated glycan cleavage. ASMS2012).

20. *PMM2* gene PCR amplification and sequencing

Primers used for PCR amplification and sequencing of *PMM2* gene (OMIM# 601785) are described in Table 5.

Table 5. Oligonucleotides used for PCR amplification of the *PMM2* gene.

Amplicon	Primer name	Oligonucleotide sequence (5´-3') (TmºC)	Size (bp)
Exon 1	PMM2_Ex1F	CCTCCTCTTCCCGACGTGCC (62.5)	254
	PMM2_Ex1R	CTACCCTGGGTGGTCGATAGC (59.1)	
Exon 2	PMM2_Ex2F	GACTTATGTACTTGTGTTACCC (50.7)	266
	PMM2_Ex2R	CTTGGCAACAATATCCTCATAAG (51.6)	
Exon 3	PMM2_Ex3F	CTGGAGTTTAGCGGTTTTATTG (51.9)	261
	PMM2_Ex3R	CCTAGAGGCATTCATTGTG (53.4)	
Exon 4	PMM2_Ex4F	GCTCCTGCTAAATCAAGTAAC (51.2)	218
	PMM2_Ex4R	CCTATTTGGAGAATGCCCAC (53.4)	
Exon 5	PMM2_Ex5F	GGAGAAACTCTGTCACCCTT (54.1)	173
	PMM2_Ex5R	CATAAACCCAGCCATTCACC (53.9)	
Exon 6	PMM2_Ex6F	CCTACCTTTGTGGCCAGTAG (55.1)	208
	PMM2_Ex6R	CTCTGGGAAATGGGTATCCAAG (55.1)	
Exon 7	PMM2_Ex7F	CACCTTTTGCCTTTGTGTGC (55.5)	205
	PMM2_Ex7R	CCATCAAGCGCAAATGCAAC (55.9)	
Exon 8	PMM2_Ex8F	CCAGGGTCACATCAGCAATG (56.2)	260
	PMM2_Ex8R	TGAGCACGTGTGGGAGGACC (61.9)	

21. Prediction of splice-site, enhancer or silencers, and binding of splicing factors in the *PMM2* gene

Splice scores for the natural and cryptic donor and acceptor splice-sites were determined using BDGP software (www.fruitfly.org/seq_tools/splice.html). Predictions of the presence of exonic splice enhancer or silencer sequences were made using ESEfinder software (http://rulai.cshl.edu/cgi-bin/tools/ESE3) [96], Rescue-ESE [97] and PESX [98] were used to analyze the binding of several splicing factors.

22. Genotyping of IVS4 +21 G>C and IVS2 -13 G>A *PMM2* mutations

Genotyping of two genetic changes identified in *PMM2* intronic regions in 2 CDG-like patients (IVS4 +21 G>C and IVS2 -13 G>A) was done in 61 and 47 healthy subjects by PCR-allelic specific restriction assay with MnlI enzyme (Fermentas) and by sequencing, respectively.

23. Platelet purification and proteomic analysis

Blood from one CDG-like patient (P15), one PMM2-CDG patient and four healthy controls was anticoagulated with 3.8% (w/v) trisodium citrate (9:1) and platelet rich plasma (PRP) was obtained after centrifugation at 150 g for 15 min. Analysis of platelet glycoproteins GPIbα, GPIIIa, von Willebrand Factor (vWF), thrombospondin and platelet endothelial aggregation receptor 1 (PEAR1) was done by western blot using specific home-made antibodies generated in the laboratory of Dr. K. Freson (University of Leuven, Belgium), as described before. For proteomic analysis forty mL of blood were collected into vacutainer ACD tubes (Becton Dickinson, Helsingborg, Sweden) and platelet rich plasma (PRP) was obtained after centrifugation at 150 g for 15 min.

For proteomic analysis, PRP was transferred to a plastic tube containing PGE1 (1 μM final concentration), diluted with an equal volume of ACD pH 6.5 (7 mM citric acid, 93 mM sodium citrate, and 140 mM dextrose) and centrifuged at 150 g for 10 min to remove contaminant erythrocytes. The supernatant was further centrifuged at 720 g for 10 min and the platelet pellet was washed twice in JNL buffer (130 mM NaCl, 10 mM trisodium citrate, 9 mM $NaHCO_3$, 6 mM dextrose, 0.9 mM MgCl2, 0.81 mM KH_2PO_4, and 10 mM Tris, pH 7.4).Glycoproteins from platelet pellets were obtained by lectin affinity chromatography, using the total glycoprotein spin columns (Qproteome® Total Glycoprotein Kit, Qiagen). This approach is based on the selective binding capacity to specific glycan moieties of

specific lectins, namely concanavalin A (ConA) and Wheat germ agglutinin (WGA). Glycoproteins were then precipitated in thricloroacetic acid in acetone and solubilised in DIGE buffer (7M urea, 2M thiourea, 4% CHAPS, 30mM Tris, pH8.0).For DIGE analysis, glycoprotein samples were labelled as previously described [99,100]. Briefly, equal quantities of platelet proteins from the patients and healthy controls (50 μg for each sample) were labelled with 400 pmol of Cy3 or Cy5, whereas the pooled internal standard was labelled with Cy2 (GE Healthcare, Munich, Germany). The samples from any condition were never labelled with all Cy3 or Cy5 to control any specific dye labelling artefacts that might occur. The internal standard consisted of a pool of all the samples of the experiments. Different pairs of Cy3- and Cy5-labeled samples (each containing 50 μg of protein) were combined and mixed with a 50-μg aliquot of the Cy2-labeled pooled standard. These mixtures were diluted 1:1 with lyses buffer containing 0.5% IPG buffer (pH 3–10) and 1.2% Destreak and applied by cup loading on IPG strips (NL pH 3–11, 24 cm). The first dimension was carried out in an IPGphor system (GE Healthcare) using the following conditions: 1.5 h at 150 V, 1 h at 500 V in gradient, 2 h at 1,000 V in gradient, 3 h at 8,000 V in gradient and 8 h at 8,000 V. IPG strips were next incubated in equilibration buffer (6 M urea, 30% glycerol, 2% SDS, and 50 mM Tris–HCl pH 8.8) with 65 mM DTT for 20 min and in equilibration buffer with 200 mM iodoacetamide and 0.02% bromophenol blue for an additional 20 min. The second dimensional separations were carried out on 10% SDS-polyacrylamide gels on the EttanDaltSix system (GE Healthcare).Labelled proteins were visualized using the Typhoon Trio imager (GE Healthcare). The Cy2, Cy3, and Cy5 components of each gel were individually imaged using excitation/emission wavelengths specific for Cy2 (488/520 nm), Cy3 (532/580 nm) and Cy5 (633/670 nm). Gel analysis was performed using DeCyder 2D Differential Analysis Software v6.5 (GE Healthcare).Gels used to pick spots of interest were visualized with silver staining [101]. The gel pieces were destained [102] and digested by trypsin (Promega, Madison, WI). Upon concentrating and desalting tryptic fragments using Millipore C18 ZipTips (Millipore, Bedford, MA, USA), samples were mixed in a 1:1 ratio with α-cyano-4-hydroxycinnamic acid matrix (saturated solution in 50% ACN and 2.5% TFA) and spotted onto the target plate. MS/MS analyses were performed on a 4800 MALDI-TOF/ TOF instrument (Applied Biosystems, Foster City, CA). The instrument was calibrated with Applied Biosystems Calibration Mixture 1. Measurements were taken in the positive ion mode between 900 and 3,000 m/z. Sequences were automatically acquired by scanning first in MS mode and selecting the 15 most intense ions for MS/MS using an exclusion list of peaks arising from tryptic autodigestion. Data

interpretation was carried out using the GPS Explorer software (Version 3.5), and database searching was carried out using the Mascot program by using MSDB database. MS/MS searches were conducted with the following settings: MS/MS tolerance for precursor and fragment ions of 1 Da, carbamidomethylation of cysteine and methionine oxidation as fixed and variable modification, respectively. Trypsin was selected as enzyme and a maximum of one missed cleavage was allowed. Using these parameters the probability-based MOWSE (Molecular Weight Search) scores greater than the given cutoff value for MS/MS fragmentation data were taken as significant ($p< 0.05$).

24. Treatment of HepG2 cells with proton and Ca^{2+} pump inhibitors and a pH neutralizing agent

Human hepatocellular liver carcinoma cells (HepG2) were treated with 300 nM bafilomycin A1, a proton-pump chemical inhibitor, 25 mM ammonium persulphate or 50 μM cyclopiazonic acid (CPA), a strong Ca^{2+} pump inhibitor for 24 hours. Intracellular and secreted antithrombin was evaluated by immunofluorescence and western blot as described before.

25. Epidemiologic studies of CDG-like disorders

The epidemiologic studies aiming to determine the prevalence and potential thrombotic relevance of CDG-like disorders and the *PMM2* p.Arg141His mutation was possible by a multicentre collaboration, and included 1619 patients with deep venous thrombosis (DVT); 1049 from different Spanish centers (Barcelona, Valencia and Murcia [87,103,104]) and 570 from Denmark [105]; 791 Spanish patients with Acute Myocardial Infarction (AMI) [106]. Our study also enrolled 2030 Spanish and 1966 Danish controls.

p.Arg141His genotyping was performed by using a custom SNP Genotyping Assay (Sigma-Aldrich, Madrid) using Locked Nucleic Acids® (LNA®) which enhance sensitivity and specificity for the assay. LNA is a novel type of nucleic acid analog containing a 2'-O, 4'-C methylene bridge. PMM2_R141Hwildtype: [Flc]ACTCA[+A]TG[+C]GT[+T]CT[+T]CTT[TAM]: FLC 5' fluorophore and TAMRA quencher; PMM2_R141H mutated: [Rox]ACTCA[+A]TG[+T]GT[+T]CT[+T]CTT[BHQ2]: Rox fluorophore and BHQ2 quencher. Between brackets the modified LNA nucleotides.

Electrophoretic western blot analyses of plasma samples were carried out to detect smaller forms of antithrombin and α1-antitrypsin, as described before.

RESULTS

Objective 1

1. Anti-FXa activity variability in healthy population

The anti-FXa activity observed in 307 Spanish Caucasian healthy blood donors had a normal distribution with a mean value of 97±8% (Figure 15). None of the subjects had antithrombin deficiency (<70% of the reference value).

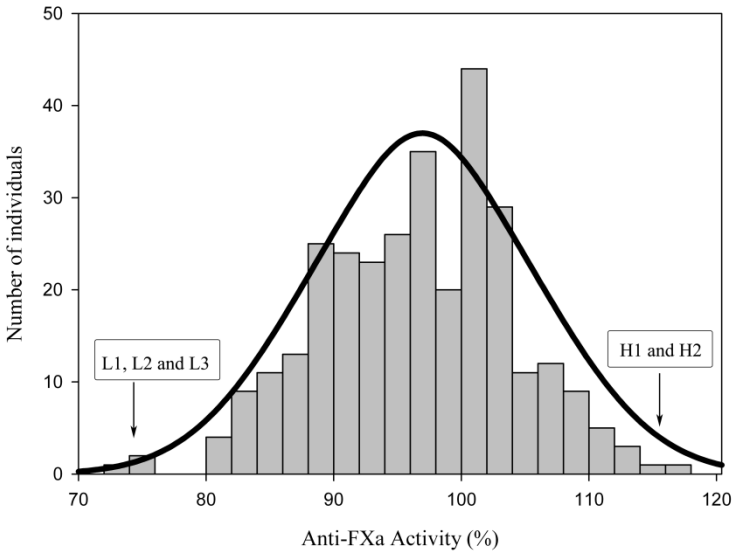


Figure 15. Distribution of anti-FXa activity in 307 Spanish blood donors. Subjects with low (L) or high (H) antithrombin levels selected for *SERPINC1* sequencing are also indicated.

2. Genetic variability of *SERPINC1* in healthy subjects with extreme anti-FXa activity, and validation of potential functional consequences

We selected five subjects with extreme antithrombin levels within the normal range: three with the lowest levels (L1: 73.5%, L2: 74.5% and L3: 75.1%) and two with highest levels (H1: 114.4% and H2: 116.1%) (Figure 15). The anti-FXa activity of these subjects did not significantly change in at least two independent determinations with different samples. We sequenced 15,375 bp of the *SERPINC1* gene, including the potential promoter, exons, introns and the 3' region (Figure 16A). Comparison with the reference sequence revealed 18 polymorphisms and a new g.2085 T>C genetic change (Figure 16 and Table 6). This substitution, identified in a subject with high antithrombin levels (H1), was located in the potential promoter region, 228 bp upstream from the translation initiation codon (Figure 16A). Interestingly, the g.2085 T>C transition could have potential transcriptional consequences, as it eliminates GATA-1 and c-Myb binding sites and creates a new site for AP-1 (Figure 16B).

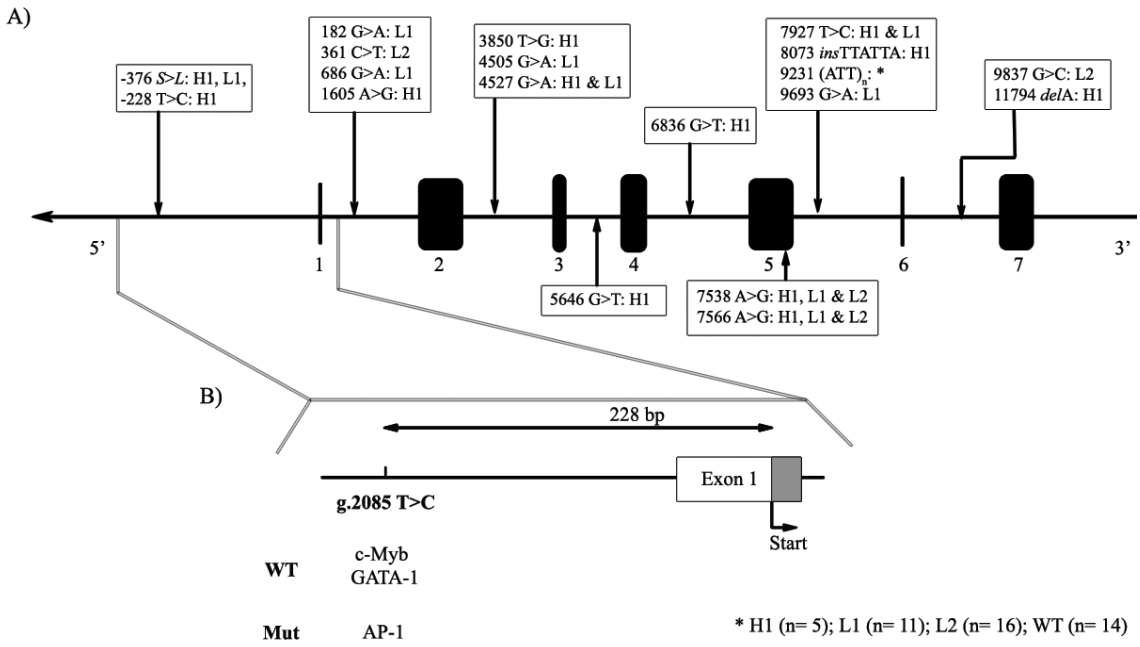


Figure 16. Genetic variations identified in the *SERPINC1* gene. A) Localization in the *SERPINC1* locus of genetic variations identified in this study. The subjects carrying these variations are also indicated. *S*: small nucleotide series; *L*: large nucleotide series: B) Potential transcriptional binding sites affected by the g.2085 T>C modification located -228 bp from initiation codon. Numbering is based on genomic DNA sequence, nucleotide +1 corresponds to the A of the ATG initiation codon of the *Homo sapiens* antithrombin gene (GenBank accession number NG_012462.1).

The genotyping of this modification in 131 healthy subjects by HRM/Razor probe method (Figure 14) confirmed that this genetic change is a polymorphism. Actually, during this study, this change was already described as a SNP in the HapMap (rs78972925). We identified the g.2085 T>C transition in heterozygous state in seven out of 131 subjects (5%). However, this polymorphism only showed a trend to be associated with high anti-FXa activity (99±11% in carriers vs. 96±9% in non-carriers; p= 0.375) (Table 7).

Table 6. Polymorphisms identified in healthy blood donors with low (L1: 73.5%; L2: 74.5% and L3: 75.1%) and high (H1: 114.4% and H2: 116.1%) antithrombin levels. ins: insertion; I: Intron; E: exon; WT: Wild-Type allele; ND: not determined; S: small nucleotide series; L: large nucleotide series.

Polymorphism	Change	Location	Position	H1	H2	L1	L2	L3
g.2085 T>C (rs78972925)	T>C	5'	-228	T/C	T/T	T/T	T/T	T/T
rs3138521	S>L	5'	-376	S/L	S/S	S/L	ND	S/S
rs2227589	G>A	I 1	182	G/G	G/G	G/A	G/G	G/G
rs2227590	C>T	I 1	361	C/C	C/C	C/C	T/C	C/C
rs2227592	G>A	I 1	686	G/G	G/G	G/A	G/G	G/G
rs2227596	A>G	I 1	1605	A/G	A/A	A/A	ND	A/A
rs2227603	T>G	I 2	3850	T/G	T/T	T/T	ND	T/T
rs941988	G>A	I 2	4505	G/G	G/G	G/A	ND	G/G
rs941989	G>A	I 2	4527	G/A	G/G	G/A	ND	G/G
rs2227607	G>T	I 3	5646	G/T	G/G	G/G	ND	G/G
rs2227609	G>T	I 4	6836	G/T	G/G	G/G	G/G	G/G
rs5877	A>G	E 5	7536	A/G	A/A	A/G	A/G	A/A
rs5878	A>G	E 5	7566	A/G	A/A	A/G	A/G	A/A
rs1799876	T>C	I 5	7927	C/C	T/T	C/C	ND	T/T
rs5778785	*ins*TTATTA	I 5	8073	ins/ins	WT	WT	WT	WT
rs57195053	$(ATT)_n$*	I 5	9231	14/5	14/14	14/11	14/16	14/14
rs2759328	G>A	I 5	9693	G/G	G/G	A/A	G/G	G/G
rs677	G>C	I 6	9837	G/G	G/G	G/G	G/C	G/G
rs2227616	*del*A	I 6	11794	A/-	A/A	A/A	A/A	A/A

Numbering is based on genomic DNA sequence, nucleotide +1 corresponds to the A of the ATG initiation codon of the *Homo sapiens* antithrombin gene (GenBank accession number NG_012462.1).
* The WT allele contains $(ATT)_{14}$, while the reported allelic variant contains $(ATT)_{13}$ in the intron 5 Alu sequence.

We also studied the potential functional effect of the other polymorphisms identified in this study by correlating genotype with anti-FXa activity in healthy subjects. This study was restricted to those polymorphisms identified only in individuals with high or low antithrombin levels. As shown in Table 7, only rs2227589, located in intron 1, was significantly associated with antithrombin levels. Carriers of the polymorphic allele had a moderately reduced anti-FXa activity in plasma [65].

Table 7. Plasma anti-FXa values of blood donors according to the genotype of *SERPINC1* polymorphisms identified in selected 5 healthy subjects with extreme antithrombin levels.

Polymorphism	Subject	Genotyping method	N	Frequency	Anti-FXa activity (mean±SD)		P°
					PGA European#	Our study	
g.2085 T>C (rs78972925)	H1	HRM/Razor probe&sequencing	131	No data	T:0.973 C:0.027	WT:96±9 Pol:99±11	0.375
rs2227589	L1	Taqman probe	305	G:0.975 A:0.043	G:0.898 A:0.102	WT:97±7 Pol:95±8	**0.037**
rs2227590	L2	PCR-ASRA: Fok I	128	C:0.870 T:0.130	C:0.922 T:0.078	WT:95±8 Pol:94±9	0.793
rs2227596	H1	PCR-ASRA: Tsp 45I	127	A:0.795 G:0.205	A:0.886 G:0.114	WT:94±8 Pol:97±9	0.129
rs2227603	H1	PCR-ASRA: Fok I	101	T:0.935 G:0.065	T:0.990 G:0.010	WT:94±8 Pol:100±19	0.317
rs941988	L1	PCR-ASRA: Hinf I	97	G:0.955 A:0.045	G:0.878 A:0.122	WT:97±7 Pol:95±8	0.513
rs2227607	H1	PCR-ASRA: Nla III	138	G:0.972 T:0.028	G:0.895 T:0.105	WT:95±8 Pol:95±9	0.903
rs5778785	H1	HRM&sequencing	106	No data	WT:0.896 INS:0.104	WT:95±8 Pol:95±9	0.838
rs2759328	L1	PCR-ASRA: Fau I	149	G:0.917 A:0.083	G:0.850 A:0.150	WT:95±8 Pol:94±8	0.736
rs677	L2	PCR-ASRA: Dde I	158	G:0.842 C:0.158	G:0.930 C:0.070	WT:95±8 Pol:94±9	0.604
rs2227616	H1	PCR-ASRA: Bst EII	183	A:0.935 (-):0.065	A:0.880 (-):0.120	WT:94±7 Pol:94±8	0.923

* N: Number of healthy blood donors studied.
# Frequencies from http://www.ncbi.nlm.nih.gov/snp
° p-value obtained from the student-t statistical analysis following a dominant model.
WT: Homozygous wild type genotype; Pol: Carriers of the polymorphism (heterozygous + homozygous).
SD: standard deviation.
rs941988 and rs2227592 were completely linked, according to the Haploview 4.1
The genotypes of all studied polymorphisms were in Hardy-Weinberg equilibrium.

Haplotype analysis revealed 7 frequent haplotypes, although none seemed to be significantly associated with anti-FXa activity (Figures 17 and 18).

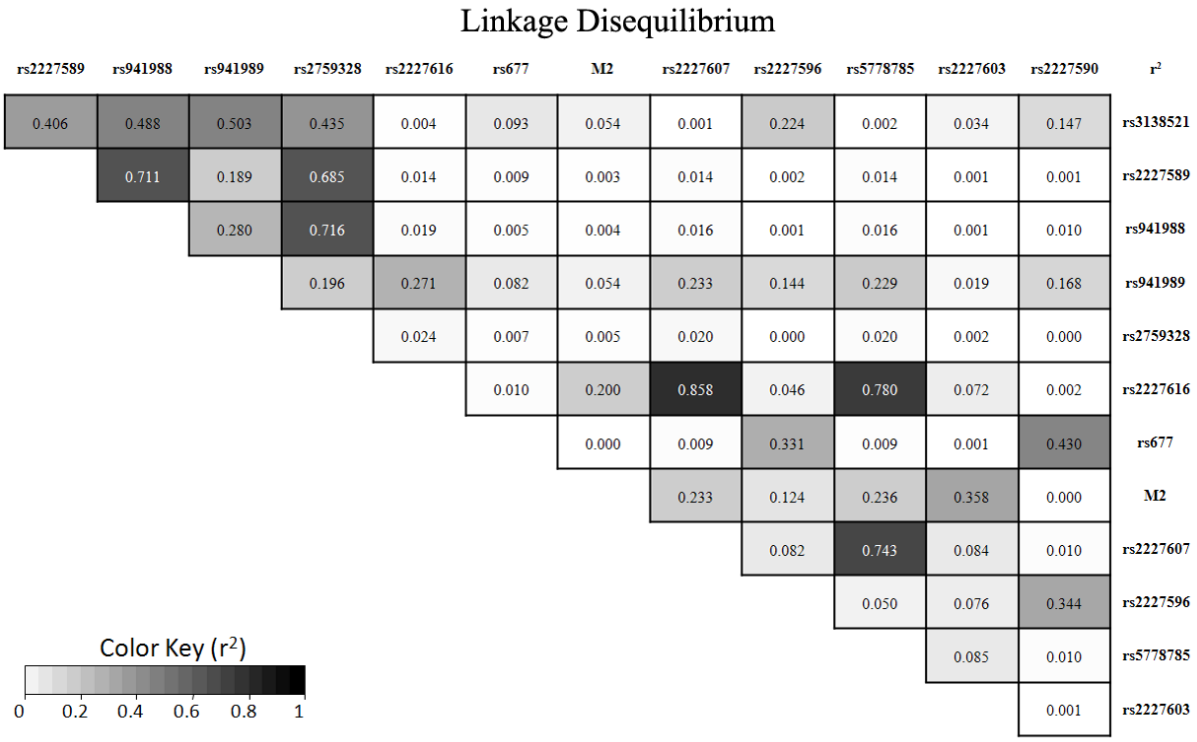


Figure 17. Linkage disequilibrium (r^2 values) between polymorphisms identified in our study. A gray scale is used to associate r^2 values.

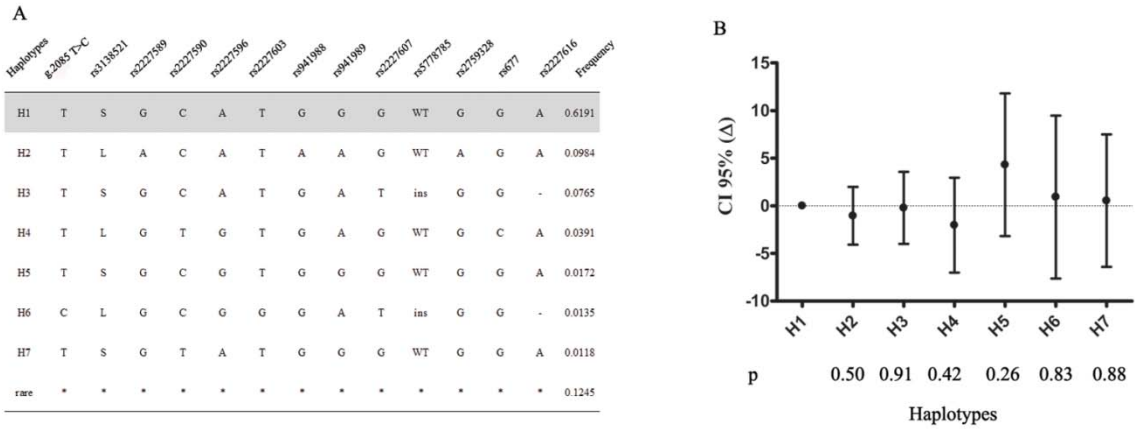


Figure 18. Haplotype analysis. A) Haplotypes identified. WT: Wild-type allele; ins: ATTATT insertion. B) Correlation between the haplotypes and anti-FXa activity.

Finally, our sequence analysis revealed three new alleles of the rs57195053. This polymorphism was originally described as a deletion of an ATT triplet from the sequence of 14 tandem triplets on intron 5. We detected alleles with 5, 11 and 16 triplets. This result together with the description of 11 new polymorphisms affecting this location (rs72125649, rs71708020, rs71799889, rs71117551, rs71563245, rs66670532, rs72352633, rs10598596, rs72304917, rs67446132 and rs57195053),

supports a single polymorphism with high allelic variability (at least from 5 to 16 ATT repeats).

3. Genetic analysis in 25 patients with moderate antithrombin deficiency

The clinical features and anti-FXa activity of these patients are shown in Table 8. In these patients we sequenced the coding region, flanking introns and the potential promoter region. We only detected three polymorphisms (Table 8). In addition to the rs3138521, whose potential functional effect was discarded in a previous study [65], five patients (P1, P7, P18, P20 and P23) carried two linked polymorphisms (rs2227588 and rs61827938) located in the potential promoter region. However, none of these polymorphisms modified potential binding sites for transcription factors. Moreover, the functional effect of these polymorphisms was discarded after genotyping in 126 blood donors (carriers: 93±9; non-carriers: 95±8; p: 0.342).

Table 8. Clinical and genetic features of patients with moderate antithrombin deficiency.

N	Sex	Thrombotic event	Age first event	Anti-FXa	Alterations
P1	Male	DVT	47	72	rs2227588(A/G)* rs61827938 (A/G)
P2	Male	Recurrent DVT	66	77	No
P3	Male	DVT	56	88	rs3138521 S/L
P4	Male	Recurrent DVT + PE	66	86	No
P5	Male	DVT	59	79	No
P6	Male	DVT	35	70	No
P7	Male	Recurrent arterial thrombosis	48	78	rs3138521(L/L) rs2227588(A/G) rs61827938 (A/G)
P8	Male	Venous thrombosis	42	74	rs3138521 S/L
P9	Male	Recurrent DVT	65	86	No
P10	Male	Recurrent DVT	67	81	No
P11	Male	Recurrent DVT	68	79	No
P12	Male	DVT	68	78	No
P13	Male	Recurrent DVT + PE		72	No
P14	Female	Mesenteric venous thrombosis	40	89	rs3138521(S/L)
P15	Male	DVT	45	89	No
P16	Male	DVT+Arterial thrombosis	47	89	No
P17	Female	DVT	60	80	No
P18	Female	Cerebral venous thrombosis	27	88	rs61827138 C/T; rs3138521 S/L; rs2227588 A/G; **g.1091C>T**
P19	Male	PE	74	73	No
P20	Male	Myocardial acute infraction	36	88	rs61827938 C/T; rs3138521 S/L; rs2227588 A/G;
P21	Male	Recurrent DVT	48	81	rs3138521 S/L
P22	Male	Arterial thrombosis	34	70	No
P23	Female	Pulmonary thrombosis	24	89	rs2227588 A/G; rs61827938 C/T
P24	Male	DVT	74	84	No
P25	Male	Stroke	88	76	rs3138521(S/L)

DVT: Deep venous thrombosis. PE: Pulmonary embolism.

rs3138521: S: small nucleotide series; L: large nucleotide series.
* rs61827938 and rs2227588 are in complete disequilibrium linkage.

Interestingly in one patient (P18) we identified a new genetic modification in heterozygous state not previously described, g.1091 C>T. The presence of this genetic modification was validated by PCR-allelic specific restriction assay using Rsa I enzyme 24h at 37 ºC. This alteration is located 1,052 bp upstream the translation initiation codon and close to rs2227588 (only 7 bp away). Further studies are required to clarify whether these two modifications affect the same allele, although all rs2227588 carriers identified in this study did not present the g.1091 C>T mutation (Table 8). Finally, *in silico* studies predicted that the g.1091 C>T transition abolished a potential upstream stimulatory factor 1 (USF-1) binding site using both prediction software (TESS and TFSEARCH), independently of the presence or absence of rs2227588.

4. *SERPINC1* analysis in subjects with classical antithrombin deficiency carrying neither gross gene deletion nor mutation affecting exons or flanking regions

Patients P26 and P27, with deep venous thrombosis and no familial antithrombin deficiency, did not present any genetic alteration in *SERPINC1* gene including 1,500 bp from the potential promoter region and 1,000 bp from the 3' region. Additional clinical details, biochemical studies and thrombophilia characterization of P27 patient are the purpose of the third objective.

Family 1 (F1) had severe family history of thrombosis and a type I antithrombin deficiency. Three sisters with DVT and recurrent events showed 50-56% of functional antithrombin activity and similar antigen levels. MLPA analysis and sequencing of both strands of exons and flanking regions were done in two affected members of F1, but no gross deletion was discovered. Moreover, we did not identify mutations in the rest of the *SERPINC1* gene, including 1,500 bp from the potential promoter region and 1,000 bp from 3' region.

In contrast, three members of family 2 had a moderate type I antithrombin deficiency, although no one has developed thrombotic events (Figure 19). The single relevant genetic change identified in the proband of family 2 was a point mutation g.2143 C>G, located 170 bp upstream from the translation initiation codon. *In silico* prediction suggested that this change could eliminate a potential binding site for the leader-binding protein-1 (LBP-1) regulatory factor (Figure 19). Then, we genotyped this

transversion in 103 healthy blood donors by PCR-allelic specific restriction assay with Bsa JI, but no one carried it, supporting that this change is not a common polymorphism. In contrast, the mutation was present in three members of family 2 with low antithrombin levels (Figure 19B).

In order to further sustain the functional relevance of the g.2143 C>G transversion, luciferase reporter assays were performed. The plasmid containing the mutated promoter was associated with 30-50% reduction of luciferase activity in two different cell lines (HepG2 and PLC-PRF-5) (Figure 19C).

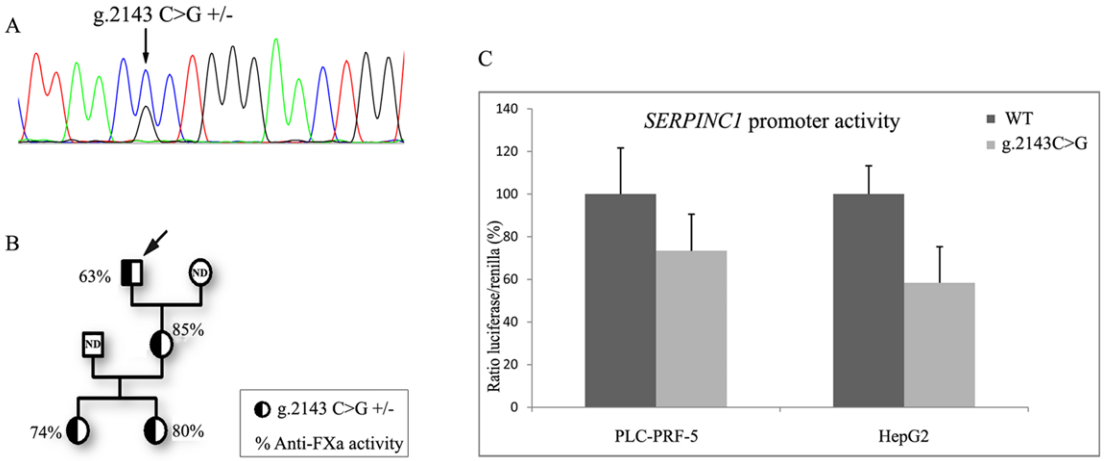


Figure 19. Identification and functional characterization of the g.2143 C>G modification. A) Electropherogram of the mutation identified in family 2. B) Family 2 pedigree. Proband is pointed by an arrow. Carriers of the g.2143 C>G modification are represented with the semi filled symbol. ND: not determined. Plasma anti-FXa activities (%) are also shown. C) Luciferase reporter assay of *SERPINC1* promoter activity with the wild-type and mutated g.2143G allele in HepG2 and PLC-PRF-5 cell lines. Results were obtained from 5 independent experiments.

Objective 2

The results of the previous objective have provided quite clear evidences about the low genetic variability in *SERPINC1* gene and the minor functional consequences to antithrombin levels. These findings encouraged the search for other potential modulating genes of plasma antithrombin levels by using a GWAS approach, aiming to find associations between genotypes and plasma anti-FXa activity.

1. **GWAS analysis. Association between genotype and plasma anti-FXa activity in the GAIT study**

The plasma anti-FXa activity in the GAIT study had a normal distribution, with a medium value of 109.05% of the reference plasma and 154% and 78% as extreme values. Results of the GWAS are summarized in figure 20. No SNP was found associated with plasma anti-FXa activity at a genome-wide significance level (figure 20). For validation analysis we selected the 10 SNPs with the strongest association (p< 4x10E-05). Interestingly, 2 additional polymorphisms affecting *LARGE*, one of the genes identified, also showed significant association with anti-FXa activity, and were also selected to be validated (rs762057 and rs240082). Two of the *LARGE* SNPs (rs713703 and rs762057) displayed high linkage disequilibrium (D^2= 0.81). Table 9 displays the list of these polymorphisms, showing the p-value for the association with anti-FXa activity, the chromosomal location, and the gene potentially affected.

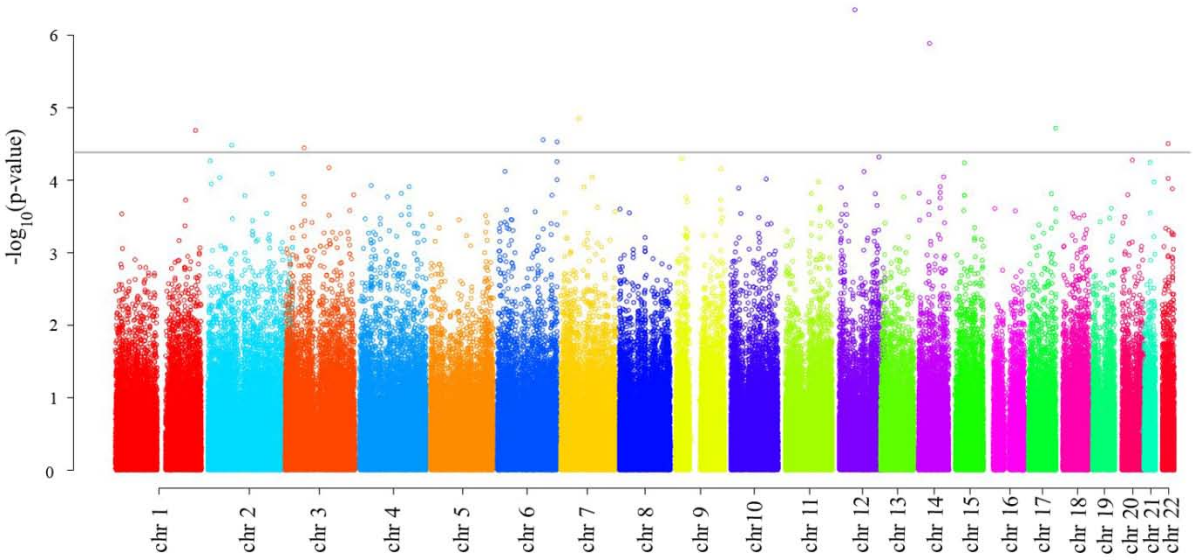


Figure 20: Manhattan plot GWAS with antithrombin phenotype. The threshold of significance to select candidate SNPs for validation is also shown.

Table 9. Single nucleotide polymorphisms (SNPs) associated with anti-FXa activity in the GWAS of the GAIT study and that were selected for validation studies.

SNPS	PVAL	BSNP	V-EXPL	CHR	LOC	GENE
rs10880942	0.00000044	-0.734718	0.081805	12	44905397	*SLC38A1*
rs2356895	0.00000128	-0.454476	0.088341	14	50898162	*LOC283553*
rs1860867	0.0000141	0.741308	0.050958	7	51762794	*COBL*
rs9896932	0.0000191	-1.361577	0.028271	17	77848326	*CD7*
rs1411771	0.0000206	0.361303	0.063374	1	230241398	*DISC1*
rs2152192	0.0000277	0.765150	0.061933	6	130510819	*SAMD3*
rs13193455	0.0000295	0.362923	0.029200	6	170354404	*LOC154449*
rs713703	0.0000313	0.330586	0.058187	22	32286080	*LARGE*
rs11681944	0.0000328	-0.338764	0.062070	2	68176083	*C1D*
rs6768189	0.0000357	0.429999	0.051721	3	54446050	*CACNA2D3*
rs762057	0.0000944	0.310900	0.052599	22	32280513	*LARGE*
rs240082	0.0010970	0.707098	0.039564	22	32428417	*LARGE*

2. Validation study

The 10 SNPs that showed stronger statistical association with anti-FXa in the GWAS, as well as the 2 additional *LARGE* SNPs were genotyped in 307 blood donors from another Spanish region used as a validation cohort. Only rs762057 maintained a significant association with anti-FXa levels in the validation cohort (p= 0.047) (Table 10). Multivariate analysis including age, gender and rs2227589, a SNP in *SERPINC1* gene previously reported to be associated with plasma anti-FXa activity [65], increased the significance between rs762057 and anti-FXa activity (p= 0.020) (Table 10).

Table 10. Genotype-phenotype analysis in the validation study.

SNP	Gene	Genotype	Anti-FXa	Crude p	Adjusted p
1. rs10880942	*SLC38A1*	T/T	96.3±7.4		
		T/C+C/C	97.1±8.1	0.444	0.414
2. rs2356895	*LOC283553*	T/T	96.2±7.4		
		T/C+C/C	96.8±7.8	0.510	0.240
3. rs1860867	*COBL*	G/G	96.7±7.4		
		G/A+A/A	95.1±8.4	0.178	0.240
4. rs9896932	*CD7*	A/A	96.3±7.5		
		A/G+G/G	96.7±8.3	0.800	0.870
5. rs1411771	*DISC1*	T/T	95.7±7.9		
		T/C+C/C	97.0±7.2	0.122	0.078
6. rs2152192	*SAMD3*	G/G	96.4±7.6		
		G/A+A/A	96.2±7.4	0.839	0.998
7. rs13193455	*LOC154449*	G/G	96.2±7.5		
		G/A+A/A	96.6±7.6	0.651	0.705
8. rs762057	*LARGE*	G/G	95.3±6.9		
		G/A+A/A	97.1±7.7	**0.047**	**0.020**
9. rs11681944	*C1D*	G/G	96.3±8.0		
		G/A+A/A	96.6±7.1	0.730	0.655
10. rs6768189	*CACNA2D3*	G/G	96.8±7.8		
		G/A+A/A	95.6±6.9	0.158	0.198
11. rs713703	*LARGE*	T/T	95.2±7.2		
		T/C + C/C	96.8±7.5	0.082	0.087
12. rs240082	*LARGE*	A/A	96.4±7.7		
		A/G + G/G	96.0±6.9	0.800	0.733

Finally, *LARGE* haplotype analysis in the validation cohort revealed 5 frequent haplotypes, one of them (H2) significantly associated with anti-FXa activity (p= 0.030) (Table 11).

Table 11: *LARGE* haplotypes identified in the validation study and their correlation with anti-FXa activity.

Haplotype	rs762057	rs713703	rs240082	Frequency	Difference (95% CI)	P
1	G	T	A	0.4523	0.00	---
2	A	C	A	0.403	**1.4 (0.12-2.68)**	**0.033**
3	G	C	A	0.0666	2.24 (-0.34-4.81)	0.09
4	A	T	A	0.0399	1.8 (-1.36-4.95)	0.27
5	G	T	G	0.0258	1.18 (-3.17-5.52)	0.6
rare	*	*	*	0.0124	-1.42 (-8.16-5.31)	0.68

Consistently in both cohorts, the polymorphic A allele associated with a moderate higher plasma anti-FXa activity (2-4%) (Figure 21).

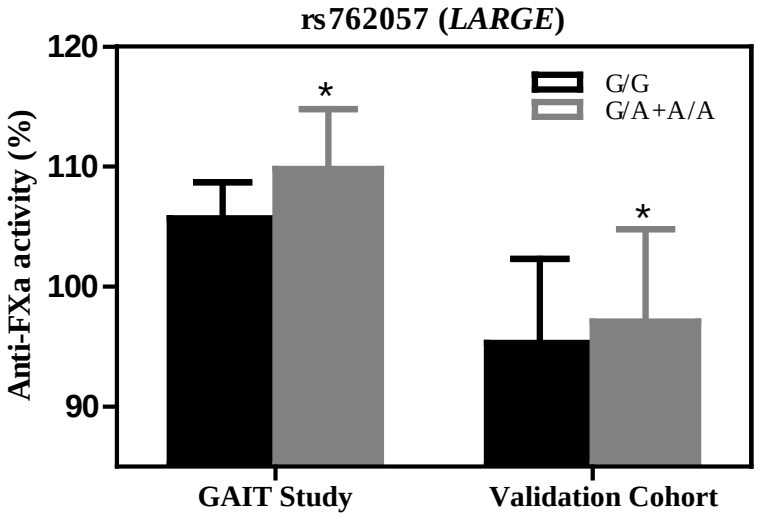


Figure 21: Plasma anti-FXa activity according to the rs762057 genotype in the GAIT and validation studies. *p<0.05

3. Thrombotic relevance of the *LARGE* rs762057 polymorphism

Although the A allele only mildly correlated with higher antithrombin levels, the key relevance of this anticoagulant encouraged to evaluate whether or not the rs762057 polymorphism affected the risk of venous thrombosis in our case-control study. Genotyping was successful in 849 patients and 862 controls. The allele or genotype frequencies were similar in cases than in controls (Figure 22).

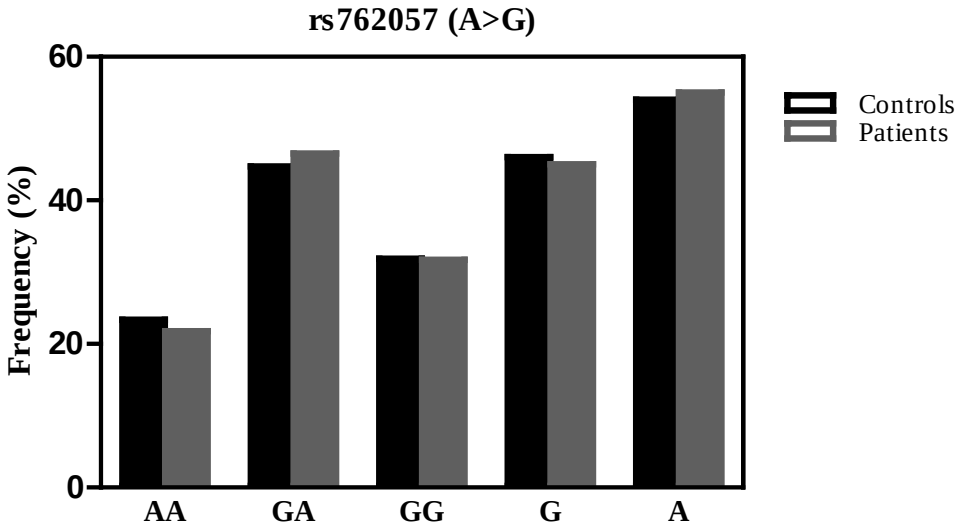


Figure 22. Genotype and allelic frequencies of rs762057 in the case/control study including 849 patients with venous thrombosis and 862 controls.

Actually, both the univariate and multivariate analysis including age, gender and genetic prothrombotic risk factors (FV Leiden, prothrombin G20210, antithrombin Cambridge II A384S and ZPI Arg67Stop) revealed that this polymorphism did not significantly modify the risk of venous thrombosis (p= 0.92 and p= 0.94, respectively).

4. Functional studies

In order to verify the potential role of *LARGE* as a modulating gene of antithrombin identified by genetic association studies, further functional studies were performed.

Firstly, as the human *LARGE* gene is ubiquitously expressed, with highest levels in heart, brain and skeletal muscle [107], we evaluated the expression of this gene in mononuclear cells from 10 healthy subjects. Interestingly, the expression of this gene in mononuclear cells showed a significant and positive correlation (r^2= 0.8) with plasma anti-FXa activity (Figure 23A).

Since *LARGE* codes an enzyme involved in post-translational glycosylation, and glycosylation of antithrombin plays a relevant role in the function of this serpin, particularly in the heparin affinity [24,77,108], our first hypothesis considered that differential expression or function of *LARGE* could result in distinct glycosylation of antithrombin. In order to verify this hypothesis, proteomic and glycomic studies were done with the main plasma antithrombin glycoform (α-antithrombin) purified from subjects with the highest and lowest *LARGE* expression. However, their molecular

masses were very similar, and glycomic studies showed fluctuations but not significant differences on the level or type of glucidic components (Figure 24). These results suggested that the association with anti-FXa activity might be explained by a quantitative defect rather than by qualitative differences caused by the differential *LARGE* expression. Actually, plasma antigen levels of antithrombin also significantly correlated with *LARGE* expression in these healthy subjects (Figure 23B).

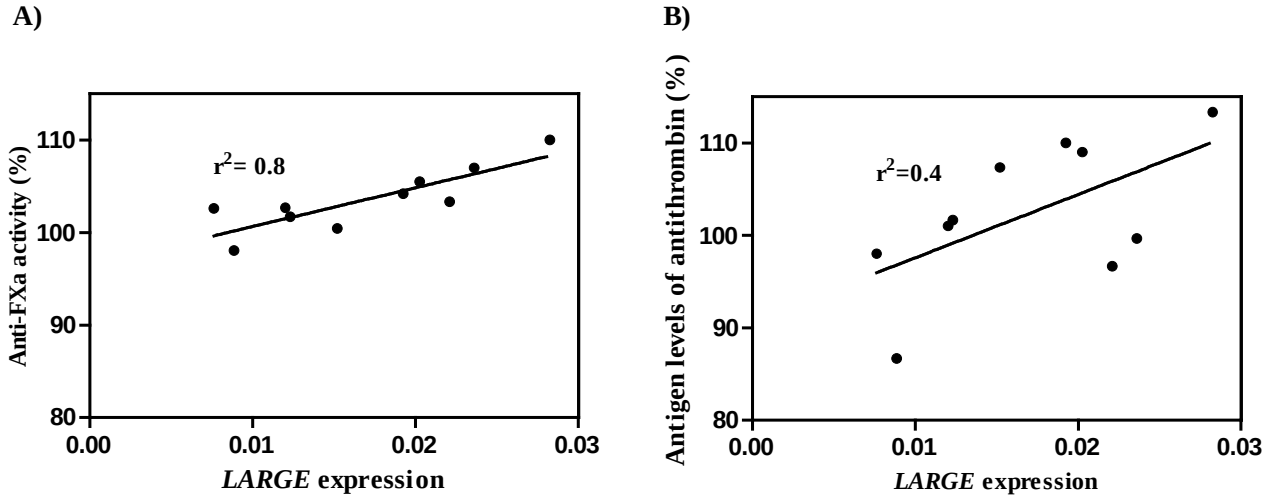
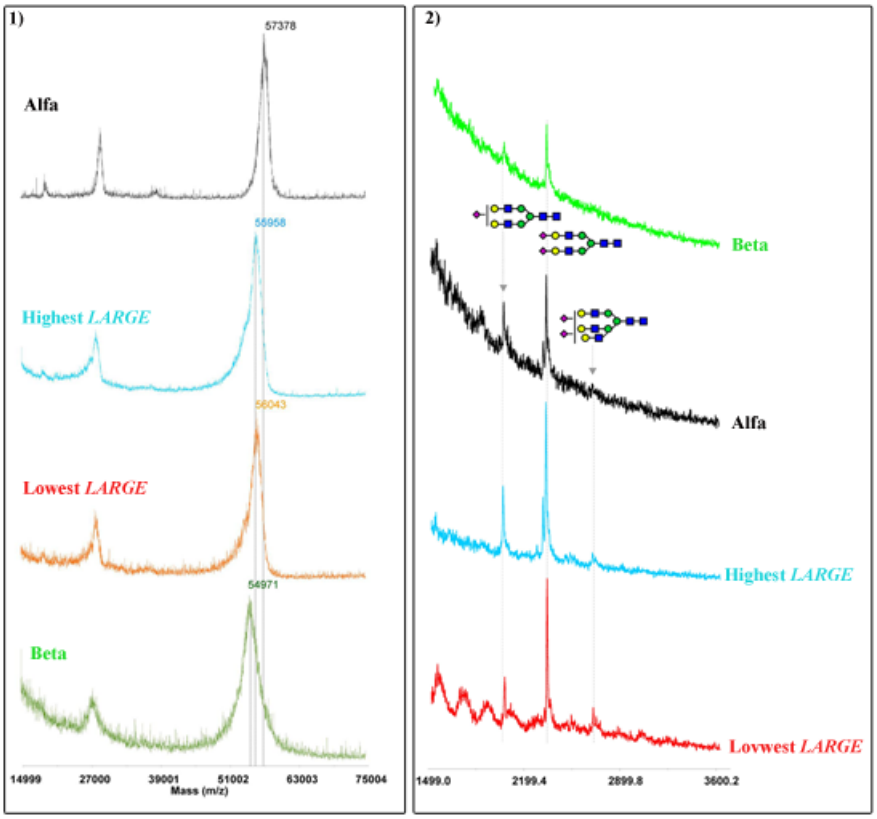


Figure 23: Correlation between mononuclear *LARGE* expression and plasma anti-FXa activity (A) and antigen levels (B) in 10 healthy subjects. LARGE expression was evaluated by specific transcription probe. Values are expressed as % of the reference value obtained in a pool of 100 healthy subjects**.**

A)



B)

1)

				others
Alfa	8.05 (45.361)	72.43 (408.331)	10.53 (59.345)	9.00 (50.723)
Highest *LARGE*	6.47 (49.124)	76.32 (579.135)	9.18 (69.626)	8.03 (60.942)
Lowest *LARGE*	11.12 (12.987)	72.51 (84.697)	8.76 (10.226)	7.61 (8.891)
Beta	10.50 (0.472)	38.16 (1.715)	0.00 (0.000)	51.34 (2.307)

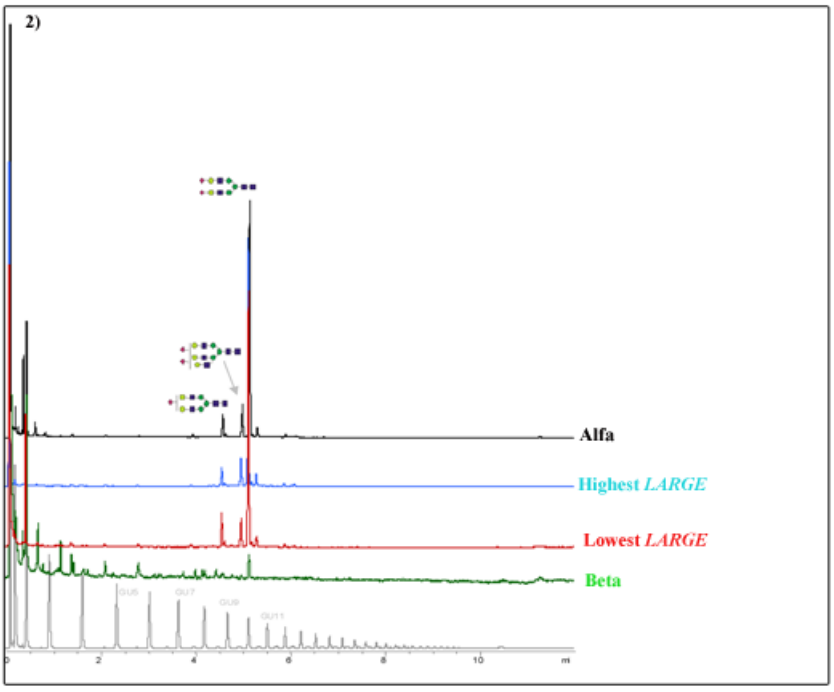


Figure 24. Glycomic and proteomic analysis of healthy subjects with the highest (blue) and lowest (red) *LARGE* expression. The main glycoform α (black), and the minor β □glycoform (green), both purified from a pool of 100 healthy blood donors are also shown. The β glycoform has 3 *N*-glycans since it lacks *N*-glycosylation at N-135. A) MALDI TOF mass spectrometric analysis of: 1) Intact glycoproteins; 2) 2AB-labeled *N*-glycans. B) HPLC data. 1) Distribution of the glycan structures of antithrombin specimens. Values are represented as % of total glycan pool. Between brackets are the absolute fluorescence units. 2) HILIC HPLC profiles of antithrombin specimens.

Unfortunately, these results can be questioned because of the weak expression of *LARGE* in mononuclear cells and the moderate differences found in healthy subjects with the highest and lowest *LARGE* expression (6.2-fold: 0.028 and 0.0045 units relatives to the expression of the constitutive gene, respectively).

To strongly sustain the relevance of *LARGE* on antithrombin levels, we carried out silencing experiments of *LARGE* in HepG2 and HEK-EBNA cell lines. According to our data, 60 hours after transfection with *LARGE* silencers, the expression of this gene was silenced around 30% in both cell lines (Figure 25). At this time, secretion of antithrombin to the conditioned medium in both cell lines treated with *LARGE* silencers was significantly reduced in silenced cells, 4-fold by western blot (Figure 26A) and 10-fold by ELISA (0.01 ± 0.01 mg/ml compared to 0.15 ± 0.20 mg/ml of control cells). The reduction was more significant in HEK-EBNA cells (Figure 26A). However, according to electrophoretic data, secreted antithrombin from silenced cells shows similar size to that of control cells (Figure 26A Interestingly, anti-FXa activity in the conditioned medium of *LARGE* silenced cells was 59 ± 30% and 11 ± 12% of that found in control cells transfected with the scramble siRNA or without siRNA in HepG2 and HEK-EBNA respectively. The reduction of antithrombin secretion paralleled with a moderate intracellular retention of this serpin according to the immunofluorescence and western blot results (Figure 26B).

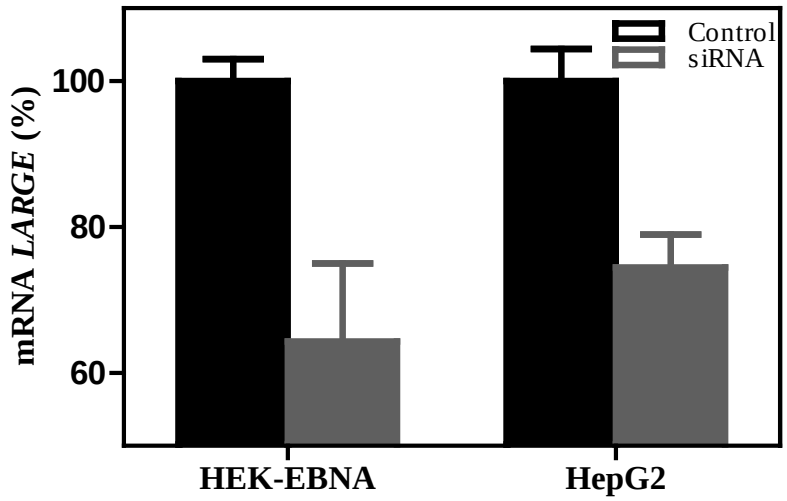


Figure 25: *LARGE* gene expression in HEK-EBNA and HepG2 cell lines 60 hours after transfection with *LARGE* siRNAs. Values are expressed as percentages respect to control cells (100%) (N= 3).

Moreover, aiming to determine the mechanisms underlying the modulation of antithrombin by *LARGE*, we measured *SERPINC1* expression in these cells. Silencing of *LARGE* did not significantly modify *SERPINC1* mRNA levels (Figure 26C).

In HepG2 cells, we also studied the effect of *LARGE* silencing on other *N*-glycoproteins: prothrombin, transferrin and other hepatic serpin: α1-antitrypsin. As shown in figure 26A, silencing of *LARGE* also reduced the secretion of all other proteins evaluated, although antithrombin seemed to be the most affected.

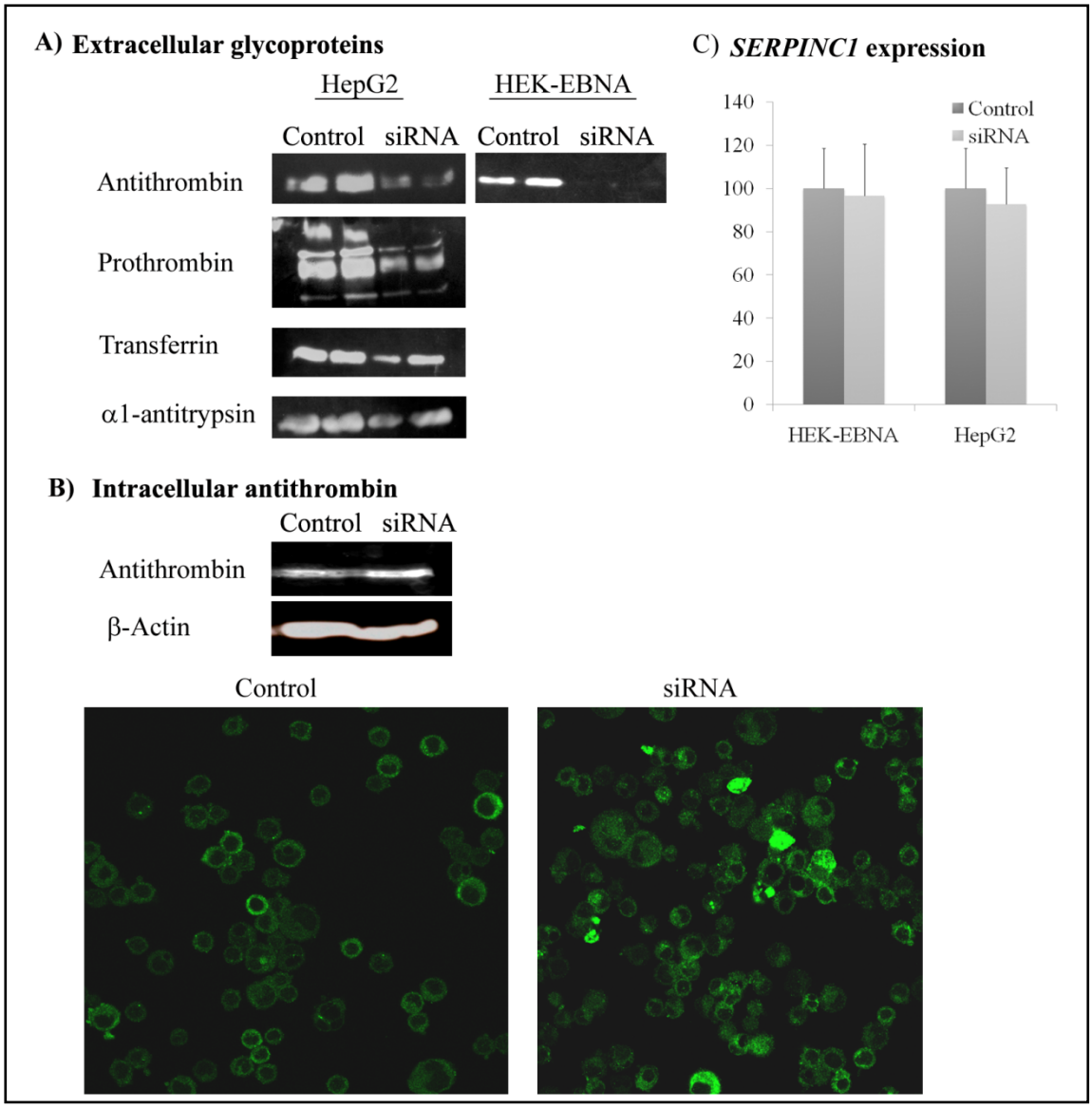


Figure 26. *LARGE* gene silencing in HepG2 and HEK-EBNA cell lines. A) Consequences on proteins secreted to the conditioned medium evaluated by immunoblotting. B) Effect on intracellular antithrombin from HepG2 cells analyzed by immunofluorescence and immunoblotting. C) Effect on the levels of *SERPINC1* expression in HEK-EBNA and HepG2 cell lines. Immunoblots and immunofluorescence figures are representative of at least 3 independent experiments.

Objective 3

The analysis of our large cohort of 89 patients with thrombosis and severe deficiency of antithrombin has identified the mutations or deletions responsible for antithrombin deficiency in 95% of these cases. However, the remaining 5% of patients with antithrombin deficiency and no mutation in the *SERPINC1* gene is an excellent group to identify new mechanisms responsible for the deficiency of this important anticoagulant.

1. Diagnosis of a congenital disorder of glycosylation (PMM2-CDG) in a patient with antithrombin deficiency and recurrent venous thrombosis

Our first study focused on three patients with thrombosis and severe antithrombin deficiency but no genetic defect on the *SERPINC1* gene. Patient P26 and F1 had type I antithrombin deficiency. Western blot analysis only confirmed the reduced levels of antithrombin in plasma, but did not reveal abnormal antithrombin forms.

Patient P27 is an 18 year old Portuguese male with recurrent left-leg venous thrombosis. The first event happened at 11 years old without any known concomitant risk factor, including fever or immobilization. Oral anticoagulation with warfarin was established for 5 months. The patient suffered a second spontaneous event in the same localization at 12 years old. Oral anticoagulation with warfarin was re-established. Unexpectedly, a new thrombotic event took place 6 months later under stable oral anticoagulant therapy with INR 2-3. A fourth thrombotic event was diagnosed at 17 years old, again under oral anticoagulant therapy with INR 2-3. Now, the patient is being treated with higher warfarin dosis (INR 2.5-3.5).

No familial history of thrombosis was reported. Thrombophilic analysis revealed no prothrombotic polymorphisms (Factor V Leiden or prothrombin G20210A). No antiphospholipid antibodies were detected. Protein C and protein S had values within the normal range. The single thrombophilic defect identified was an antithrombin deficiency (67% plasma anti-FXa activity *versus* 100% in a pool of plasma from 100 healthy blood donors). Antigen levels of antithrombin in plasma were also low (60%), suggesting a classical heterozygous type I antithrombin deficiency. However, the sequencing of the potential promoter, as well as the coding and flanking regions of the *SERPINC1* gene [109] revealed no significant mutation. Large gene deletions were excluded by MLPA studies. Interestingly, electrophoretic analysis of plasma

antithrombin performed on denaturing and native conditions revealed the presence of an abnormal antithrombin (Figure 27A). The reduced size of this abnormal band, observed in SDS gels under reducing conditions, and the absence of mutations in the coding region encouraged the search for a post-translational modification. We evaluated an impaired *N*-glycosylation, the only post-translational modification described in antithrombin that also affects the function and clearance of this anticoagulant serpin [108]. We treated plasma from the patient and a control with *N*-glycosidase F, as described before. As shown in figure 27B, *N*-glycan removal uniformed the size of plasma antithrombins in SDS gels sustaining an *N*-glycosylation defect for about 20% of antithrombin molecules in the plasma of the proband.

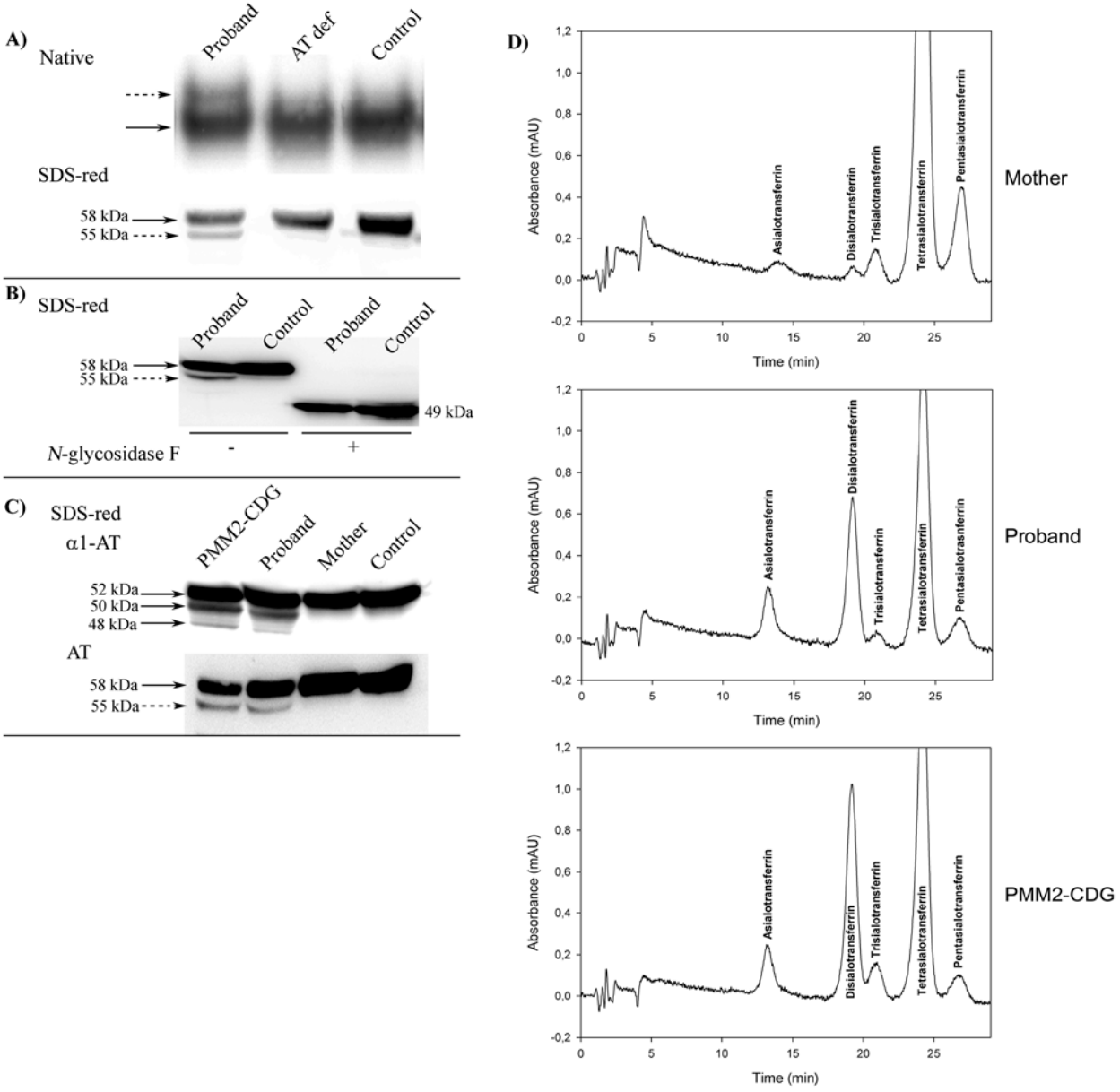


Figure 27. Identification of PMM2-CDG in the proband (P27). A) Native and SDS-PAGE under reducing conditions showing the normal (arrow) and abnormal (dashed arrow) antithrombin present in plasma of the proband, a patient with congenital type I antithrombin deficiency (AT def) (*SERPINC1* p.Arg117Stop) and a control subject. **B)** SDS-PAGE and western blot of antithrombin from plasma untreated (-) or treated (+) with *N*-glycosidase F. **C)** SDS-PAGE pattern of α1-antitrypsin (α1-AT) and antithrombin (AT) in the proband, his mother, a control and a PMM2-CDG patient (*PMM2* p.Phe119Leu and p.Arg141His). Arrows mark the abnormal glycoforms of these molecules. **D)** HPLC profile of transferrin glycoforms of the proband, his mother and a PMM2-CDG patient (*PMM2* p.Phe119Leu and p.Arg141His).

This abnormal antithrombin was not the beta-glycoform as it has different electrophoretic mobility in SDS-PAGE (Figure 28) and among other evidences, this defect was not specific of antithrombin as smaller forms of α1-antitrypsin were also detected in plasma of this patient (Figure 27C).

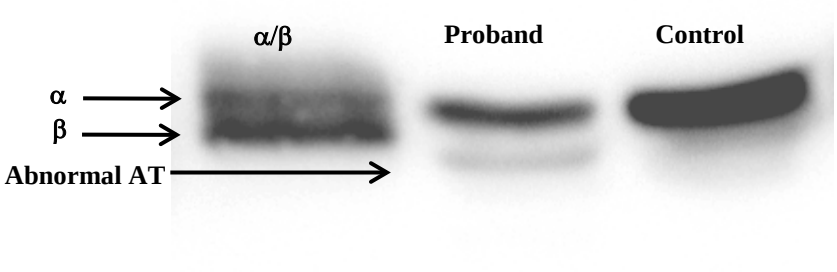


Figure 28. Electrophoretic pattern of plasma antithrombin from a control and the proband (P27) compared to that of a mixture of alpha and beta antithrombin (AT) purified from plasma of healthy subjects.

All these biochemical and clinical data focus the search to congenital disorders of glycosylation (CDG), a broad group of autosomal recessive disorders [110]. Actually, antithrombin and α1-antitrypsin electrophoretic profiles were very similar to that displayed by a patient with the most common CDG: PMM2-CDG, previously called CDG-Ia (Figure 27C). PMM2-CDG patients usually are compound heterozygous for mutations in *PMM2*, the gene encoding phosphomannomutase 2, a cytosolic enzyme that catalyzes the conversion of mannose 6-phoshate to mannose 1-phosphate, a substrate for GDP-mannose synthesis, required in *N*-glycosylation [110]. HPLC analysis of transferrin glycoforms pointed to a CDG-Ia diagnosis by the presence of asialotransferrin and a significant increase of disialotransferrin (Figure 27D). *PMM2* sequencing in the proband identified two mutations: one of maternal origin in exon 5,

and the other in exon 8, responsible for the p.Arg141His and p.Cys241Ser amino acid substitutions, respectively (Figure 29). The same combination has been reported previously in a PMM2-CDG patient with mild clinical phenotype [111]. In addition, the proband had strabismus, a thyroid nodule, kyphosis and moderate mental retardation without ataxia. The digestive, hepatic, renal and cardiovascular systems were normal.

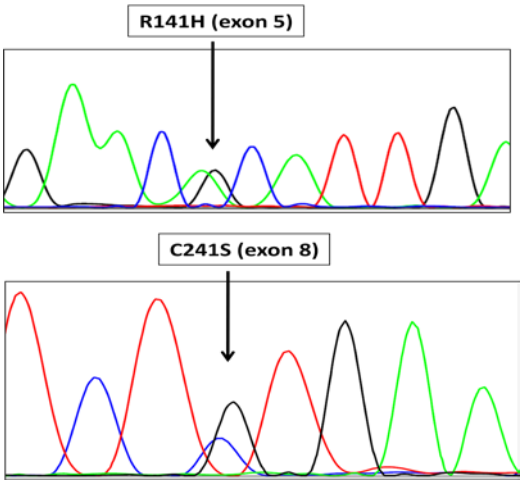


Figure 29. Electropherograms of mutations identified in the *PMM2* gene of the proband (P27).

2. Identification of a CDG-like disorder

The diagnosis of a PMM2-CDG disorder in a patient with antithrombin deficiency but without mutation in the *SERPINC1* gene encouraged the search for this o similar defect among the 25 patients with venous thrombosis and a moderate reduction of anti-FXa activity (70-90%, Table 8) but no mutations in *SERPINC1* gene. Thus, we studied their plasma antithrombin by electrophoretic analysis, including native and SDS gels under reducing and no reducing conditions.

Interestingly, four patients (P2, P15, P16, and P22) showed abnormal plasma antithrombin pattern in SDS-PAGE under reducing conditions (Figure 30A). They presented a smaller band not present in controls. The amount of this abnormal band never reached more than 15% of total antithrombin although varied in these patients, being P15 the patient with higher levels (Figure 30A). Purification by FPLC of this smaller antithrombin from P15, followed by mass spectrometry quantification revealed an abnormal antithrombin of 56.4 kDa. Treatment of plasma of this patient and a control with *N*-glycosidase F uniformed abnormal bands, suggesting again an *N*-glycosylation defect. Further analysis of plasma antithrombin from this patient by treatment with neuraminidase and galactosidase revealed incomplete sialic acid and galactose content

(Figure 30B). Then, we compared the electrophoretic pattern of these patients with a PMM2-CDG patient. As illustrated in figure 30C for P15, the pattern of antithrombin, α1-antitrypsin and transferrin glycoproteins was really similar. Further evidences strengthen the similarities of these patients with PMM2-CDG patients. Thus, HPLC and Q-TOF also confirmed the presence of abnormally glycosylated proteins in these patients. Transferrin glycoform analyses by HPLC revealed the presence of asialotransferrin and a significant increase of disialotransferrin in these patients (Figure 31). Moreover, Q-TOF analysis showed increase levels of aglycan and monoglycan structures of transferrin (Figure 32). Importantly, the electrophoretic and HPLC patterns of parents of PMM2-CDG patients with a single *PMM2* mutation displayed was similar than that of controls (Figure 27).

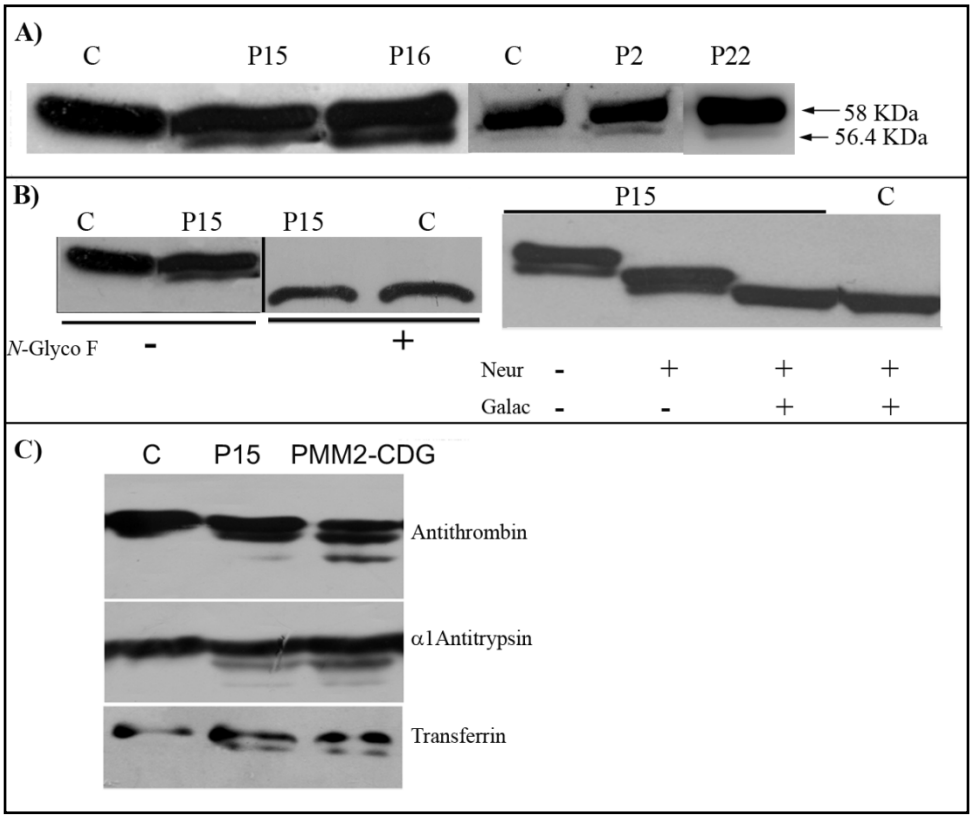


Figure 30. Electrophoresis analysis in SDS-PAGE under reducing conditions of patients with a CDG-like disorder (P15, P16, P2 and P22). A) Electrophoretic pattern of plasma antithrombin. B) Treatment of plasma antithrombin with *N*-glycosydase F (*N*-Glyco F), neuraminidase (Neu) and galactosidase (Galac). C) Electrophoretic pattern of antithrombin, α1antitrypsin and transferrin of P15 comparing with a control and a PMM2-CDG patient. C: Control.

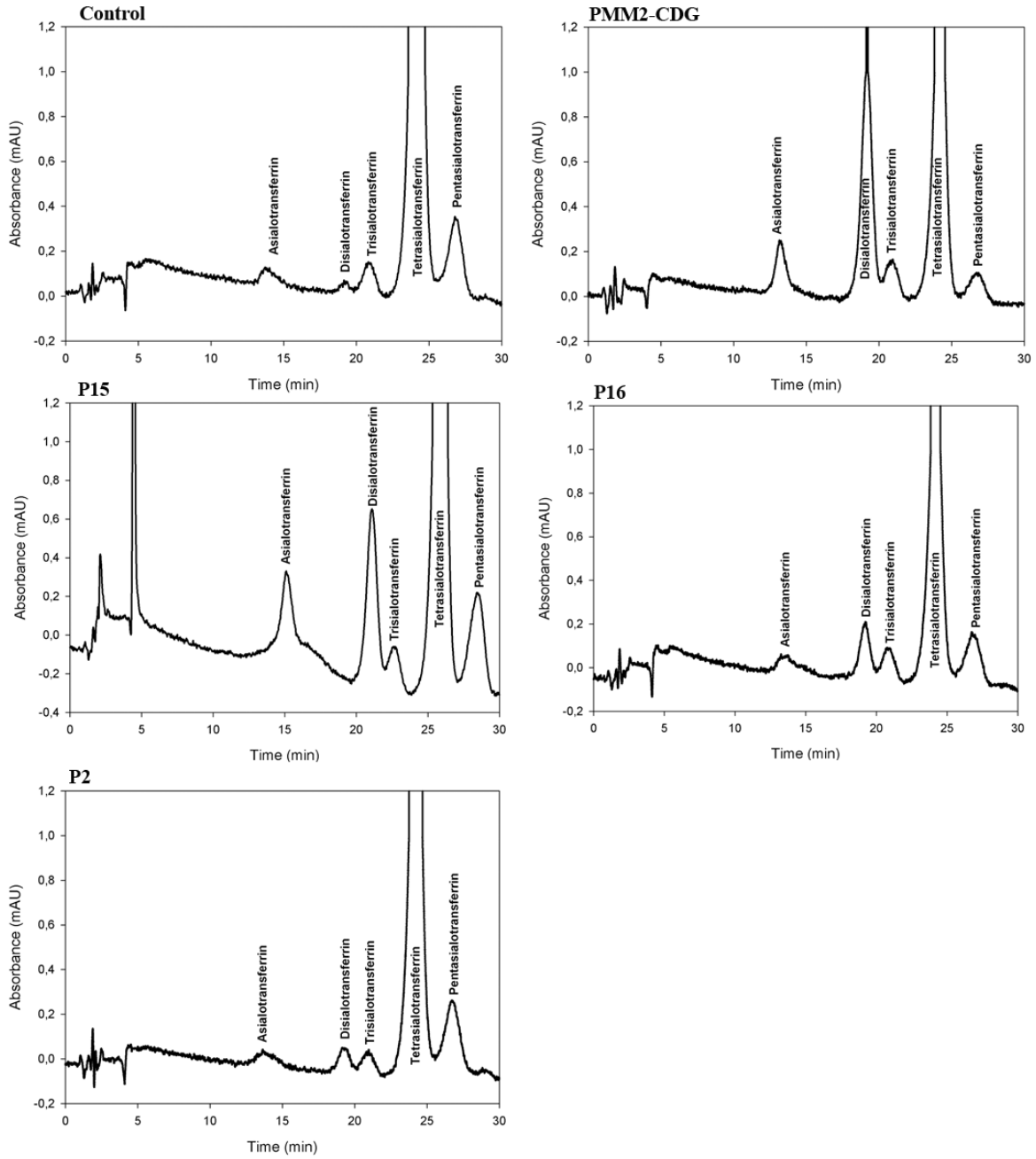
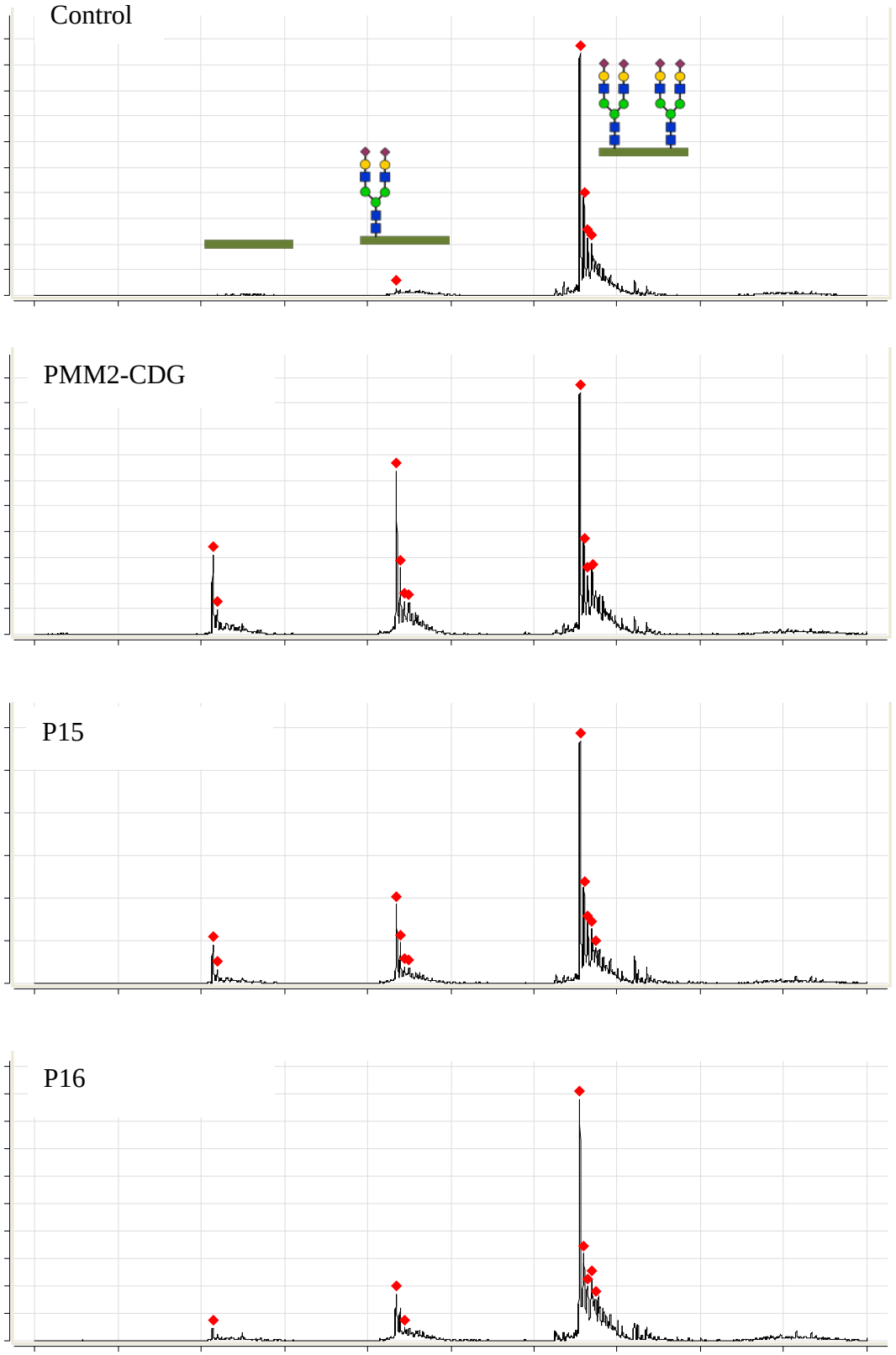


Figure 31. HPLC profiles of plasma transferrin glycoforms of CDG-like patients (P15, P16 and P2), a control and a PMM2-CDG patient.


Control
PMM2-CDG
P15
P16

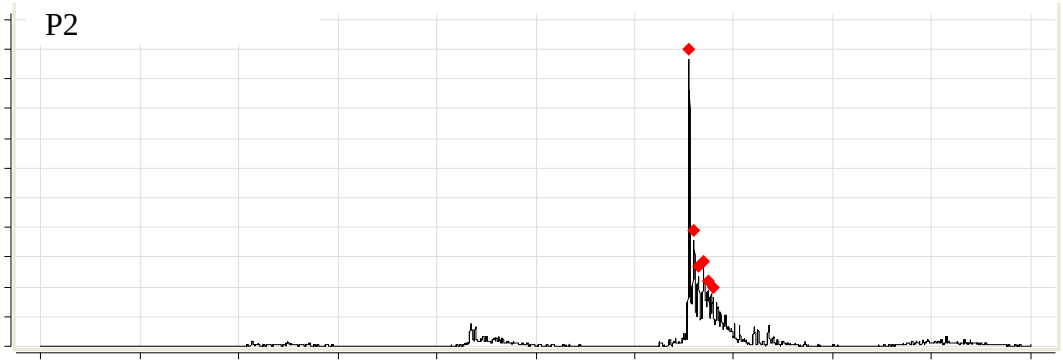


Figure 32. Q-TOF mass spectrometry of transferrin glycoprotein of CDG-like patients (P2, P15 and P16), a PMM2-CDG patient and a control. Peaks correspond to transferrin with 0, 1, or 2 *N*-glycans. Data are expressed as ratio (aglyco = 0 glycans; monoglyco = 1 glycan; diglyco = 2 glycans).

Finally, *PMM2* sequencing identified a different heterozygous mutation in three out of four CDG-like patients: P15, P16 and P2 (c.422 G>A -p.Arg141His-, IVS4 +21 G>C and IVS2 -13 G>A, respectively) (Figure 33).

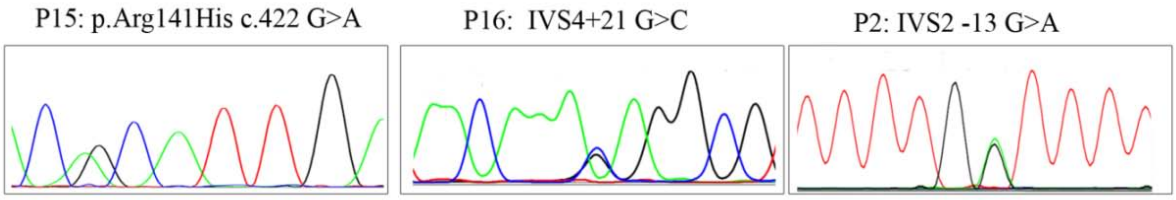


Figure 33. Electropherograms of *PMM2* mutations found in CDG-like patients (P15, P16 and P2).

The p.Arg141His mutation has a strong functional effect and is the most frequent genetic defect identified in PMM2-CDG patients, particularly in the north of Europe [112]. The remaining two mutations, IVS4 +21 G>C and IVS2 -13 G>A, could have significant effects on the splicing process according to *in silico* modelling as both generate new branch sites with 1.5 and 2.2 scores, respectively (threshold= 0). Moreover, ESE Finder software identified two potential serine/arginine-rich proteins (SR proteins) which are involved in regulating and selecting splice sites (SRSF5 and SRSF1) able to bind to the mutated sequence of IVS4+21 G>C with 3.08 and 3.07 scores, respectively.

We discarded these *PMM2* mutations as common polymorphisms in the Spanish population since the p.Arg141His substitution was found in 5 heterozygous subjects out of 1,570 healthy controls (0.3%, allele frequency 0.0016), and the IVS4 +21 G>C and IVS2 -13 G>A mutations were not identified in 61 and 47 healthy subjects, respectively.

However, these intronic mutations have never been described in PMM2-CDG patients [112].

Hence, both a similar laboratory phenotype and a common clinical feature (thrombosis) of these patients with a single *PMM2* mutation suggest the existence of a new disorder that we have called CDG-like.

3. Follow-up of patient P15

Five years follow-up of patient P15 allowed us to study the evolution of abnormal antithrombin, α1-antitrypsin and transferrin glycoforms. As shown in figure 34, the amount of abnormal α1-antitrypsin glycoforms is highly variable, being hardly detectable at certain time points. As expected, there is a positive correlation between the amounts of these abnormal α1-antitrypsin glycoforms with the peak of asialotransferrin. Interestingly, we also detected an inverse correlation of these parameters with plasma antithrombin activity (Figure 34).

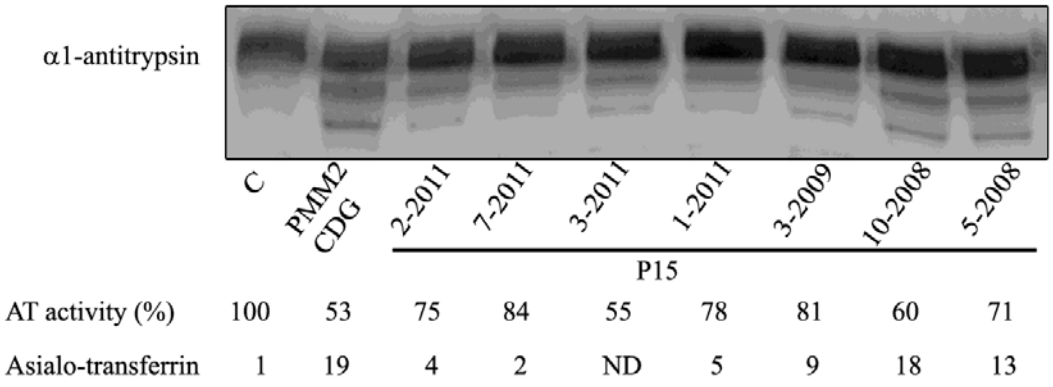


Figure 34. Evolution of α1-antitrypsin glycoforms, plasma antithrombin (AT) activity (% of anti-FXa compared with a pool of 100 healthy controls) and asialo-transferrin (area of the asialotransferrin peak in HPLC) levels in patient P15 with CDG-like during five years of follow-up. The date of each sample (month-year) is indicated. C: healthy control.

4. Proteomic analysis of platelet glycoproteins in PMM2-CDG and CDG-like patients

All abnormal glycosylated proteins identified in CDG-like and PMM2-CDG patients were of hepatic origin. We wanted to study whether these disorders could also affect *N*-glycoproteins from other cells or tissues. Most studies were done with platelet proteins due to the key haemostatic role of platelets. Western blot analysis did not

reveal abnormal glycoforms of GPIbα, GPIIIa, von Willebrand factor (vWF), thrombospondin or PEAR1 neither in PMM2-CDG nor CDG-like patients (Figure 35). We were not able to detect smaller glycoforms of many other proteins from different origin: immunoglobulin, plasminogen activator inhibitor 1 (PAI-1) or tissue factor pathway inhibitor (TFPI) (Data not shown).

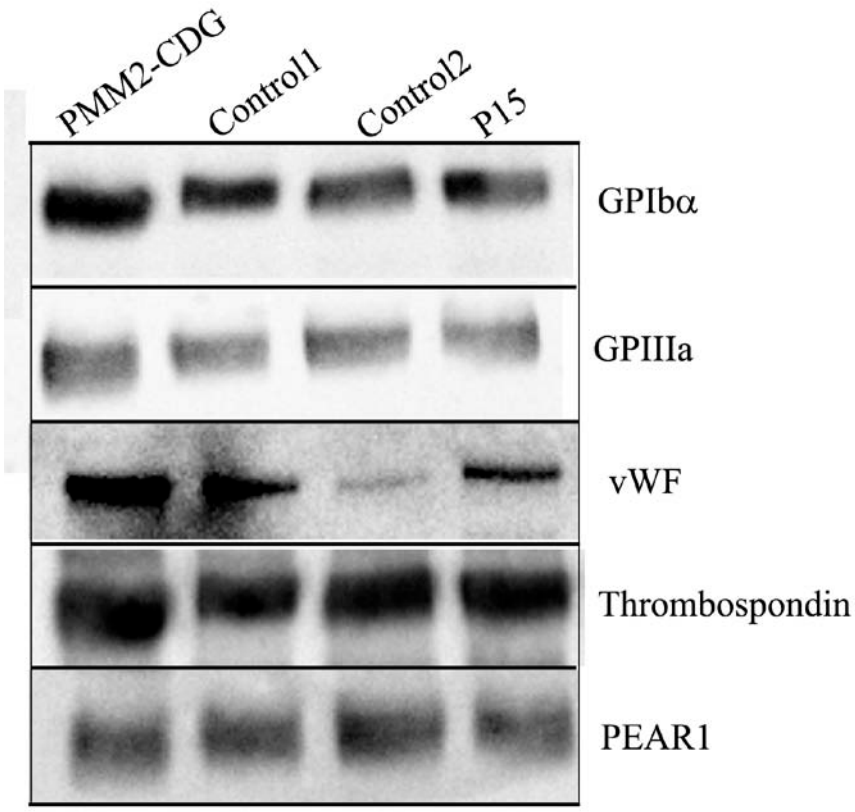


Figure 35. Western blot of platelet glycoproteins in a PMM2-CDG patient, a CDG-like patient (P15) and two controls.

We also performed a proteomic study of platelet glycoproteins. The 2D pattern of a PMM2-CDG patient was similar to that of a CDG-like patient (P15), but they were different to that of controls (Figure 36). Interestingly, MALDI-TOF analysis of tryptic peptides from unique spots of PMM2-CDG and CDG-like patients, which were compatible with a glycosylation defect due to the smaller size and increased pI; revealed the presence of hepatic proteins: α1-antitrypsin, transferrin or fibrinogen (Figure 36). However, this study did not identify any abnormal platelet glycoprotein in PMM2-CDG or CDG-like patients.

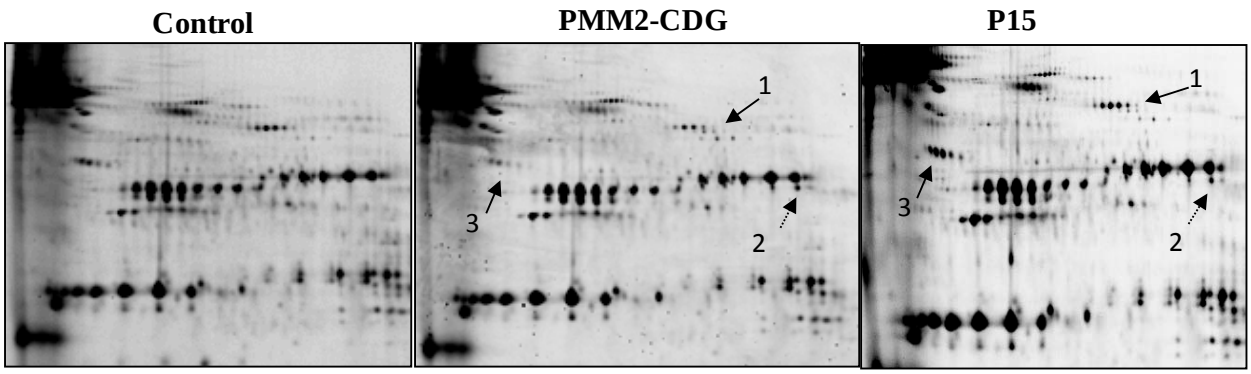


Figure 36. 2D gel of platelet glycoproteins from a control, a PMM2-CDG patient and a CDG-like patient (P15). Arrows show spots further analysed by mass spectrometry. Peptides identified correspond to hepatic proteins: 1 transferrin 2: β-fibrinogen, 3: α1-antitrypsin.

5. Search for additional factors involved in CDG-like disorders

The identification of a single *PMM2* defect in 3 patients with CDG-like disorders together with the fact that these patients do only share a clinical feature with PMM2-CDG patients (thrombosis) strongly suggest the presence of an additional factor that exacerbate the effects of the single *PMM2* mutation. P15 also had Darier disease which was inherited by his mother. We originally hypothesized that this combination might explain the CDG-like disorder. We speculated that the mutation of the Ca^{2+} pump that features the Darier disease [113] might finally affect the pH of Golgi, leading to the mislocalization of galactosyl-transferases, as described for inhibitors of the proton pump or pH neutralizing agents [114]. In order to test this hypothesis we treated HepG2 cells with bafilomycin A1, a proton-pump chemical inhibitor. As indicated in figure 37, this inhibitor caused a significant reduction in the levels of antithrombin secreted to the medium, potentially by its toxic effects on this cells, but the antithrombin secreted showed an abnormal glycosylation. Similarly, ammonium persulphate, an agent that neutralizes the pH inside Golgi, caused formation of endosomes (Figure 38), as previously described [114], and reduced glycosylation of secreted antithrombin (Figure 37). However, treatment of HepG2 cells with cyclopiazonic acid (CPA), a strong Ca^{2+} pump inhibitor, had a toxic effect that reduced the levels of secreted antithrombin, but had negligible effects on antithrombin glycosylation (Figure 37).

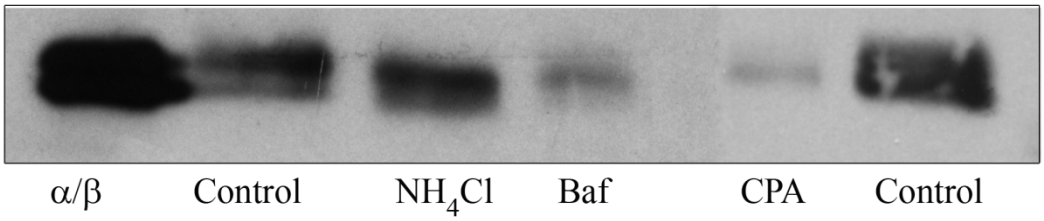


Figure 37. Electrophoretic analysis of antithrombin secreted to the conditioned medium of HepG2 treated with bafilomycin A1 (Baf), anmonium persulphate (NH_4Cl) and cyclopiazonic acid (CPA); α/β a mixture of α and β glycoforms of antithrombin purified from plasma of healthy subjects.

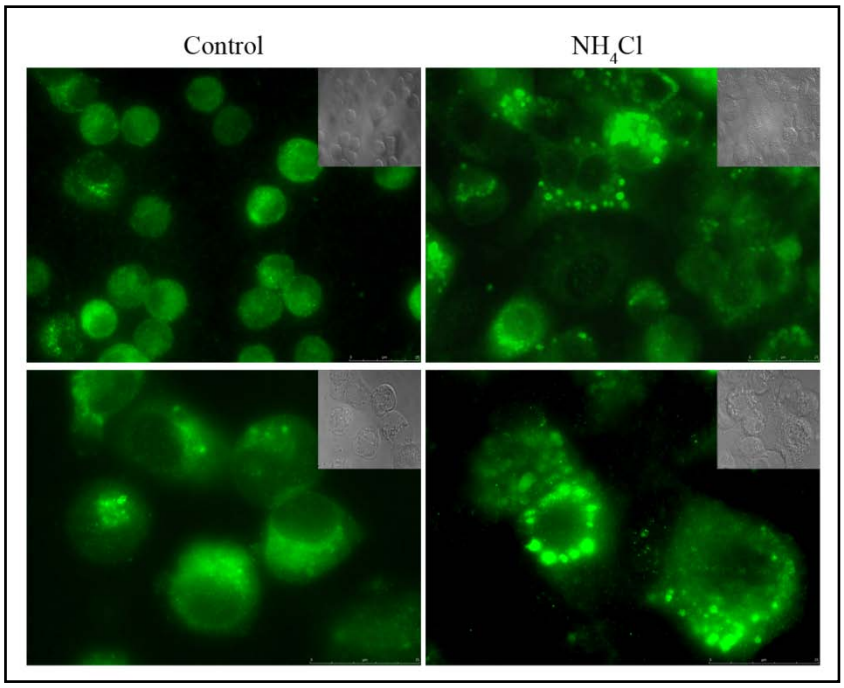


Figure 38. Immunoflourescence of intracellular antithrombin from HepG2 cells treated with anmonium persulphate (NH_4Cl).

Moreover, no further CDG-like patient reported Darier disease. In contrast, and interestingly, all four CDG-like patients identified in this study reported high alcohol consumption.

6. Epidemiologic analysis evaluating potential CDG-like disorders in patients with thrombosis

We firstly evaluated the prevalence of CDG-like disorders in two different thrombotic diseases: DVT and AMI evaluating the electrophoretic pattern of α1-antitrypsin in plasma samples from 571 Spanish patients with DVT, 262 Spanish patients with AMI and 393 Spanish controls. Patients and controls with an abnormal

pattern compatible with a CDG-like disorder were verified by electrophoretic analysis of plasma antithrombin. A mild CDG-like pattern was identified in 4 DVT patients, 2 AMI patients and only one control. Table 12 indicates the clinical data and results of *PMM2* sequencing in these subjects. Interestingly, the control and one DVT patient carried the p.Arg141His mutation and one patient carried a new mutation in intron 2 of *PMM2* (IVS2-14insT) in compound heterozygosity with Glu197Ala.

Table 12. Clinical and genetic features of patients and controls with a CDG-like pattern. DVT: patients with deep venous thrombosis; AMI: patients with acute myocardial infarction; WT: Wild Type. The age of the first event is between brackets.

Patient	Gender	Age	Recurrence	Risk factors	Familial history	*PMM2*
DVT1	Male	42(40)	No	No	Yes	WT
DVT2	Female	47(38)	No	FVL	No	E197A&IVS2-14InsT*
DVT3	Male	35(34)	Yes	No	ND	WT
DVT4	Female	36(36)	No	No	No	R141H
AMI1	Male	42(42)	No	No	Yes	rs11074924 (T/T)
AMI2	Male	48(48)	No	No	No	rs11074924 (T/T)
Control	Male	36	No	No	No	R141H

*** New mutation in intron 2 of *PMM2***

Then, we screened the prevalence of p.Arg141His, the most frequent *PMM2* alteration identified in PMM2-CDG patients, also identified in CDG-like patients. The genotyping was done in two case-control studies from different populations with different allele frequency for this mutation: Denmark and Spain. Figure 39 shows the results of this study. The p.Arg141His mutation had higher prevalence in patients with thrombosis than controls but differences did not reach statistical significance.

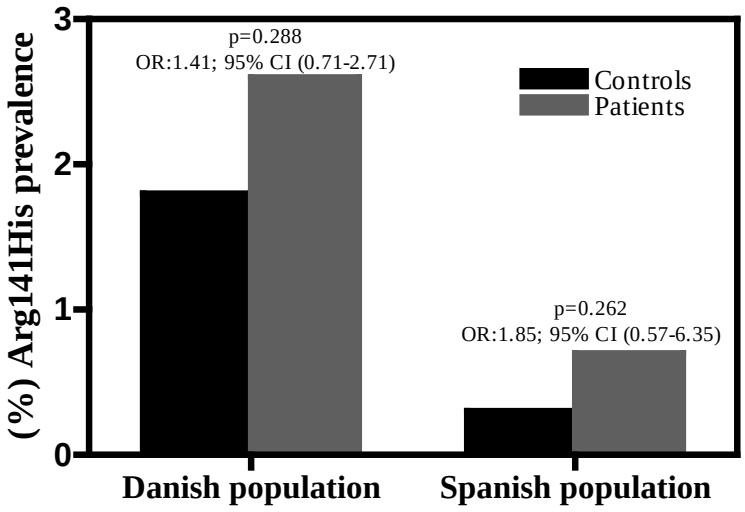


Figure 39. Prevalence of *PMM2* p.Arg141His mutation in patients and controls.

DISCUSSION

The great importance of antithrombin in the physiopathology of the haemostatic system is due, in some extent, to the so efficient anticoagulant mechanism of this serpin and the wide variety of procoagulant proteases that is able to inhibit [115]. Furthermore, since 1965 when O Egeberg described the first family with deficiency associated with a high incidence of venous thrombosis [2] more than 296 different mutations have been identified in patients with venous thrombosis and antithrombin deficiency. Although the clinical features of carriers of congenital antithrombin deficiency may be quite heterogeneous, the thrombotic risk associated with antithrombin deficiency is in general high, particularly in patients with type I deficiency whose risk is about 20-50 fold respect healthy people [57]. Fortunately, the prevalence of antithrombin type I deficiency is very low, even among patients with venous thrombosis (1 %) [116]. In contrast, subjects carrying the antithrombin Cambridge II mutation (p.Ala384Ser), relatively frequent (0.2% in our population), have a moderate risk of thrombosis, because it only mildly impairs the anti-FIIa activity of antithrombin [76]. In agreement with these data, moderate reduction in antithrombin levels has recently been suggested to mildly increase the thrombotic risk [70]. All these data strongly encourage investigating all potential mechanisms that might impair antithrombin levels or function. This search has rendered excellent results in patients with congenital antithrombin deficiency, as the causative mutation is identified in a high percentage of cases (up to 90% according to recent large studies [74], and 95% in our large cohort of patients with antithrombin deficiency). However, these studies are restricted to the *SERPINC1* gene, the gene encoding antithrombin, and all genetic defects identified are gross gene deletions or mutations affecting exons or flanking regions. The practically absence of genetic modifications associated with reduced levels of this key anticoagulant through regulatory mechanisms is surprising, probably reflecting that mutations affecting these regions are very rare or have small functional consequences. Moreover, few regions with potential regulatory role have been described in the *SERPINC1* gene. So far, only one report from our group has reported a very moderate effect associated to a common polymorphism in intron 1 [65], which also explains its negligible thrombotic consequences [117-119].

This study aimed to identify new genes and genetic modifications involved in the regulation or modulation of antithrombin levels or function, and to characterize the mechanism(s) involved. To achieve this objective, we followed three different approaches:

1. To identify new regulatory regions in the *SERPINC1* gene, we sequenced a 15,000 bp of this gene in three groups of samples with different antithrombin levels.
2. To identify antithrombin modulating genes, a Genome Wide Association Study (GWAS) and further experimental validations were done.
3. To identify new mechanisms leading to antithrombin deficiency, we studied the biochemical features of plasma antithrombin from patients with antithrombin deficiency and no mutations in *SERPINC1* gene.

The result of this study reveals a fascinating complex heterogeneity of new mechanisms affecting the levels of antithrombin with potential pathological consequences.

1) We analyzed 1,500 bp of the promoter region of the *SERPINC1* gene in 29 subjects with antithrombin deficiency without genetic modifications in the coding or splicing regions and no gross gene deletions. This study identified the first mutation in the promoter region associated with antithrombin deficiency. Certainly, the g.2143 C>G change identified in family 2 does not abolish the transcription of the mutated allele, as demonstrated by the antithrombin levels observed in carriers, but significantly impairs it. This functional effect was further confirmed by reporter assays in different cell lines. Moreover, our results verify the regulatory relevance of the region of at least 200 bp upstream the ATG start codon. The g.2143 C>G modification is located in one of the four protected regions identified in two independent reports [61,120]. Thus, it is located at the end of region C proposed by Fernandez-Rachubinski and coworkers [61], and in the middle of region II according to the nomenclature of Tremp and co-workers [120] (Figure 40). Interestingly, it has been suggested that the hepatic nuclear factor 4 (HNF4) interacts with this region [61] (Figure 40).

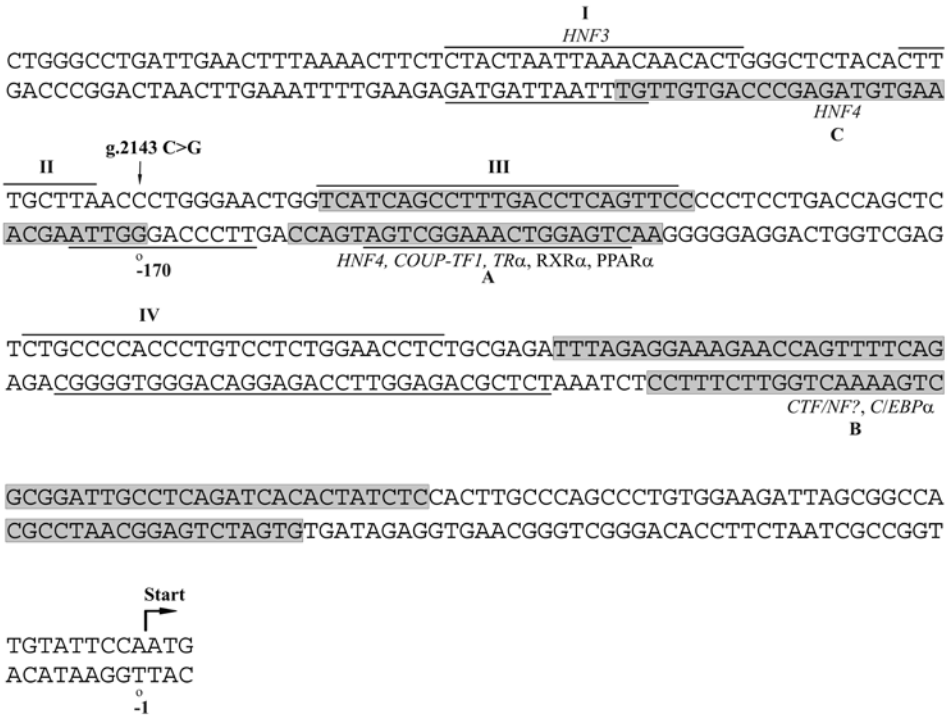


Figure 40. Localization of g.2143 C>G in potential promoter regulatory regions of the *SERPINC1* gene identified by DNase protection assay. Sequences highlighted in grey (A-C) were described by Fernandez-Rachubinski and coworkers [61] while regions underlined (I-IV) were described by Tremp and coworkers [61,120]. The interacting transcription factors identified are indicated below each sequence

Moreover, preliminary results in one additional patient with moderate antithrombin deficiency have identified one new mutation g.1091 C>T, with a potential regulatory effect in the most distal promoter. Further familial and reporter studies are required to verify the regulatory relevance of this genetic modification as well as to identify the mechanism(s) involved.

Therefore, our results support that this promoter region should be included in the molecular analysis of subjects with antithrombin deficiency, particularly in those subjects with type I deficiency and without mutation in the coding region of *SERPINC1*.

However, we did not detect mutations or any functional polymorphism in the promoter of 24 patients with moderate antithrombin deficiency or two patients with severe antithrombin deficiency. These data strongly suggest the presence of additional regulatory elements located up or down stream the region studied. Actually, the identification of a family (F1) and one patient (P26) who belongs to the same geographic area (Galicia) than F1, with type I antithrombin deficiency and severe thrombosis who had neither gross deletion nor functional mutations in the 15,925 bp of

the *SERPINC1* gene, encourage further search for remote regulatory regions. The recent ENCODE project may assist this search [121].

Regulatory regions affecting the *SERPINC1* gene might also be studied in the general population. Actually, the anti-FXa activity, the method widely used in thrombophilic test to detect antithrombin deficiency, shows a great variability in the 307 healthy donors with normal distribution (70-130%) [122]. Factors such as sex, body mass index, oral contraceptives or race have been associated with antithrombin variability [90], but genetic factors must also play a relevant role according to the high heritability of this trait (h= 0.486) [64]. The sequence of the *SERPINC1* locus in five subjects with extreme antithrombin levels within the normal range (75% and 115%) selected from 307 healthy donors aimed to detect genetic variations in the *SERPINC1* gene involved in such a great interindividual variability. We verified the minor functional effect of the rs2227589 polymorphism located in intron 1, and we found a polymorphism in the promoter region that deserved our attention. The g.2085 T>C change identified in a subject with high antithrombin levels could potentially modify the binding of different transcription factors. However, our study revealed that this polymorphism had no significant effect on antithrombin levels. The low allelic frequency of this polymorphism (0.027) encourages confirming these results in further studies including more subjects.

Despite its small size, the dissection of the genetic architecture of the *SERPINC1* gene in healthy subjects with extreme antithrombin levels revealed two relevant findings. A) The genetic variability of this gene is very low. Actually, we only identified 22 polymorphisms, 1.4 polymorphisms/1,000 bp, a low rate. Particular interest has the nearly absence of polymorphisms in the coding region of this gene. We have only identified two linked polymorphisms in exon 5, which in addition did not modify the aminoacidic sequence (rs5877 and rs5878). A deep review of the literature and SNPs data bases (Bayston T, Lane D. Imperial College of London. Antithrombin mutation database. http://www1.imperial.ac.uk/departmentofmedicine/divisions/experimentalmedicine/haematology/coag/antithrombin/; http://www.genecards.org/cgi-bin/carddisp.pl?gene=SERPINC1&snp=267&search=serpinc1&rf=/home/genecards/current/website/carddisp.pl#snp) only reveals 19 missense polymorphisms (Table 13), some of them with low prevalence or associated with antithrombin deficiency.

Table 13. Missense polymorphisms identified in the *SERPINC1* gene.

Chr. position	mRNA pos	dbSNP rs# cluster id	MAF	dbSNP allele	Protein residue	Amino acid pos*	Clinical interpretation**
173873072	1469	rs5879		T	Asn [N]	450	
				C	Asn [N]	450	
173873198	1343	rs201411353	0.0005	C	Asn [N]	408	
				T	Asn [N]	408	
173876599	1326	rs138743710		T	Ser [S]	403	Probable-pathogenic
				G	Ala [A]	403	
173876634	1291	rs201541724		A	Gln [Q]	391	
				G	Arg [R]	391	
173878720	1242	rs149006854		A	Met [M]	375	Probable-pathogenic
				C	Leu [L]	375	
173878832	1130	rs5878	0.4675	G	Gln [Q]	337	
				A	Gln [Q]	337	
173878838	1124	rs192187532	0.0037	A	Val [V]	335	
				G	Val [V]	335	
173878862	1100	rs5877	0.4528	G	Val [V]	327	
				A	Val [V]	327	
173878868	1094	rs200857896		G	Ala [A]	325	
				C	Ala [A]	325	
173878933	1029	rs150507232		C	Leu [L]	304	
				T	Leu [L]	304	
173878960	1002	rs201381904	0.0005	A	Met [M]	295	
				G	Val [V]	295	
173878961	1001	rs184508490	0.0005	T	Arg [R]	294	
				C	Arg [R]	294	
173878985	977	rs139463995		C	His [H]	286	Probable-pathogenic
				G	Gln [Q]	286	
173879035	927	rs139275128		T	Leu [L]	270	
				C	Leu [L]	270	
173879905	868	rs144084678		T	Ile [I]	250	Probable-pathogenic
				C	Thr [T]	250	
173879935	838	rs200861147	0.0005	G	Ser [S]	240	
				A	Asn [N]	240	
173879958	815	rs141292014		C	Asp [D]	232	
				T	Asp[D]	232	

Chr. position	mRNA pos	dbSNP rs# cluster id	MAF	dbSNP allele	Protein residue	Amino acid pos*	Clinical interpretation**
173879961	812	rs140452859	0.0014	T	Thr [T]	231	
				C	Thr [T]	231	
173879976	797	rs2227627	0.0009	T	Thr [T]	226	
				C	Thr [T]	226	
173879984	789	rs146733468		G	Asp [D]	224	Probable-pathogenic
				A	Asn [N]	224	
173880967	713	rs183416252	0.0005	C	Tyr [Y]	198	
				T	Tyr [Y]	198	
173881032	648	rs143521873		T	Cys [C]	177	Probable-pathogenic
				C	Arg [R]	177	
173881122	558	rs2227606	0.0023	G	Ala [A]	147	Probable-pathogenic
				A	Thr [T]	147	
173881135	545	rs199525344	0.0005	A	Thr [T]	142	
				C	Thr [T]	142	
173883748	470	rs200810296	0.0005	A	Thr [T]	117	
				G	Thr [T]	117	
173883749	469	rs139392083		T	Met [M]	117	Probable-pathogenic
				C	Thr [T]	117	
173883788	430	rs200118419		T	Val [V]	104	Probable-pathogenic
				A	Asp [D]	104	
173883797	421	rs199895690	0.0005	G	Cys [C]	101	Probable-pathogenic
				C	Ser [S]	101	
173883834	384	rs147266200	0.0005	T	Cys [C]	89	Pathogenic
				C	Arg [R]	89	
173883987	231	rs145771113		A	Asn [N]	38	
				G	Asp [D]	38	
173884000	218	rs147676453		A	Gln [Q]	33	
				T	His [H]	33	
				C	His [H]	33	
173884010	208	rs2227624	0.0018	A	Glu [E]	30	
				T	Val [V]	30	
173884012	206	rs188274879	0.0005	T	Cys [C]	29	
				C	Cys [C]	29	
173884014	204	rs147918976		C	Arg [R]	29	
				T	Cys [C]	29	
173886369	148	rs61736655	0.0005	A	Asn [N]	10	

Chr. position	mRNA pos	dbSNP rs# cluster id	MAF	dbSNP allele	Protein residue	Amino acid pos*	Clinical interpretation**
				C	Thr [T]	10	

*: numbering from first Met.
**: Clinical interpretation based on identification in patients with antithrombin deficiency (Pathological) or the same residue affected in a patient with antithrombin deficiency, affecting a highly conserved or functional residue (probable pathogenic).
Lines in gray: missense polymorphisms
Lines in white: silent polymorphisms
Data from http://www.ncbi.nlm.nih.gov/sites/varvu?gene=462&rs=267598177
http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?geneId=462

Antithrombin deficiency associated to missense mutations may be caused by mechanisms common to many other proteins, such as RNA instability. However, the extraordinary functional and conformational sensitivity of this molecule [79] could also contribute to explain the low prevalence of missense polymorphisms. Antithrombin, as a member of the serpin superfamily has to fold into a metastable conformation. Minor modification affecting the folding of this molecule, particularly those mutations affecting key structural residues, might impair the folding process and the correct intracellular pathway. The number of candidate residues is really high due to the conservation among serpins. Moreover, antithrombin also has functional domains highly sensitive to even minor mutations. Thus, changes of residues involved in the interaction with heparin and the latter activation process, residues involved in the interaction with the target proteases or its translocation process after the proteolytic cleavage might have deleterious consequences that could impair or abolish the inhibitory capacity of this molecule. All these data suggest that most changes, even affecting a single residue, might have significant consequences leading to antithrombin deficiency. Further genetic studies in patients with thrombosis but not significant antithrombin deficiency are encouraged to verify if mild mutations with potential pathological relevance might exist.

B) The genetic variability of *SERPINC1* plays a minor role on the interindividual variability of antithrombin levels in plasma. Only one polymorphism in intron 1 (rs2227589) explains a minor percentage of the interindividual variation (7%).

The low genetic variability of *SERPINC1* together with the high heritability of antithrombin levels [64] support the existence of modulating genes that might be involved in s variability, and could play a role in the course or severity of thrombotic disorders [123]. These genes might encode for transcriptional factors, molecules

involved in folding or secretion, enzymes implicated in posttranslational modifications, proteases or miRNAs. An effort to identify these potential modulating genes of antithrombin has also been done in this study.

2) The first objective has confirmed the low genetic variability of *SERPINC1* locus and its minor functional influence on antithrombin levels. Therefore we focused our search to other genes that could indirectly modulate antithrombin by a multi-stage approach, the same approach followed by a very recent study that extended the search to protein C and protein S [124]. The potential identification of modulators of key anticoagulant genes might help to identify new genetic risk factors for venous thrombosis. Moreover, this study might also identify new mechanisms or pathways involved in the folding, secretion, function or clearance of antithrombin, which may also be extrapolated to other homologous proteins. A similar approach might be considered for other haemostatic factors, with interindividual variability not fully explained by the genetic variability of the coding gene, particularly considering the few studies that have tried to identify modifying genes of haemostatic factors [123].

However, both multi-stage approaches failed to find SNPs associated with antithrombin levels at genome-wide significance. These negative results, despite the high heritability of antithrombin levels, strongly suggest that different elements with potential moderate effect might modulate the levels of this key anticoagulant and strength the requirement of additional approaches using different experimental studies, like those used in this objective of our study, to identify antithrombin-modulating genes. Thus, our replication analysis has evaluated 12 SNPs with milder association with anti-FXa activity in the GWAS, three rs762057, rs713703 and rs240082 affecting *LARGE*. One SNP, rs762057, maintained the association with anti-FXa activity in the validation cohort, and interestingly, in both cohorts the polymorphic allele associated with a modest increase of anti-FXa activity in plasma (2-4%). This result also explains the absence of thrombotic relevance for this polymorphism in the case-control study. Moreover, the *LARGE* H2 haplotype defined by rs762057, rs713703 and rs240082, also associated with a modest increase of anti-FXa activity. The relevance of these three *LARGE* SNPs on the heritability of anti-FXa-levels was minor: 4.3%, 4.9% and 4.0% for rs762057, rs713703 and rs240082, respectively, but rose to 7.8% when considering the three SNPs together. It is possible that other polymorphisms not included in the chip, haplotypes or rare mutations of *LARGE* might have stronger functional

consequences. Unfortunately, the name of this gene reflects its length and genetic variability. It expands more than 756,000 bp and contains more than 7,790 polymorphisms, a size and genetic variability that may complicate the studies required to dissect the genetic architecture of this gene and to evaluate its potential functional and pathological relevance.

As the GWAS approach and the validation study identified *LARGE* as a candidate antithrombin-modulating gene based only on association data, additional experimental evidences were required to sustain a potential role of *LARGE* on the indirect regulation of the levels of this anticoagulant. Thus silencing experiments confirmed a role for *LARGE* modulating antithrombin levels. Moreover, these results may also open new mechanisms or pathways involved in the folding, secretion, function or clearance of this important anticoagulant, which may also be extrapolated to other homologous proteins.

Additionally, our study also opens new attractive roles for LARGE, a protein largely unknown. The activity of LARGE is required for the right function alpha-dystroglycan (α-DG) as mutations in the *LARGE* gene have been identified in congenital muscular dystrophy patients with brain abnormalities [125]. Early evidences suggested that LARGE participates in the formation of a phosphoryl glycan branch on *O*-linked mannose or it modifies complex *N*- and mucin *O*-glycans [126,127]. Indeed, site-directed mutagenesis of the transferase-like domains abolishes its glycosylation capability, suggesting that LARGE may function as a glycosyltransferase [126]. Despite LARGE plays a critical role in biosynthesis of the functional *O*-glycans of α-dystroglycan (α-DG) [128], it is also described that its over expression competes to modify GlcNAc terminals with Gal to generate the functional glycans not only in *O*-linked but also in *N*-glycans in α-DG [129]. Moreover, a recent study over expressing LARGE suggested that this protein can also mediate phosphoryl glycosylation on *N*-linked glycans of non-α-DG proteins [130]. Finally, an excellent and recent study demonstrated that LARGE could act as a bifunctional glycosyltransferase, with both xylosyltransferase and glucuronyltransferase activities, which produced repeating units of [–3-xylose–α1,3–glucuronic acid-β1–] [131]. How could LARGE modulate antithrombin levels? Since reduced expression of *LARGE* did not affect the expression of *SERPINC1*, we can rule out an indirect role of LARGE on the transcriptional regulation of antithrombin. According to our results, a direct effect on the glycomic features of antithrombin might also be discarded. The reduced expression of *LARGE* seems to down-regulate the secretion of antithrombin, without significant intracellular

accumulation, probably reflecting a degradation of abnormal folding proteins [132]. These data together with the impaired secretion of other proteins (α1-antitrypsin, prothrombin or transferrin) observed under silencing of *LARGE* encouraged us to suggest a new function for LARGE in intracellular folding and/or secretion. The fact that the main affected protein among all tested is antithrombin, a protein with an heparin binding domain [133], together with the fact that the glycan produced by LARGE resembles heparin-heparan sulfate (HS) and chondroitin-dermatan sulfate (CS-DS) glycosaminoglycans (GAGs) [131], make attractive this hypothesis. Further studies are required to verify this hypothesis and to define the exact mechanism involving LARGE on the folding, secretion and degradation pathways of glycoproteins, particularly antithrombin, and to determine the final effect on the haemostatic equilibrium, as LARGE might also reduce the secretion of prothrombotic proteins such as prothrombin.

In conclusion, through a hypothesis free GWAS approach followed by functional experiments, we have identified the first modulating gene of antithrombin: *LARGE*. This is a bifunctional glycosyltransferase that produces a glycan that resembles heparin-HS– and CS-DS–GAGs linear polysaccharides. Our study also found that under expression of LARGE also impairs the secretion of many other proteins, suggesting that LARGE could also have additional roles, including a potential role in the appropriate folding and/or secretion pathways of multiple proteins.

3) In the deep seeking to find new genetic elements involved in the risk of thrombosis, rare mutations affecting different elements of the haemostatic system are emerging candidates [8,9]. The identification of *LARGE* as a potential antithrombin modulating gene opens new perspectives in thrombophilia. Accordingly, mutations affecting other genes able to influence antithrombin levels or the whole haemostatic system may be alternative thrombophilic candidates. In the third objective of the present study, we identified a mechanism leading to severe antithrombin deficiency with a significant risk of thrombosis. The analysis of patients with thrombosis and antithrombin deficiency but without mutations or deletions in the *SERPINC1* gene was crucial to identify these new mechanisms.

We firstly focused on the analysis of antithrombin in a thrombophilic patient with severe deficiency and early and recurrent venous thrombosis. This study allowed us to identify a particular thrombophilic defect that, affecting the *N*-glycosylation pathway, disturbs this key anticoagulant serpin. This patient (P27) developed four severe

thrombotic episodes and antithrombin deficiency was the single thrombophilic factor identified. The initial identification of abnormal antithrombin glycoforms by electrophoretic analysis followed by different biochemical studies, a deeply glycoprotein study and a *PMM2* molecular characterization has provided the final accurate diagnosis of Congenital Disorder of Glycosylation type Ia (CDG-Ia, also called PMM2-CDG).

Congenital Disorders of Glycosylation (CDG) result from defects in the assembly, transfer, and processing of *N*-linked oligosaccharides. Since 1980, when the first *N*-glycosylation defect was described and characterized [134] over 45 types of CDGs have been recognized. Although very rare, CDGs are the most rapidly expanding group among the currently known 3,000 monogenetic inherited diseases. All known CDGs have a recessive inheritance except EXT1/EXT2-CDG which is dominantly inherited and MGAT1-CDG which is X linked [110].

CDG-Ia (PMM2-CDG, OMIM#212065) is by far the most prevalent protein *N*-glycosylation defect with more than 700 affected individuals worldwide. The prevalence could be as high as 1:20,000 [135]. Phosphomannomutase (PMM) 2 deficiency is a (cytosolic) defect in the second step of the mannose pathway (transforming mannose 6-phosphate into mannose 1-phosphate), which normally leads to the synthesis of GDP-mannose. This nucleotide sugar is the donor of the mannoses used in the endoplasmic reticulum to assemble the dolichol-pyrophosphate oligosaccharide precursor (Figure 41). Deficiency of GDP-mannose causes hypoglycosylation of numerous glycoproteins, including serum proteins, lysosomal enzymes, and membranous glycoproteins [136].

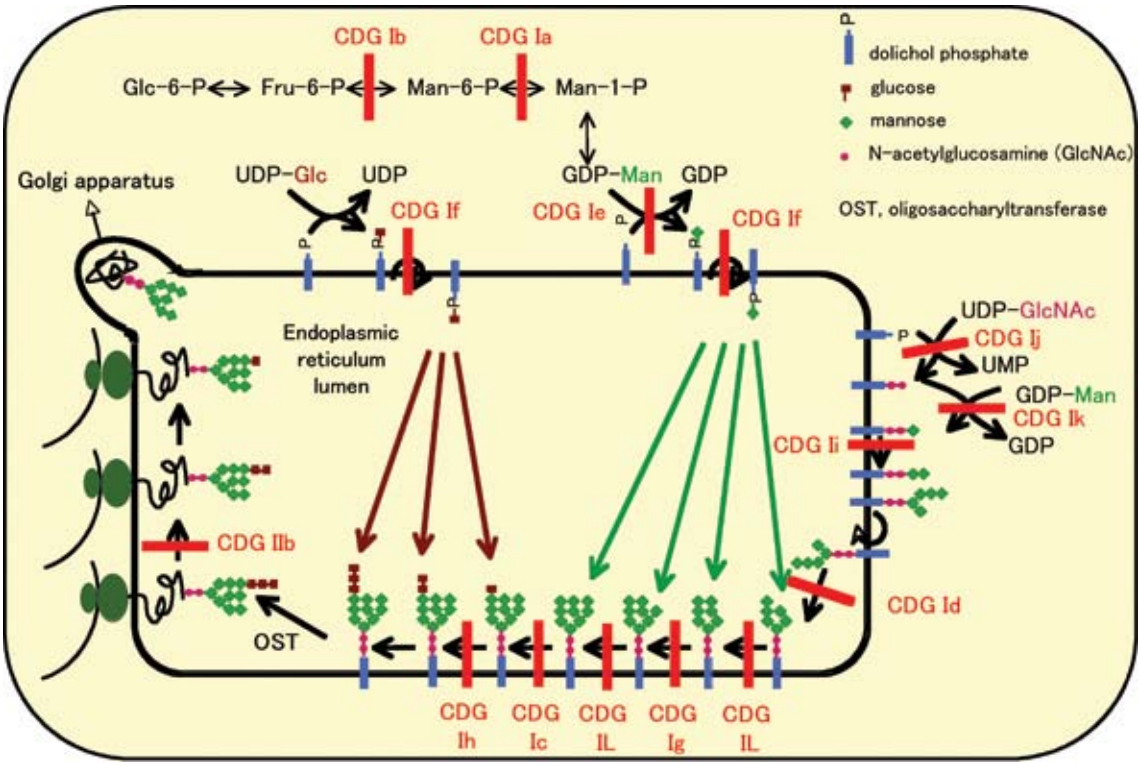


Figure 41*: N*-glycosylation of proteins and its defects. The figure shows the enzymatic steps taking place in the cytosol and in the endoplasmic reticulum. Enzymatic defects in any of these steps leads to CDG (congenital disorder of glycosylation). *N*-glycosylation also involves steps taking place in the Golgi apparatus (not shown), which may also be defective.

The clinical spectrum associated to CDGs in general and PMM2-CDG in particular, is very large. The nervous system is affected in all patients, and most other organs are involved in a variable way. The neurological picture comprises alternating internal strabism and other abnormal eye movements, axial hypotonia, psychomotor retardation, ataxia, and hyporeflexia. After infancy, symptoms include retinitis pigmentosa, often stroke-like episodes, and sometimes epilepsy. During the first year(s) of life, there are variable feeding problems such as anorexia, vomiting, and diarrhea. These can lead to severe failure to thrive. Other features include variable dysmorphy (large hypoplastic/dysplastic ears, abnormal subcutaneous adipose tissue distribution), hepatomegaly, skeletal abnormalities, and hypogonadism. Some infants develop pericardial effusion and/or cardiomyopathy. At the other end of the spectrum are patients with a very mild phenotype (no dysmorphy, mild psychomotor retardation). Patients often have an extroverted and happy appearance. There is an increased mortality in the first years of life due to vital organ involvement or severe infection. Approximately 90 mutations have been identified in the *PMM2* gene. The mutation

leading to the p.Arg141His substitution is present in about 75% of the alleles of Central European patients [112].

The haemostatic system in these patients is also relatively impaired. Up to 55% of PMM2-CDG patients show acute events such as ‘‘stroke-like’’ episodes, ischemic cerebral vascular accidents, and at least 10 cases with either venous or arterial thrombosis have been described [137]. Acute vascular events occurred in patients younger than 15 years, especially during fever and prolonged immobilization [137]. These accidents can be recurrent in the same patient [137]. Only few studies have evaluated haemostatic factors [137,138] and platelets [139] in PMM2-CDG with the aim to identify the mechanism(s) involved in the development of these vascular events. PMM2-CDG patients consistently have antithrombin deficiency [137,138] with abnormal hypoglycosylated antithrombin molecules [137]. This might contribute to increase the risk of thrombotic events described in these patients. However, there are no differences in antithrombin levels between patients with and without vascular events [137].

P27 is the first report of a PMM2-CDG patient with documented and recurrent venous thrombosis. Moreover, the patient suffered from thrombotic events without a clear triggering factor such as fever, catheterization or immobilization. Finally, two recurrent thrombotic events appeared under oral anticoagulant therapy. This unusual clinical severity suggests that additional haemostatic or vascular defects than those described in PMM2-CDG patients (or additional unknown thrombophilic anomalies) might be present in this patient.

This case also highlights another important conclusion concerning the diagnostic procedure in patients with antithrombin deficiency. Direct sequencing and MLPA analysis of the *SERPINC1* gene in patients with thrombosis and antithrombin deficiency is highly effective, but these studies failed to find any genetic modification in up to 10% of these patients [73,74]. Our study suggests that a diagnosis of PMM2-CDG might be suspected among the cases, particularly in patients with antithrombin deficiency without a family history of thrombosis, as well as in cases with a lower molecular weight plasma antithrombin.

Finally, this new thrombophilic factor could be more frequent than expected. The electrophoretic analysis of antithrombin in 25 adult patients with thrombosis and moderate antithrombin deficiency but no mutation in the *SERPINC1* gene has identified four patients with smaller forms of antithrombin also caused by defects in *N*-

glycosylation. Further studies in these patients confirmed that this is caused by a non-specific defect, as other hepatic glycoproteins (α1-antitrypsin and transferrin) were also affected according to electrophoresis, HPLC and Q-TOF methods. Moreover, molecular analysis of *PMM2* gene identified heterozygous mutations in three of them. The importance of these findings is not only restricted to the high frequency of this new disorder, but also in the clinical features of these patients and the pathological mechanism.

The patient with a biochemical phenotype closer to the PMM2-CDG disorder, P15, carried the *PMM2* c.422G>A missense mutation responsible for the change p.Arg141His, the most frequent alteration in PMM2-CDG patients. Around 37% of PMM2-CDG patients are heterozygous for this mutation, which has never been observed in the homozygous state, not even in North Europe where this alteration reaches the frequency of a polymorphism (1/72 in Dutch and Danish populations) [140,141]. This mutation does not abolish but significantly reduces the phosphomannomutase activity (to a very low residual activity, <1%) [142]. In P15, the p.Arg141His mutation was of paternal origin, but as described for parents of PMM2-CDG patients, heterozygous carriers have neither biochemical phenotype nor clinical symptoms. P15 also had a Darier syndrome which was inherited by his mother. Our first hypothesis considered that the combination of a mutation in the *ATP2A2* gene (the gene affected in patients with Darier disease) and in the *PMM2* gene might result in a CDG-like disorder. However, it was discarded by several evidences: 1) no reduced glycosylation of proteins, reduced antithrombin levels or thrombosis has been identified in patients with Darier syndrome, 2) the use of Ca^{2+}-pump inhibitors did not modify glycosylation of proteins in HepG2 cells, and more importantly, 3) the variable laboratory data of P15 during 5 years of follow up, 4) the absence of neurological symptoms in P15, and 5) any other CDG-like patient identified in our study had a Darier disease. Moreover, all these data argued against the combination of two congenital defects to explain the CDG-like disorder and strongly suggest a role for an acquired factor in combination with a *PMM2* defect.

Two additional CDG-like patients P16 and P2 carried *PMM2* intronic mutations (IVS4 +21 G>C and IVS2 -13 G>A, respectively). *In silico* studies revealed these mutations could generate new branch-sites. Branch-site sequence is generally located between 10 and 50 nucleotides upstream of the 3'consensus splice site (AG). Interestingly, both mutations are located within these distances. Although, mutations in

branch sites seems to be a relatively rare class of splice mutations several instances of point mutations within the branch point sequence have been published, for example one in the *COL5A1* gene associated with type II Ehlers–Danlos syndrome [143] and two others in a DNA repair gene in xeroderma pigmentosum [144]. Therefore branch-site mutations are possibly often undiagnosed [145].

These *PMM2* intronic mutations were not identified in healthy people, but their functional effect has not been elucidated yet. Further studies of the PMM2 activity in fibroblast or *PMM2* transcript analysis, and/or exon trap studies should be performed to better characterize the functional consequences of these intronic changes. However, we can expect a lower functional effect than that of p.Arg141His according to the reduced levels of underglycosylated proteins detected in carriers of these intronic mutations.

Finally, the last CDG-like patient, P22, has no *PMM2* mutations. It is possible that mutations affecting other regulatory regions of *PMM2* gene might be present in this patient, but we can't discard a mutation in other gene involved in CDG disorders. Indeed, there are 45 genes identified so far associated with CDGs [136], a number that will grow since there are patients with clinical and biochemical features of CDG without mutation in any of these genes [146].

Interestingly, all four CDG-like patients described high alcohol intake, and hepatic transaminases moderately elevated. Severe alcoholism has been previously associated with glycosylation defects [147]. Briefly, prolonged exposure to ethanol generates reactive oxygen species (ROS) [148]. Dolichol (an isoprene derivate precursor of the early steps of *N*-glycosylation pathway, Figure 41) is particularly vulnerable to the lipoperoxidative effects of ethanol-induced ROS. Indeed, Cottalasso et al. demonstrated that ethanol treatment causes a profound reduction in the dolichol (phosphate) content of rat liver microsomes and Golgi apparatus [148]. A sufficient level of microsomal dolichol-P is a prerequisite for the initiation of *N*-linked glycosylation (Figure 41). Thus, ethanol exposure may seriously impede glycoprotein synthesis and secretion. Ethanol also affects mannosyltransferase activity, an enzyme mediating mannosylation of dolichol-P during the build-up of the lipid-linked oligosaccharide (LLO) [148] (Figure 41). Furthermore, sialyltransferase and galactosyltransferase are inhibited, while sialidase activity is stimulated in response to ethanol administration [148,149]. These latter effects of ethanol, causing disruption of terminal sialylation during the second phase of *N*-linked glycosylation in the Golgi apparatus, are probably of minor importance, considering that carbohydrate-deficient

transferrin (CDT) –the best indicator of chronic alcohol abuse currently available- mainly consists of di- and asialotransferrin (type I TIEF pattern), indicating defects in LLO assembly during the first phase of *N*-linked glycosylation in the Endoplasmic Reticulum [149,150]. Moreover, in addition to the consequences on dolichol depletion, degradation of vital membrane constituents by ethanol exposure affects plasma membrane integrity and vesicular trafficking [148,151]. Finally, alcohol inhibits gluconeogenesis pathway [152] and in consequence could interfere the early steps of the *N*-glycosylation pathway because of the glucose content reduction. Thus, reduced levels of fructose-6-phosphate in alcoholic patients might be compensated by increased synthesis through the mannose-6-phosphate isomerase (MPI) activity, reducing the synthesis of mannose-1 phosphate, thereby aggravating a single *PMM2* defect as shown in Figure 42. Actually, the most recent progress on PMM2-CDG treatment consists on mannose supplementation and MPI inhibition [153].

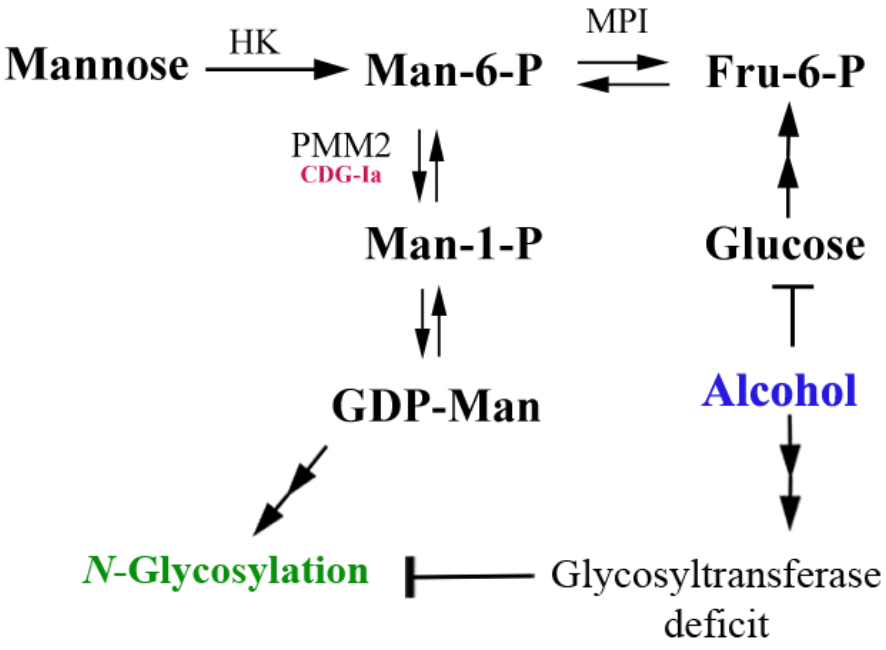


Figure 42. Potential interference of alcohol on the *N*-glycosylation metabolic pathway. HK: Hexokinase; MPI: Mannose-6 phosphate isomerase; PMM2: phosphomannomutase; Man-6-P: mannose-6-phosphate; Fru-6-P: fructose-6-phosphate; Man-1-P: mannose-1-phosphate.

However, it is important to point out that isolated alcohol abuse, even at higher levels than that found in CDG-like patients does not cause the CDG electrophoretic or HLPC pattern. We have studied plasma samples from one patient with severe alcoholism (218 g/day). The electrophoretic pattern of α1-antitrypsin was normal

(Figure 43) and HPLC analysis of transferrin glycoforms only showed a moderate increase of the asialotransferrin content, while di- or trisilaotransferrin were normal (Figure 44).

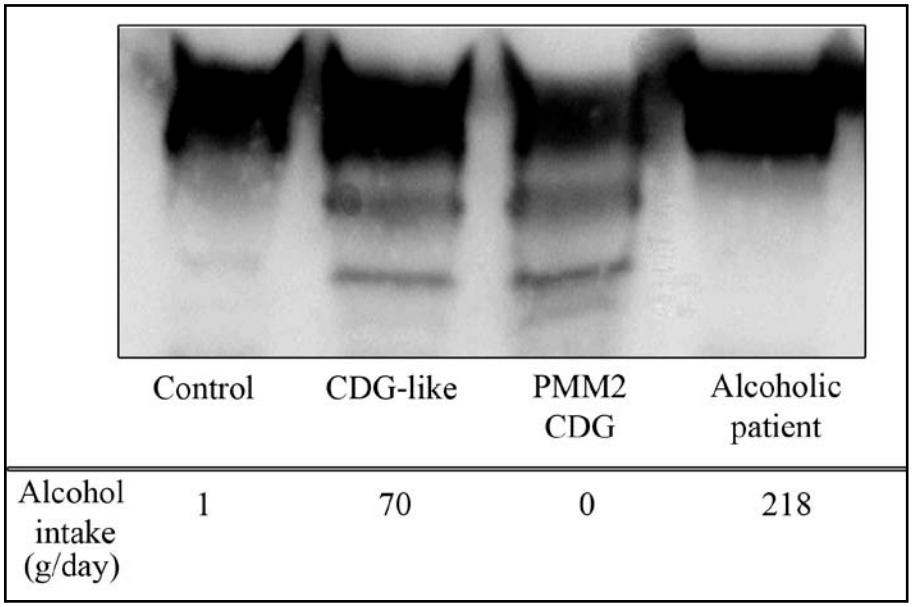


Figure 43. Electrophoretic pattern of α1-Antitrypsin of an alcoholic patient comparing with a control, PMM2-CDG and CDG-like patients. Alcohol intake is indicated below.

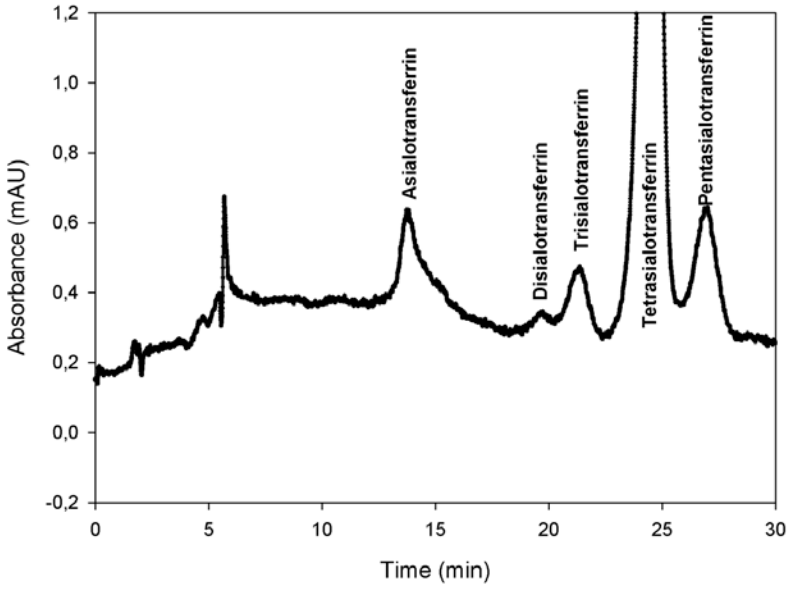


Figure 44. HPLC profile of plasma transferrin glycoforms from an alcoholic patient.

All these results strongly suggest that CDG-like disorders might be the result of the combination of a moderate congenital defect affecting the *N*-glycosylation pathway (mainly *PMM2*) and an acquired factor also affecting this pathway, being alcohol an excellent candidate. Further studies using animal models will be required to confirm that this combination produces CDG-like disorders. It is also possible that other

combinations of different genes and acquired factors might also render CDG or CDG-like disorders.

This study also aimed to evaluate the incidence of this new thrombophilic disorder in different thromboembolic diseases. Our preliminary epidemiological studies aiming to determine the incidence in thrombotic diseases of CDG-like disorders suggest a relatively high incidence in both venous and arterial thrombosis. It is true that the high incidence of CDG-like disorders among patients with moderate antithrombin deficiency (4/25: 16%) make us to expect higher prevalence of CDG-like in thrombotic disorders than that found in our study, particularly in DVT and specially in the north of Europe, where the p.Arg141His is a relatively common polymorphism. However, to fully understand the thrombotic relevance of these new disorders, it would be necessary to perform a multivariate analysis including these acquired factors that exacerbate the *PMM2* defect, at least high alcohol intake. Moreover, we can't forget that the abnormal glycosylation pattern used to identify the CDG-like disorder is highly variable (the best example is P15 during the 5 years of study) and depends on the combination with the acquired factor. It is therefore possible that our epidemiological study, performed with plasma samples of retrospectively recruited patients had underestimated the incidence of these disorders. Further studies using samples obtained during the acute event might answer this question.

Finally, it is necessary to investigate the mechanism responsible for the high risk of thrombosis described in CDG and CDG-like patients. The deficiency of antithrombin is an excellent candidate, but we do not know if this is cause or consequence of the disorder. Both anticoagulant and procoagulant hepatic proteins are affected in these patients (antithrombin and prothrombin are the best examples). Trying to identify further defects on haemostatic elements we studied platelet glycoproteins. Surprisingly, we observed no major modifications on all proteins tested, even using high sensitive methods such as proteomic studies. These results sustain that the impaired glycosylation pathway of PMM2-CDG and CDG-like patients has different consequences depending on the protein and particularly the cell type. Our results suggest that cells with a high synthesis and secretory rate like liver are potentially more affected by these disorders. Further studies are therefore required to answer how disorders of glycosylation affect the haemostatic system and to identify the mechanism underlying the increased risk of thrombosis of these patients.

CONCLUSIONS

The result of this study reveals a fascinating complex heterogeneity of new mechanisms affecting the levels of antithrombin a key anticoagulant serpin, with potential pathological consequences.

1) We identified the first mutations in the potential promoter of *SERPINC1* gene that reduced the levels of plasma antithrombin. This region must be included in molecular analysis aiming to identify the genetic defect involved in congenital antithrombin deficiency.
2) *SERPINC1* locus has low genetic variability with minor influence in the heterogeneity of the plasma antithrombin in the general population.
3) We have identified the first antithrombin modulating gene. The reduced levels or activity of LARGE, potentially through its xylosyltransferase and glucuronyltransferase activities that produce a heparin-like glycosaminoglycan, impair the secretion of antithrombin, a heparin-binding protein.
4) We diagnose a PMM2-CDG disorder in a young patient with recurrent venous thrombosis and antithrombin deficiency.
5) We identified a new CDG-like disorder based on the similar electrophoretic, HPLC or Q-TOF pattern of glycoforms of hepatic proteins than of PMM2-CDG patients, probably as the result of combination of a *PMM2* mutation and alcoholism.
6) CDG-like, as PMM2-CDG, may be considered as new thrombophilic disorder and must be studied in patients with venous thrombosis, particularly in those with moderate or severe antithrombin deficiency and no mutation or deletion in the *SERPINC1* gene.

BIBLIOGRAPHY

Reference List

1. Lippi G, Favaloro EJ, Franchini M, Guidi GC. Milestones and perspectives in coagulation and hemostasis. *Semin Thromb Hemost* 2009; **35**; 9-22.

2. Egeberg O. Inherited antithrombin deficiency causing thrombophilia. *Thromb Diath Haemorrh* 1965; **13**; 516-30.

3. Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, Coll I, Felices R, Stone W, Fontcuberta J, Blangero J. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. *Am J Hum Genet* 2000; **67**; 1452-9.

4. Bertina RM. Genetic approach to thrombophilia. *Thromb Haemost* 2001; **86**; 92-103.

5. Gohil R, Peck G, Sharma P. The genetics of venous thromboembolism. A meta-analysis involving approximately 120,000 cases and 180,000 controls. *Thromb Haemost* 2009; **102**; 360-70.

6. Germain M, Saut N, Greliche N, Dina C, Lambert JC, Perret C, Cohen W, Oudot-Mellakh T, Antoni G, Alessi MC, Zelenika D, Cambien F, Tiret L, Bertrand M, Dupuy AM, Letenneur L, Lathrop M, Emmerich J, Amouyel P, Tregouet DA, et al. Genetics of venous thrombosis: insights from a new genome wide association study. *PLoS One* 2011; **6**; e25581.

7. Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, de Andrade M. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost* 2012; **10**; 1521-31.

8. Miyawaki Y, Suzuki A, Fujita J, Maki A, Okuyama E, Murata M, Takagi A, Murate T, Kunishima S, Sakai M, Okamoto K, Matsushita T, Naoe T, Saito H, Kojima T. Thrombosis from a prothrombin mutation conveying antithrombin resistance. *N Engl J Med* 2012; **366**; 2390-6.

9. Simioni P, Tormene D, Tognin G, Gavasso S, Bulato C, Iacobelli NP, Finn JD, Spiezia L, Radu C, Arruda VR. X-linked thrombophilia with a mutant factor IX (factor IX Padua). *N Engl J Med* 2009; **361**; 1671-5.

10. Law RH, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, Rosado CJ, Langendorf CG, Pike RN, Bird PI, Whisstock JC. An overview of the serpin superfamily. *Genome Biol* 2006; **7**; 216.

11. Hunt LT, Dayhoff MO. A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha 1-proteinase inhibitor. *Biochem Biophys Res Commun* 1980; **95**; 864-71.

12. Bird PI. Serpins and regulation of cell death. *Results Probl Cell Differ* 1998; **24**; 63-89.

13. Irving JA, Pike RN, Lesk AM, Whisstock JC. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res* 2000; **10**; 1845-64.

14. Whisstock JC, Bottomley SP. Molecular gymnastics: serpin structure, folding and misfolding. *Curr Opin Struct Biol* 2006; **16**; 761-8.

15. Rau JC, Beaulieu LM, Huntington JA, Church FC. Serpins in thrombosis, hemostasis and fibrinolysis. *J Thromb Haemost* 2007; **5 Suppl 1**; 102-15.

16. Silverman GA, Bird PI, Carrell RW, Church FC, Coughlin PB, Gettins PG, Irving JA, Lomas DA, Luke CJ, Moyer RW, Pemberton PA, Remold-O'Donnell E, Salvesen GS, Travis J, Whisstock JC. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *J Biol Chem* 2001; **276**; 33293-6.

17. Bode W, Huber R. Structural basis of the endoproteinase-protein inhibitor interaction. *Biochim Biophys Acta* 2000; **1477**; 241-52.

18. Cho YL, Chae YK, Jung CH, Kim MJ, Na YR, Kim YH, Kang SJ, Im H. The native metastability and misfolding of serine protease inhibitors. *Protein Pept Lett* 2005; **12**; 477-81.

19. Huntington JA, Read RJ, Carrell RW. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* 2000; **407**; 923-6.

20. Belorgey D, Hagglof P, Karlsson-Li S, Lomas DA. Protein misfolding and the serpinopathies. *Prion* 2007; **1**; 15-20.

21. Davies MJ, Lomas DA. The molecular aetiology of the serpinopathies. *Int J Biochem Cell Biol* 2008; **40**; 1273-86.

22. Cugno M, Zanichelli A, Foieni F, Caccia S, Cicardi M. C1-inhibitor deficiency and angioedema: molecular mechanisms and clinical progress. *Trends Mol Med* 2009; **15**; 69-78.

23. Quinsey NS, Greedy AL, Bottomley SP, Whisstock JC, Pike RN. Antithrombin: in control of coagulation. *Int J Biochem Cell Biol* 2004; **36**; 386-9.

24. Martinez-Martinez I, Ordonez A, Navarro-Fernandez J, Perez-Lara A, Gutierrez-Gallego R, Giraldo R, Martinez C, Llop E, Vicente V, Corral J. Antithrombin Murcia (K241E) causing antithrombin deficiency: a possible role for altered glycosylation. *Haematologica* 2010; **95**; 1358-65.

25. Picard V, Ersdal-Badju E, Bock SC. Partial glycosylation of antithrombin III asparagine-135 is caused by the serine in the third position of its N-glycosylation consensus sequence and is responsible for production of the beta-antithrombin III isoform with enhanced heparin affinity. *Biochemistry* 1995; **34**; 8433-40.

26. Olson ST, Chuang YJ. Heparin activates antithrombin anticoagulant function by generating new interaction sites (exosites) for blood clotting proteinases. *Trends Cardiovasc Med* 2002; **12**; 331-8.

27. McCoy AJ, Pei XY, Skinner R, Abrahams JP, Carrell RW. Structure of beta-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity. *J Mol Biol* 2003; **326**; 823-33.

28. Swedenborg J. The mechanisms of action of alpha- and beta-isoforms of antithrombin. *Blood Coagul Fibrinolysis* 1998; **9 Suppl 3**; S7-10.

29. Luxembourg B, Delev D, Geisen C, Spannagl M, Krause M, Miesbach W, Heller C, Bergmann F, Schmeink U, Grossmann R, Lindhoff-Last E, Seifried E, Oldenburg J, Pavlova A. Molecular basis of antithrombin deficiency. *Thromb Haemost* 2011; **105**; 635-46.

30. Berry LR, Thong B, Chan AK. Comparison of recombinant and plasma-derived antithrombin biodistribution in a rabbit model. *Thromb Haemost* 2009; **102**; 302-8.

31. Carlson TH, Atencio AC, Simon TL. Comparison of the behaviour in vivo of two molecular forms of antithrombin III. *Biochem J* 1985; **225**; 557-64.

32. Frebelius S, Isaksson S, Swedenborg J. Thrombin inhibition by antithrombin III on the subendothelium is explained by the isoform AT beta. *Arterioscler Thromb Vasc Biol* 1996; **16**; 1292-7.

33. Kamp P, Strathmann A, Ragg H. Heparin cofactor II, antithrombin-beta and their complexes with thrombin in human tissues. *Thromb Res* 2001; **101**; 483-91.

34. Swedenborg J. The mechanisms of action of alpha- and beta-isoforms of antithrombin. *Blood Coagul Fibrinolysis* 1998; **9 Suppl 3**; S7-10.

35. Witmer MR, Hatton MW. Antithrombin III-beta associates more readily than antithrombin III-alpha with uninjured and de-endothelialized aortic wall in vitro and in vivo. *Arterioscler Thromb* 1991; **11**; 530-9.

36. Stein PE, Carrell RW. What do dysfunctional serpins tell us about molecular mobility and disease? *Nat Struct Biol* 1995; **2**; 96-113.

37. Whisstock J, Skinner R, Lesk AM. An atlas of serpin conformations. *Trends Biochem Sci* 1998; **23**; 63-7.

38. Gils A, Declerck PJ. Structure-function relationships in serpins: current concepts and controversies. *Thromb Haemost* 1998; **80**; 531-41.

39. Wardell MR, Abrahams JP, Bruce D, Skinner R, Leslie AG. Crystallization and preliminary X-ray diffraction analysis of two conformations of intact human antithrombin. *J Mol Biol* 1993; **234**; 1253-8.

40. Johnson DJ, Langdown J, Li W, Luis SA, Baglin TP, Huntington JA. Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation. *J Biol Chem* 2006; **281**; 35478-86.

41. Carrell RW, Stein PE, Fermi G, Wardell MR. Biological implications of a 3 A structure of dimeric antithrombin. *Structure* 1994; **2**; 257-70.

42. Jin L, Abrahams JP, Skinner R, Petitou M, Pike RN, Carrell RW. The anticoagulant activation of antithrombin by heparin. *Proc Natl Acad Sci U S A* 1997; **94**; 14683-8.

43. Olson ST, Bjork I, Sheffer R, Craig PA, Shore JD, Choay J. Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement. *J Biol Chem* 1992; **267**; 12528-38.

44. Huntington JA. Mechanisms of glycosaminoglycan activation of the serpins in hemostasis. *J Thromb Haemost* 2003; **1**; 1535-49.

45. Langdown J, Belzar KJ, Savory WJ, Baglin TP, Huntington JA. The critical role of hinge-region expulsion in the induced-fit heparin binding mechanism of antithrombin. *J Mol Biol* 2009; **386**; 1278-89.

46. Petitou M, Duchaussoy P, Herbert JM, Duc G, El HM, Branellec JF, Donat F, Necciari J, Cariou R, Bouthier J, Garrigou E. The synthetic pentasaccharide fondaparinux: first in the class of antithrombotic agents that selectively inhibit coagulation factor Xa. *Semin Thromb Hemost* 2002; **28**; 393-402.

47. Izaguirre G, Rezaie AR, Olson ST. Engineering functional antithrombin exosites in alpha1-proteinase inhibitor that specifically promote the inhibition of factor Xa and factor IXa. *J Biol Chem* 2009; **284**; 1550-8.

48. Martinez-Martinez I, Ordonez A, Pedersen S, de la Morena-Barrio ME, Navarro-Fernandez J, Kristensen SR, Minano A, Padilla J, Vicente V, Corral J. Heparin affinity of factor VIIa: implications on the physiological inhibition by antithrombin and clearance of recombinant factor VIIa. *Thromb Res* 2011; **127**; 154-60.

49. Li W, Johnson DJ, Esmon CT, Huntington JA. Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat Struct Mol Biol* 2004; **11**; 857-62.

50. Pike RN, Buckle AM, le Bonniec BF, Church FC. Control of the coagulation system by serpins. Getting by with a little help from glycosaminoglycans. *FEBS J* 2005; **272**; 4842-51.

51. Im H, Ahn HY, Yu MH. Bypassing the kinetic trap of serpin protein folding by loop extension. *Protein Sci* 2000; **9**; 1497-502.

52. Corral J, Rivera J, Martinez C, Gonzalez-Conejero R, Minano A, Vicente V. Detection of conformational transformation of antithrombin in blood with

crossed immunoelectrophoresis: new application for a classical method. *J Lab Clin Med* 2003; **142**; 298-305.

53. Mushunje A, Evans G, Brennan SO, Carrell RW, Zhou A. Latent antithrombin and its detection, formation and turnover in the circulation. *J Thromb Haemost* 2004; **2**; 2170-7.

54. O'Reilly MS, Pirie-Shepherd S, Lane WS, Folkman J. Antiangiogenic activity of the cleaved conformation of the serpin antithrombin. *Science* 1999; **285**; 1926-8.

55. Bae JS, Rezaie AR. Mutagenesis studies toward understanding the intracellular signaling mechanism of antithrombin. *J Thromb Haemost* 2009; **7**; 803-10.

56. Guerrero JA, Teruel R, Martinez C, Arcas I, Martinez-Martinez I, de la Morena-Barrio ME, Vicente V, Corral J. Protective role of antithrombin in mouse models of liver injury. *J Hepatol* 2012;

57. Lane DA, Bayston T, Olds RJ, Fitches AC, Cooper DN, Millar DS, Jochmans K, Perry DJ, Okajima K, Thein SL, Emmerich J. Antithrombin mutation database: 2nd (1997) update. For the Plasma Coagulation Inhibitors Subcommittee of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost* 1997; **77**; 197-211.

58. Olds RJ, Lane DA, Chowdhury V, De Stefano V, Leone G, Thein SL. Complete nucleotide sequence of the antithrombin gene: evidence for homologous recombination causing thrombophilia. *Biochemistry* 1993; **32**; 4216-24.

59. Bock SC, Levitan DJ. Characterization of an unusual DNA length polymorphism 5' to the human antithrombin III gene. *Nucleic Acids Res* 1983; **11**; 8569-82.

60. Prochownik EV. Relationship between an enhancer element in the human antithrombin III gene and an immunoglobulin light-chain gene enhancer. *Nature* 1985; **316**; 845-8.

61. Fernandez-Rachubinski FA, Weiner JH, Blajchman MA. Regions flanking exon 1 regulate constitutive expression of the human antithrombin gene. *J Biol Chem* 1996; **271**; 29502-12.

62. Ochoa A, Brunel F, Mendelzon D, Cohen GN, Zakin MM. Different liver nuclear proteins binds to similar DNA sequences in the 5' flanking regions of three hepatic genes. *Nucleic Acids Res* 1989; **17**; 119-33.

63. Sheffield WP, Wu YI, Blajchman MA. Antithrombin: Structure and Function. In: High KA, Roberts HR, editors. Molecular Basis of Thrombosis and Hemostasis.New York: Marcel Dekker, Inc.; 1995. p. 355-77.

64. Souto JC, Almasy L, Borrell M, Gari M, Martinez E, Mateo J, Stone WH, Blangero J, Fontcuberta J. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation* 2000; **101**; 1546-51.

65. Anton AI, Teruel R, Corral J, Minano A, Martinez-Martinez I, Ordonez A, Vicente V, Sanchez-Vega B. Functional consequences of the prothrombotic SERPINC1 rs2227589 polymorphism on antithrombin levels. *Haematologica* 2009; **94**; 589-92.

66. Hernandez-Espinosa D, Ordonez A, Minano A, Martinez-Martinez I, Vicente V, Corral J. Hyperglycaemia impairs antithrombin secretion: Possible contribution to the thrombotic risk of diabetes. *Thromb Res* 2009;

67. Hernandez-Espinosa D, Mota R, Minano A, Ordonez A, Yelamos J, Vicente V, Corral J. In vivo effects of hyperthermia on the functional and conformational characteristics of antithrombin. *J Thromb Haemost* 2007; **5**; 963-70.

68. Hernandez-Espinosa D, Minano A, Martinez C, Perez-Ceballos E, Heras I, Fuster JL, Vicente V, Corral J. L-asparaginase-induced antithrombin type I deficiency: implications for conformational diseases. *Am J Pathol* 2006; **169**; 142-53.

69. Hernandez-Espinosa D, Ayala I, Castells MT, Garcia-Perez B, Martin-Castillo A, Minano A, Arcas I, Vicente V, Corral J. Intracellular retention of hepatic serpins caused by severe hyperlipidemia. *Liver Int* 2006; **26**; 708-15.

70. Bucciarelli P, Passamonti SM, Biguzzi E, Gianniello F, Franchi F, Mannucci PM, Martinelli I. Low borderline plasma levels of antithrombin, protein C and protein S are risk factors for venous thromboembolism. *J Thromb Haemost* 2012;

71. Ishiguro K, Kojima T, Kadomatsu K, Nakayama Y, Takagi A, Suzuki M, Takeda N, Ito M, Yamamoto K, Matsushita T, Kusugami K, Muramatsu T, Saito H. Complete antithrombin deficiency in mice results in embryonic lethality. *J Clin Invest* 2000; **106**; 873-8.

72. Lijfering WM, Brouwer JL, Veeger NJ, Bank I, Coppens M, Middeldorp S, Hamulyak K, Prins MH, Buller HR, Van Der Meer J. Selective testing for thrombophilia in patients with first venous thrombosis: results from a retrospective family cohort study on absolute thrombotic risk for currently known thrombophilic defects in 2479 relatives. *Blood* 2009; **113**; 5314-22.

73. Caspers M, Pavlova A, Driesen J, Harbrecht U, Klamroth R, Kadar J, Fischer R, Kemkes-Matthes B, Oldenburg J. Deficiencies of antithrombin, protein C and protein S - Practical experience in genetic analysis of a large patient cohort. *Thromb Haemost* 2012; **108**; 247-57.

74. Luxembourg B, Delev D, Geisen C, Spannagl M, Krause M, Miesbach W, Heller C, Bergmann F, Schmeink U, Grossmann R, Lindhoff-Last E, Seifried E, Oldenburg J, Pavlova A. Molecular basis of antithrombin deficiency. *Thromb Haemost* 2011; **105**; 635-46.

75. Aguila S, Martinez-Martinez I, Collado M, Llamas P, Anton AI, Martinez-Redondo C, Padilla J, Minano A, de la Morena-Barrio ME, Garcia-Avello A, Vicente V, Corral J. Compound heterozygosity involving Antithrombin

Cambridge II (p.Arg416Ser) in antithrombin deficiency. *Thromb Haemost* 2013; **109**

76. Corral J, Hernandez-Espinosa D, Soria JM, Gonzalez-Conejero R, Ordonez A, Gonzalez-Porras JR, Perez-Ceballos E, Lecumberri R, Sanchez I, Roldan V, Mateo J, Minano A, Gonzalez M, Alberca I, Fontcuberta J, Vicente V. Antithrombin Cambridge II (A384S): an underestimated genetic risk factor for venous thrombosis. *Blood* 2007; **109**; 4258-63.

77. Martinez-Martinez I, Navarro-Fernandez J, Ostergaard A, Gutierrez-Gallego R, Padilla J, Bohdan N, Minano A, Pascual C, Martinez C, de la Morena-Barrio ME, Aguila S, Pedersen S, Kristensen SR, Vicente V, Corral J. Amelioration of the severity of heparin-binding antithrombin mutations by posttranslational mosaicism. *Blood* 2012; **120**; 900-4.

78. Hernandez-Espinosa D, Ordonez A, Vicente V, Corral J. Factors with conformational effects on haemostatic serpins: implications in thrombosis. *Thromb Haemost* 2007; **98**; 557-63.

79. Corral J, Vicente V, Carrell RW. Thrombosis as a conformational disease. *Haematologica* 2005; **90**; 238-46.

80. Yamasaki M, Li W, Johnson DJ, Huntington JA. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 2008; **455**; 1255-8.

81. Raja SM, Chhablani N, Swanson R, Thompson E, Laffan M, Lane DA, Olson ST. Deletion of P1 arginine in a novel antithrombin variant (antithrombin London) abolishes inhibitory activity but enhances heparin affinity and is associated with early onset thrombosis. *J Biol Chem* 2003; **278**; 13688-95.

82. Martinez-Martinez I, Navarro-Fernandez J, Aguila S, Minano A, Bohdan N, de la Morena-Barrio ME, Ordonez A, Martinez C, Vicente V, Corral J. The infective polymerization of conformationally unstable antithrombin mutants may play a role in the clinical severity of antithrombin deficiency. *Mol Med* 2012; **18**; 762-70.

83. Maclean PS, Tait RC. Hereditary and acquired antithrombin deficiency: epidemiology, pathogenesis and treatment options. *Drugs* 2007; **67**; 1429-40.

84. Corral J, Huntington JA, Gonzalez-Conejero R, Mushunje A, Navarro M, Marco P, Vicente V, Carrell RW. Mutations in the shutter region of antithrombin result in formation of disulfide-linked dimers and severe venous thrombosis. *J Thromb Haemost* 2004; **2**; 931-9.

85. Corral J, Aznar J, Gonzalez-Conejero R, Villa P, Minano A, Vaya A, Carrell RW, Huntington JA, Vicente V. Homozygous deficiency of heparin cofactor II: relevance of P17 glutamate residue in serpins, relationship with conformational diseases, and role in thrombosis. *Circulation* 2004; **110**; 1303-7.

86. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**; 97-101.

87. Corral J, Gonzalez-Conejero R, Soria JM, Gonzalez-Porras JR, Perez-Ceballos E, Lecumberri R, Roldan V, Souto JC, Minano A, Hernandez-Espinosa D, Alberca I, Fontcuberta J, Vicente V. A nonsense polymorphism in the protein Z-dependent protease inhibitor increases the risk for venous thrombosis. *Blood* 2006; **108**; 177-83.

88. Sole X, Guino E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics* 2006; **22**; 1928-9.

89. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**; 1198-211.

90. Conlan MG, Folsom AR, Finch A, Davis CE, Marcucci G, Sorlie P, Wu KK. Antithrombin III: associations with age, race, sex and cardiovascular disease risk factors. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. *Thromb Haemost* 1994; **72**; 551-6.

91. Llop E, Gallego RG, Belalcazar V, Gerwig GJ, Kamerling JP, Segura J, Pascual JA. Evaluation of protein N-glycosylation in 2-DE: Erythropoietin as a study case. *Proteomics* 2007; **7**; 4278-91.

92. Mushunje A, Zhou A, Carrell RW, Huntington JA. Heparin-induced substrate behavior of antithrombin Cambridge II. *Blood* 2003; **102**; 4028-34.

93. Hernandez-Espinosa D, Minano A, Martinez C, Perez-Ceballos E, Heras I, Fuster JL, Vicente V, Corral J. L-asparaginase-induced antithrombin type I deficiency: implications for conformational diseases. *Am J Pathol* 2006; **169**; 142-53.

94. Mushunje A, Evans G, Brennan SO, Carrell RW, Zhou A. Latent antithrombin and its detection, formation and turnover in the circulation. *J Thromb Haemost* 2004; **2**; 2170-7.

95. Helander A, Husa A, Jeppsson JO. Improved HPLC method for carbohydrate-deficient transferrin in serum. *Clin Chem* 2003; **49**; 1881-90.

96. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; **3**; 285-98.

97. Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science* 2002; **297**; 1007-13.

98. Zhang XH, Chasin LA. Computational definition of sequence motifs governing constitutive exon splicing. *Genes Dev* 2004; **18**; 1241-50.

99. Di MM, Thys C, Waelkens E, Overbergh L, D'Hertog W, Mathieu C, De VR, Peerlinck K, Van GC, Freson K. An integrated proteomics and genomics analysis to unravel a heterogeneous platelet secretion defect. *J Proteomics* 2011; **74**; 902-13.

100. Di MM, Van GC, Freson K. Recent advances in platelet proteomics. *Expert Rev Proteomics* 2012; **9**; 451-66.

101. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 1996; **68**; 850-8.

102. Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS, Mische SM. Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis* 1999; **20**; 601-5.

103. Medina P, Navarro S, Estelles A, Vaya A, Woodhams B, Mira Y, Villa P, Migaud-Fressart M, Ferrando F, Aznar J, Bertina RM, Espana F. Contribution of polymorphisms in the endothelial protein C receptor gene to soluble endothelial protein C receptor and circulating activated protein C levels, and thrombotic risk. *Thromb Haemost* 2004; **91**; 905-11.

104. Roldan V, Ordonez A, Marin F, Zorio E, Soria JM, Minano A, Espana F, Gonzalez-Conejero R, Pineda J, Estelles A, Fontcuberta J, Vicente V, Corral J. Antithrombin Cambridge II (A384S) supports a role for antithrombin deficiency in arterial thrombosis. *Thromb Haemost* 2009; **101**; 483-6.

105. Severinsen MT, Johnsen SP, Tjonneland A, Overvad K, Dethlefsen C, Kristensen SR. Body height and sex-related differences in incidence of venous thromboembolism: a Danish follow-up study. *Eur J Intern Med* 2010; **21**; 268-72.

106. Medina P, Navarro S, Corral J, Zorio E, Roldan V, Estelles A, Santamaria A, Marin F, Rueda J, Bertina RM, Espana F. Endothelial protein C receptor polymorphisms and risk of myocardial infarction. *Haematologica* 2008; **93**; 1358-63.

107. Peyrard M, Seroussi E, Sandberg-Nordqvist AC, Xie YG, Han FY, Fransson I, Collins J, Dunham I, Kost-Alimova M, Imreh S, Dumanski JP. The human LARGE gene from 22q12.3-q13.1 is a new, distinct member of the glycosyltransferase gene family. *Proc Natl Acad Sci U S A* 1999; **96**; 598-603.

108. McCoy AJ, Pei XY, Skinner R, Abrahams JP, Carrell RW. Structure of beta-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity. *J Mol Biol* 2003; **326**; 823-33.

109. de la Morena-Barrio ME, Anton AI, Martinez-Martinez I, Padilla J, Minano A, Navarro-Fernandez J, Aguila S, Lopez MF, Fontcuberta J, Vicente V, Corral J. Regulatory regions of SERPINC1 gene: Identification of the first mutation associated with antithrombin deficiency. *Thromb Haemost* 2012; **107**; 430-7.

110. Jaeken J. Congenital disorders of glycosylation (CDG): it's (nearly) all in it! *J Inherit Metab Dis* 2011; **34**; 853-8.

111. Grünewald S, Schollen E, Van Schaftingen E, Jaeken J, Matthijs G. High residual activity of PMM2 in patients' fibroblasts: possible pitfall in the diagnosis of CDG-Ia (phosphomannomutase deficiency). *Am J Hum Genet* 2001; **68**; 347-54.

112. Matthijs G, Schollen E, Bjursell C, Erlandson A, Freeze H, Imtiaz F, Kjaergaard S, Martinsson T, Schwartz M, Seta N, Vuillaumier-Barrot S, Westphal V, Winchester B. Mutations in PMM2 that cause congenital disorders of glycosylation, type Ia (CDG-Ia). *Hum Mutat* 2000; **16**; 386-94.

113. Savignac M, Edir A, Simon M, Hovnanian A. Darier disease : a disease model of impaired calcium homeostasis in the skin. *Biochim Biophys Acta* 2011; **1813**; 1111-7.

114. Axelsson MA, Karlsson NG, Steel DM, Ouwendijk J, Nilsson T, Hansson GC. Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the O-glycosylation of mucins. *Glycobiology* 2001; **11**; 633-44.

115. Bjork I, Olson ST. Antithrombin. A bloody important serpin. *Adv Exp Med Biol* 1997; **425**; 17-33.

116. Kottke-Marchant K, Duncan A. Antithrombin deficiency: issues in laboratory diagnosis. *Arch Pathol Lab Med* 2002; **126**; 1326-36.

117. Austin H, De Staercke C., Lally C, Bezemer ID, Rosendaal FR, Hooper WC. New gene variants associated with venous thrombosis: a replication study in White and Black Americans. *J Thromb Haemost* 2011; **9**; 489-95.

118. Bezemer ID, Bare LA, Doggen CJ, Arellano AR, Tong C, Rowland CM, Catanese J, Young BA, Reitsma PH, Devlin JJ, Rosendaal FR. Gene variants associated with deep vein thrombosis. *JAMA* 2008; **299**; 1306-14.

119. Guerra-Shinohara EM, Bertinato JF, Tosin BC, Cordeiro da SK, Burlacchini de Carvalho MH, Pulcineli Vieira FR, Zugaib M, Cerda A, Morelli VM. Polymorphisms in antithrombin and in tissue factor pathway inhibitor genes are associated with recurrent pregnancy loss. *Thromb Haemost* 2012; **108**; 693-700.

120. Tremp GL, Duchange N, Branellec D, Cereghini S, Tailleux A, Berthou L, Fievet C, Touchet N, Schombert B, Fruchart JC, . A 700-bp fragment of the human antithrombin III promoter is sufficient to confer high, tissue-specific expression on human apolipoprotein A-II in transgenic mice. *Gene* 1995; **156**; 199-205.

121. Thomas DJ, Rosenbloom KR, Clawson H, Hinrichs AS, Trumbower H, Raney BJ, Karolchik D, Barber GP, Harte RA, Hillman-Jackson J, Kuhn RM, Rhead BL, Smith KE, Thakkapallayil A, Zweig AS, Haussler D, Kent WJ. The ENCODE Project at UC Santa Cruz. *Nucleic Acids Res* 2007; **35**; D663-D667.

122. Tait RC, Walker ID, Islam SI, McCall F, Conkie JA, Mitchell R, Davidson JF. Influence of demographic factors on antithrombin III activity in a healthy population. *Br J Haematol* 1993; **84**; 476-80.

123. Westrick RJ, Ginsburg D. Modifier genes for disorders of thrombosis and hemostasis. *J Thromb Haemost* 2009; **7 Suppl 1**; 132-5.

124. Oudot-Mellakh T, Cohen W, Germain M, Saut N, Kallel C, Zelenika D, Lathrop M, Tregouet DA, Morange PE. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol* 2012; **157**; 230-9.

125. Kanagawa M, Saito F, Kunz S, Yoshida-Moriguchi T, Barresi R, Kobayashi YM, Muschler J, Dumanski JP, Michele DE, Oldstone MB, Campbell KP. Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell* 2004; **117**; 953-64.

126. Aguilan JT, Sundaram S, Nieves E, Stanley P. Mutational and functional analysis of Large in a novel CHO glycosylation mutant. *Glycobiology* 2009; **19**; 971-86.

127. Patnaik SK, Stanley P. Mouse large can modify complex N- and mucin O-glycans on alpha-dystroglycan to induce laminin binding. *J Biol Chem* 2005; **280**; 20851-9.

128. Brockington M, Torelli S, Prandini P, Boito C, Dolatshad NF, Longman C, Brown SC, Muntoni F. Localization and functional analysis of the LARGE family of glycosyltransferases: significance for muscular dystrophy. *Hum Mol Genet* 2005; **14**; 657-65.

129. Hu Y, Li ZF, Wu X, Lu Q. Large induces functional glycans in an O-mannosylation dependent manner and targets GlcNAc terminals on alpha-dystroglycan. *PLoS One* 2011; **6**; e16866.

130. Zhang P, Hu H. Differential glycosylation of alpha-dystroglycan and proteins other than alpha-dystroglycan by like-glycosyltransferase. *Glycobiology* 2012; **22**; 235-47.

131. Inamori KI, Hara Y, Willer T, Anderson ME, Zhu Z, Yoshida-Moriguchi T, Campbell KP. Xylosyl- and glucuronyltransferase functions of LARGE in alpha-dystroglycan modification are conserved in LARGE2. *Glycobiology* 2012;

132. Aebi M, Bernasconi R, Clerc S, Molinari M. N-glycan structures: recognition and processing in the ER. *Trends Biochem Sci* 2010; **35**; 74-82.

133. Smith JW, Knauer DJ. A heparin binding site in antithrombin III. Identification, purification, and amino acid sequence. *J Biol Chem* 1987; **262**; 11964-72.

134. Jaeken J, Vanderschueren-Lodewyckx M, Snoeck L, Corbeel L, Wggermont E, Eeckels R. Familial psychomotor retardation with markedly fluctuating serum prolactin, FSH and GH levels, partial TBG deficiency, increased serum arylsulphatase A and increased CSF protein: a new syndrome). *Pediatr Res* 1980; **14**; 179.

135. Jaeken J, Matthijs G. Congenital disorders of glycosylation. *Annu Rev Genomics Hum Genet* 2001; **2**; 129-51.

136. Jaeken J. Congenital disorders of glycosylation. *Ann N Y Acad Sci* 2010; **1214**; 190-8.

137. Arnoux JB, Boddaert N, Valayannopoulos V, Romano S, Bahi-Buisson N, Desguerre I, de Keyzer Y., Munnich A, Brunelle F, Seta N, Dautzenberg MD, de Lonlay P. Risk assessment of acute vascular events in congenital disorder of glycosylation type Ia. *Mol Genet Metab* 2008; **93**; 444-9.

138. Young G, Driscoll MC. Coagulation abnormalities in the carbohydrate-deficient glycoprotein syndrome: case report and review of the literature. *Am J Hematol* 1999; **60**; 66-9.

139. Van Geet C, Jaeken J, Freson K, Lenaerts T, Arnout J, Vermylen J, Hoylaerts MF. Congenital disorders of glycosylation type Ia and IIa are associated with different primary haemostatic complications. *J Inherit Metab Dis* 2001; **24**; 477-92.

140. Jaeken J, Hennet T, Matthijs G, Freeze HH. CDG nomenclature: time for a change! *Biochim Biophys Acta* 2009; **1792**; 825-6.

141. Aebi M, Helenius A, Schenk B, Barone R, Fiumara A, Berger EG, Hennet T, Imbach T, Stutz A, Bjursell C, Uller A, Wahlstrom JG, Briones P, Cardo E, Clayton P, Winchester B, Cormier-Dalre V, de LP, Cuer M, Dupre T, et al. Carbohydrate-deficient glycoprotein syndromes become congenital disorders of glycosylation: an updated nomenclature for CDG. First International Workshop on CDGS. *Glycoconj J* 1999; **16**; 669-71.

142. Eklund EA, Freeze HH. The congenital disorders of glycosylation: a multifaceted group of syndromes. *NeuroRx* 2006; **3**; 254-63.

143. Burrows NP, Nicholls AC, Richards AJ, Luccarini C, Harrison JB, Yates JR, Pope FM. A point mutation in an intronic branch site results in aberrant splicing of COL5A1 and in Ehlers-Danlos syndrome type II in two British families. *Am J Hum Genet* 1998; **63**; 390-8.

144. Khan SG, Metin A, Gozukara E, Inui H, Shahlavi T, Muniz-Medina V, Baker CC, Ueda T, Aiken JR, Schneider TD, Kraemer KH. Two essential splice lariat branchpoint sequences in one intron in a xeroderma pigmentosum DNA repair gene: mutations result in reduced XPC mRNA levels that correlate with cancer risk. *Hum Mol Genet* 2004; **13**; 343-52.

145. Vuillaumier-Barrot S, le BC, de LP, Madinier-Chappat N, Barnier A, Dupre T, Durand G, Seta N. PMM2 intronic branch-site mutations in CDG-Ia. *Mol Genet Metab* 2006; **87**; 337-40.

146. Jones MA, Ng BG, Bhide S, Chin E, Rhodenizer D, He P, Losfeld ME, He M, Raymond K, Berry G, Freeze HH, Hegde MR. DDOST mutations identified by whole-exome sequencing are implicated in congenital disorders of glycosylation. *Am J Hum Genet* 2012; **90**; 363-8.

147. Landberg E, Pahlsson P, Lundblad A, Arnetorp A, Jeppsson JO. Carbohydrate composition of serum transferrin isoforms from patients with high alcohol consumption. *Biochem Biophys Res Commun* 1995; **210**; 267-74.

148. Cottalasso D, Bellocchio A, Pronzato MA, Domenicotti C, Traverso N, Gianelli MV, Marinari UM, Nanni G. Effect of ethanol administration on the level of dolichol in rat liver microsomes and Golgi apparatus. *Alcohol Clin Exp Res* 1998; **22**; 730-7.

149. Flahaut C, Michalski JC, Danel T, Humbert MH, Klein A. The effects of ethanol on the glycosylation of human transferrin. *Glycobiology* 2003; **13**; 191-8.

150. Landberg E, Pahlsson P, Lundblad A, Arnetorp A, Jeppsson JO. Carbohydrate composition of serum transferrin isoforms from patients with high alcohol consumption. *Biochem Biophys Res Commun* 1995; **210**; 267-74.

151. Guizzetti M, Chen J, Oram JF, Tsuji R, Dao K, Moller T, Costa LG. Ethanol induces cholesterol efflux and up-regulates ATP-binding cassette cholesterol transporters in fetal astrocytes. *J Biol Chem* 2007; **282**; 18740-9.

152. Krebs HA, Freedland RA, Hems R, Stubbs M. Inhibition of hepatic gluconeogenesis by ethanol. *Biochem J* 1969; **112**; 117-24.

153. Sharma V, Ichikawa M, He P, Scott DA, Bravo Y, Dahl R, Ng BG, Cosford ND, Freeze HH. Phosphomannose isomerase inhibitors improve N-glycosylation in selected phosphomannomutase-deficient fibroblasts. *J Biol Chem* 2011; **286**; 39431-8.

Spanish abstract

Resumen en castellano I

La importancia del sistema hemostático en la enfermedad tromboembólica, la principal causa de morbimortalidad de nuestra sociedad, ha incentivado la búsqueda de alteraciones genéticas asociadas con variaciones en el funcionamiento del sistema hemostático, empleando diferentes diseños experimentales. Gran parte de estos estudios se han centrado en los propios elementos de dicho sistema, con resultados bastante frustrantes. La elevada variabilidad interindividual de los fenotipos hemostáticos, junto con la alta heredabilidad de los mismos, sugieren la posible existencia de variabilidad genética en el gen que codifica dicho fenotipo. También alteraciones genéticas en otros factores podrían modular de forma indirecta los niveles de ese elemento. En cualquier caso, estas alteraciones pueden jugar un papel relevante en la funcionalidad del sistema hemostático y en el riesgo trombótico.

La antitrombina es un miembro de la superfamilia de las serpinas, inhibidores de serín proteasas, clave en el control de la hemostasia. Su amplio espectro de proteasas procoagulantes diana, que abarca moléculas tan relevantes como el FXa o la trombina, pero también otras como el FVII, FIXa, FXIa y FXIIa entre otras, y el mecanismo de inhibición tan potente y eficaz, hacen que la antitrombina sea el principal anticoagulante endógeno del organismo. Por ello, cualquier factor que afecte a esta molécula es susceptible de desequilibrar el sistema hemostático, y podría contribuir a incrementar el riesgo trombótico. La búsqueda de estos elementos se ha centrado clásicamente en el análisis molecular del gen codificante: *SERPINC1*. Así desde que en 1965 O Egeberg describió la primera familia con deficiencia de antitrombina asociada a una alta incidencia de trombosis venosa, se han identificado más de 296 mutaciones diferentes en el gen *SERPINC1* en pacientes con trombosis venosa y deficiencia de antitrombina. Todas estas mutaciones afectan la región codificante o secuencias de procesamiento de intrones y explican, por distintos mecanismos, los tipos de deficiencia de antitrombina. Sin embargo, es sorprendente la ausencia de modificaciones genéticas que a través de mecanismos regulatorios se asocian con deficiencia de antitrombina. Probablemente las mutaciones que afecten a estas regiones son poco frecuentes o tienen poca relevancia funcional. Por otra parte, no se conocen las regiones promotora y/o reguladoras del gen *SERPINC1*. Hasta la fecha, solo un estudio de nuestro grupo ha descrito un polimorfismo localizado en una zona no codificante, en el intrón 1, que regula moderadamente los niveles de antitrombina.

Por el contrario, apenas hay estudios que hayan buscado otros elementos fuera del gen codificante, capaces de afectar los niveles o funcionalidad de este relevante

anticoagulante. Este dato junto al reciente trabajo que demuestra que la reducción incluso moderada de los niveles de antitrombina también incrementa el riesgo trombótico, anima a buscar nuevos mecanismos que potencialmente afecten la funcionalidad o los niveles de antitrombina.

El objetivo principal de este estudio fue identificar nuevos genes y modificaciones genéticas involucradas en la regulación de los niveles o funcionalidad de antitrombina.

Para llevar a cabo este objetivo utilizamos tres aproximaciones diferentes.

1. Identificación de nuevas regiones reguladoras del gen *SERPINC1* mediante la secuenciación completa del gen (15,000 pb, abarcando todos los exones e intrones de este gen, así como las regiones 3' y 5') en tres grupos de sujetos con diferentes niveles de antitrombina.

2. Identificación de nuevos genes moduladores de antitrombina empleando un estudio de asociación de polimorfismos que abarcan el genoma completo (Genome Wide Association Study, GWAS) y que se complementó con diferentes estudios experimentales, incluyendo silenciamiento del gen identificado.

3. Identificación de nuevos mecanismos responsables de la deficiencia de antitrombina, mediante la caracterización bioquímica de la antitrombina plasmática de pacientes con deficiencia de antitrombina.

Para el desarrollo del proyecto hemos empleado metodología molecular que abarca desde genotipado de alteraciones puntuales empleando diferentes sistemas (sondas taqman, PCR-ASRA, etc), hasta un GWAS, pasando por la cuantificación de expresión y la secuenciación de diferentes genes (*SERPINC1*, *LARGE*, *PMM2*). También hemos usado metodología bioquímica clásica (electroforesis en geles desnaturalizantes y nativos, purificación o separación de proteínas mediante FPLC o HPLC, estudios de inhibición, etc), pero también metodología más avanzada (proteómica y glicómica). Finalmente, en este proyecto también hemos aplicado metodología celular que incluye inmunofluorescencia, estudios de reporter, o silenciamiento génico.

Desde el punto de vista metodológico también debemos destacar el amplio espectro de pacientes y controles estudiados. Así, el proyecto ha incluido una cohorte familiar (estudio GAIT), 307 donantes de sangre sanos, tres estudios caso-control de

diferentes enfermedades tromboembólicas (infarto agudo de miocardio y trombosis venosa) de dos países diferentes (España y Dinamarca) con 2,980 pacientes y 3,996 controles. Destacamos la inclusión de 118 pacientes con deficiencia de antitrombina, grupo del que seleccionamos 25 pacientes que con deficiencia (moderada o severa) de antitrombina no tenían mutaciones en el gen *SERPINC1* y han sido la base de dos de los apartados de esta tesis.

Finalmente estudiamos también 2 pacientes con trastorno de glicosilación congénita (CDG), un desorden muy raro y un paciente con severo alcoholismo, cruciales para el último apartado de nuestro estudio.

Los resultados de este estudio muestran una heterogeneidad de nuevos mecanismos implicados en la regulación de los niveles de antitrombina con potenciales consecuencias patológicas.

1) El análisis del gen *SERPINC1* completo en sujetos con deficiencia de antitrombina pero sin mutaciones en la zona codificante de dicho gen identificó la primera mutación en la región promotora del gen asociada con deficiencia de antitrombina. El cambio g.2143 C>G identificado en la familia 2 no impide la transcripción del gen pero sí la reduce de forma significativa, como muestran los niveles de antitrombina en portadores y los estudios con reporter de luciferasa en diferentes líneas celulares. El cambio g.2143 C>G se encuentra en una de las cuatro regiones protegidas descritas en dos estudios independientes; y se ha sugerido que en esta región podría interaccionar el factor nuclear hepático 4 (HNF-4) regulando la transcripción del gen. Además, en un paciente con moderada deficiencia de antitrombina identificamos otra nueva mutación g.1091 C>T situada a 1,052 pb del codón ATG, con potencial efecto regulatorio. Son necesarios estudios familiares y de reporter de luciferasa para determinar la relevancia funcional de esta segunda alteración en la regulación del gen.

Estos resultados sugieren incluir el estudio de la región promotora del gen *SERPINC1* en la caracterización molecular de la deficiencia de antitrombina, concretamente en aquéllos casos con deficiencia tipo I y sin mutaciones en las regiones codificantes de *SERPINC1*.

Sin embargo, es probable la existencia de otros elementos reguladores aguas arriba o abajo del gen *SERPINC1*. La identificación de una familia (F1) y un paciente (P26) pertenecientes a la misma área geográfica, con deficiencia tipo I y trombosis

severas y sin mutaciones en las 15,925 pb analizadas de *SERPINC1*, alienta a continuar la búsqueda de nuevas regiones reguladoras.

La búsqueda de esas regiones reguladoras también la hemos realizado con otro diseño experimental, centrándonos en población general ya que los niveles de este anticoagulante tienen una alta variabilidad y existe un elevado grado de heredabilidad para este fenotipo (h=0.486) lo que sustenta la existencia de factores genéticos implicados. Estudiamos los niveles de antitrombina en una muestra de 307 sujetos sanos. Como se esperaba encontramos una distribución normal (70-130%). Seleccionamos cinco sujetos con niveles extremos de antitrombina dentro del rango de la normalidad (75% y 115%) y secuenciamos el gen *SERPINC1* completo con el objetivo de identificar variaciones genéticas que se asocien con esa gran variabilidad. Las alteraciones identificadas con potencial efecto funcional fueron verificadas en una muestra mayor. Los resultados confirmaron el moderado efecto que el polimorfismo rs2227589 del intrón 1 tiene sobre los niveles de antitrombina. También identificamos un polimorfismo de la región promotora g.2085 T>C que *in silico* modificaba potencialmente la unión de distintos factores de transcripción. Sin embargo, el estudio de asociación de actividad y frecuencia de este cambio en 131 sujetos sanos demostró no estar asociado a los niveles de antitrombina.

La disección de la arquitectura del gen *SERPINC1* en sujetos sanos con niveles extremos de antitrombina reveló dos importantes hallazgos: A) La variabilidad genética de este gen es muy baja. Únicamente encontramos 22 polimorfismos, y solo 2 en la región codificante, ambos ligados y sin que modifiquen la secuencia aminoacídica. Estos resultados reflejan la sensibilidad conformacional y funcional de esta molécula. B) La escasa importancia de esta variabilidad genética en el control de los niveles de antitrombina.

Estos resultados junto al alto grado de heredabilidad de los niveles de antitrombina apoyan la existencia de genes moduladores. Estos genes podrían codificar factores transcripcionales, moléculas implicadas en el plegamiento o secreción de antitrombina, enzimas implicadas en modificaciones post-traduccionales, proteasas, miRNAs…

2) Por ello, nuestro segundo objetivo fue buscar nuevos genes moduladores de antitrombina. Para ello empleamos una aproximación de asociación genotipo-fenotipo analizando el genoma completo (GWAS). Esta misma aproximación se acaba de publicar recientemente por un grupo francés. Estos estudios se diferenciaban en la

población estudiada: pacientes con trombosis (estudio francés) o nuestro estudio familiar. Pero en ambos estudios, el GWAS no identificó ningún polimorfismo asociado con los niveles de antitrombina con significación estadística (p< 10E-07). Este resultado, junto a la elevada heredabilidad, sugería la existencia de elementos genético con poco peso, por lo que nosotros validamos los 10 SNPs con mayor asociación estadística (p< 4x10E-05) y dos SNPs adicionales del gen *LARGE*, un gen que codifica una glicosiltransferasa que se encontraba en la lista de los 10 genes mas asociados con los niveles de antitrombina. Precisamente un SNP de *LARGE*, rs762057, mantuvo la asociación con los niveles de anti-FXa en plasma en los 307 donantes de sangre de otra región española que se emplearon como población de validación. Tanto en el estudio de validación como en el GWAS, el alelo polimórfico incrementaba moderadamente los niveles de antitrombina (2-4%). Este ligero aumento de la actividad de antitrombina explica la ausencia de relevancia trombótica de este polimorfismo, según reflejaron los resultados del estudio de 849 pacientes con trombosis venosa y 862 controles. Además, encontramos que el haplotipo H2 del gen *LARGE* definido con los tres polimorfismos rs762057, rs713703 y rs240082 también se asociaba de forma significativa con los niveles de antitrombina (p=0.030). Para consolidar esta asociación y definir el mecanismo de la misma, realizamos experimentos de silenciamiento del gen *LARGE* en dos líneas celulares (HepG2, línea hepática y con expresión constitutiva de antitrombina y HEK-EBNA transfectada con cDNA de antitrombina humana). El silenciamiento de *LARGE* redujo la secreción de antitrombina en ambas líneas celulares confirmando el papel de *LARGE* en la modulación de los niveles de antitrombina. El mecanismo por el que actúa LARGE, una enzima con actividades xilosiltransferasa y glucuroniltransferasa, sobre los niveles de antitrombina tiene todavía que explorarse, pero el hecho de que esta enzima genere moléculas parecidas a heparina y que la antitrombina presente un domino de unión a heparina, sustenta esta asociación y sugieren que LARGE podría estar implicado en el correcto plegamiento o tráfico intracelular de la antitrombina.

3) La caracterización bioquímica de la antitrombina plasmática en pacientes con deficiencia de antitrombina sin mutaciones en *SERPINC1* ha sido crucial para identificar un nuevo mecanismo implicado en la regulación de los niveles de antitrombina y con gran riesgo trombótico. Este particular factor trombofílico afecta una modificación postraduccional, la *N*-glicosilación. El paciente P27 tuvo cuatro episodios

trombóticos en edades tempranas y una significativa deficiencia de antitrombina como único factor trombofílico. La identificación de diferentes glicoformas de antitrombina mediante análisis electroforético seguido de diferentes estudios bioquímicos y un profundo análisis de secuenciación del gen *PMM2*, gen que codifica la enzima fosfomanomutasa 2, implicada en los primeros pasos de formación del *N*-glicano precursor que se incorporará a *N*-glicoproteínas nacientes, permitió el diagnóstico final de un trastorno de glicosilación congénita tipo Ia (CDGIa o también llamado PMM2-CDG). Además, este nuevo factor trombofílico, más bien una variante, podría ser relativamente frecuente. El análisis electroforético de antitrombina y antitripsina y el estudio de glicoformas de transferrina mediante HPLC y Q-TOF en 25 pacientes con trombosis y ligera deficiencia de antitrombina sin mutaciones en *SERPINC1*, identificó cuatro pacientes con un patrón igual pero más suave que el de CDGs. En tres casos se identificó una mutación en heterocigosis en *PMM2*. Estos resultados sugieren la existencia de un nuevo desorden al que hemos llamado CDG-like. Como los cuatro pacientes tienen en común una leve afectación hepática como consecuencia de la ingesta de alcohol, especulamos que esta combinación podría explicar el fenotipo CDG y la clínica trombótica. En su conjunto estos resultados apoyan que los trastornos de glicosilación pudieran considerarse como factor trombofílico que debería estudiarse, especialmente en pacientes con moderada deficiencia de antitrombina.

Nuestro estudio muestra nuevo mecanismos implicados en la variabilidad de la antitrombina, el principal anticoagulante endógeno. Además, abre nuevas líneas de trabajo y deja cuestiones que aún quedan por resolver: la existencia de nuevas regiones reguladoras en el gen *SERPINC1* que aún deben ser identificadas, la existencia de otros genes moduladores de antitrombina que podrían tener relevancia trombótica, la incidencia de trastornos de glicosilación en trombosis, y el mecanismo por el cual estos trastornos (PMM2-CDG y CDG-like) elevan el riesgo trombótico.

Appendix I

Diffusion of scientific results related with the Thesis

Results from the present Thesis have been published or are under review in peer-reviewed journals.

A. Published:

1. **de la Morena-Barrio ME**, Anton AI, Martinez-Martinez I, Padilla, J, Minano A, Navarro-Fernández J, Águila S, López MF, Fontcuberta J, Vicente V, and Corral J. Regulatory regions of *SERPINC1* gene: identification of the first mutation associated with antithrombin deficiency. **Thromb. Haemost**. 107(3):430-7; 2012. Impact Factor: 5.044.

2. **de la Morena-Barrio ME**, Sevivas T, Martínez-Martínez I, Miñano A, Vicente V, Jaeken J, Corral J. Congenital disorder of glycosylation (PMM2-CDG) in a patient with antithrombin deficiency and severe thrombophilia. **J Thromb Haemost**. 2012 Oct 20. doi: 10.1111/jth.12031. Impact factor: 5.731.

B. Submitted:

1. **de la Morena-Barrio ME**, Buil A, Antón AI, Martínez-Martínez I, Miñano A, Gutiérrez-Gallego R, Navarro-Fernández J, Águila S, Souto JC, Vicente V, Soria JM, Corral J. Identification of antithrombin-modulating genes. Role of *LARGE*, a gene encoding a bifunctional glycosyltransferase, in the secretion of proteins?. **PLoS ONE.** Impact factor: 4.092

Regulatory regions of SERPINC1 gene: Identification of the first mutation associated with antithrombin deficiency

María Eugenia de la Morena-Barrio[1]; Ana Isabel Antón[1]; Irene Martínez-Martínez[1]; José Padilla[1]; Antonia Miñano[1]; José Navarro-Fernández[1]; Sonia Águila[1]; María Fernanda López[2]; Jordi Fontcuberta[3]; Vicente Vicente[1]; Javier Corral[1]

[1]Centro Regional de Hemodonación, University of Murcia, Spain; [2]Unidad de Hemostasia y Trombosis del Complejo Hospitalario Universitario, A Coruña, Spain; [3]Unitat d'Hemostàsia i Trombosi, Departament d'Hematologia, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

Summary

Antithrombin is the main endogenous anticoagulant. Impaired function or deficiency of this molecule significantly increases the risk of thrombosis. We studied the genetic variability of SERPINC1, the gene encoding antithrombin, to identify mutations affecting regulatory regions with functional effect on its levels. We sequenced 15,375 bp of this gene, including the potential promoter region, in three groups of subjects: five healthy subjects with antithrombin levels in the lowest (75%) and highest (115%) ranges of our population, 14 patients with venous thrombosis and a moderate antithrombin deficiency as the single thrombophilic defect, and two families with type I antithrombin defi- ciency who had neither mutations affecting exons or flanking regions, nor gross gene deletions. Our study confirmed the low genetic variability of SERPINC1, particularly in the coding region, and its minor influence in the heterogeneity of antithrombin levels. Interestingly, in one family, we identified a g.2143 C>G transversion, located 170 bp upstream from the translation initiation codon. This mutation affected one of the four regions located in the minimal promoter that have potential regulatory activity according to previous DNase footprinting protection assays. Genotype-phenotype analysis in the affected family and reporter analysis in different hepatic cell lines demonstrated that this mutation significantly impaired, although it did not abolish, the downstream transcription. Therefore, this is the first mutation affecting a regulatory region of the SERPINC1 gene associated with antithrombin deficiency. Our results strongly sustain the inclusion of the promoter region of SERPINC1 in the molecular analysis of patients with antithrombin deficiency.

Keywords

SERPINC1, antithrombin, promoter, polymorphisms, thrombosis

Correspondence to:
Dr. Javier Corral
University of Murcia
Centro Regional de Hemodonación
Ronda de Garay S/N. Murcia 30003, Spain
Tel.: +34 968341990, Fax: +34 968261914
E-mail: javier.corral@carm.es

Financial support:
This work was supported by 04515 /GERM/ 06 (Fundación Séneca), SAF2009–08993 (MCYT & FEDER), Fundación Mutua Madrileña, and RECAVA RD06/0014/0039 & RD06/0014/0016 (ISCIII & FEDER). MEMB is a holder of a predoctoral research grant from ISCIII (FI09/00190). IMM is a researcher from Fundación para la Formación e Investigación Sanitarias. JNF is postdoctoral researcher of the University of Murcia.

Received: October 11, 2011
Accepted after minor revision: November 29, 2011
Prepublished online: January 11, 2012

doi:10.1160/TH11-10-0701
Thromb Haemost 2012; 107: ■■■

Introduction

Antithrombin is an anticoagulant serpin essential for the haemostatic system. This plasma protein synthesised in the liver has an efficient mechanism of inhibition of multiple procoagulant proteases, including thrombin and factor (F) Xa (1, 2). Accordingly, complete deficiency of antithrombin has lethal consequences and the heterozygous deficiency significantly increases (10– to 40-fold) the risk of thrombosis (3). Thus, the screening of antithrombin deficiency evaluating the activity of antithrombin in plasma is included among the thrombophilic tests (4).

SERPINC1, the gene encoding antithrombin (GenBank X68793.1) was mapped on chromosome 1q23–25.1 and comprises seven exons, encompassing 13.5 kb of genomic DNA (5, 6). More than 235 different gross gene deletions and mutations affecting the coding or splicing regions of SERPINC1 have been identified in patients with antithrombin deficiency (http://www.hgmd.cf.ac.uk/ac/gene.php?gene=SERPINC1). Different mechanisms, from RNA instability to conformational or functional consequences have been proposed to explain the associated quantitative or qualitative deficiency (3, 7). However, no mutation has been so far identified in regulatory regions. Actually, the mechanisms underlying antithrombin gene expression are not well known, the promoter lacks of TATA element or GC-rich regions (6, 8, 9) and few studies, most of them using deletional analysis and reporter or DNase I footprint assays, have identified few regions potentially involved in the transcriptional regulation of this gene (9–12).

The aim of this study was the identification of gene variations affecting regulatory elements of SERPINC1 by sequencing this gene in three groups of subjects with different expression of antithrombin.

Material and methods

Subjects and blood sampling

Our study included three groups of subjects:

- Three hundred seven Spanish Caucasian healthy blood donors (138 male/169 female), with an average age of 43 years.
- Fourteen thrombophilic patients (named as P1-P14) who had a moderate antithrombin deficiency (70–90% of the reference value) as the single thrombophilic risk factor identified.
- Two families (F1 and F2) with at least three members from two generations that repetitively had type I antithrombin deficiency (**Fig. 2**). These families were selected from 64 consecutive subjects with congenital antithrombin deficiency referenced to our laboratory from different Spanish and Portuguese hospitals. Selection was based on the absence of gross deletion, mutation in the coding region, or splicing site mutations.

Blood samples were obtained by venopuncture collection into 1:10 volume of trisodium citrate. Platelet-poor plasma fractions were obtained (within 5 minutes [min] after blood collection) by centrifugation at 4 ºC for 20 min at 2,200 x g and stored at –70ºC. Genomic DNA was purified by the salting out procedure and stored at –20ºC.

All subjects gave their informed consent to enter the study, which was approved by the ethics committee of the Centro Regional de Hemodonación (Murcia, Spain) and performed according to the declaration of Helsinki, as amended in Edinburgh in 2000.

Antithrombin levels and other thrombophilic tests

Routine thrombophilic tests including protein C, protein S, anti phospholipid antibodies, FV Leiden, and prothrombin G20210A were performed, as described elsewhere (13). FXa-inhibiting activity from plasma of patients and family members was measured using a chromogenic method (HemosIL TH, Instrumentation Laboratory, Milan, Italy). Antithrombin antigen levels in plasma were determined by rocket immunoelectrophoresis. For both parameters, values were expressed as a percentage of the result observed in a pool of citrated plasma from 100 healthy subjects.

Amplification and sequencing

Polymerase chain reactions (PCR) covering the whole SERPINC1 gene and the potential promoter region were carried out using Expand Long Template Polymerase (Roche, Madrid, Spain) and the oligonucleotide sets described in **Suppl. Table 1** (available online at www.thrombosis-online.com). PCR products were purified and sequenced with ABI Prism Big Dye Terminator v3.1 Cycle sequencing kit and resolved on a 3130xl Genetic Analyzer (Applied Biosystems, Madrid, Spain). Comparison with the reference sequence (GenBank accession number NG_012462.1) was performed with SeqScape v2.5 software (Applied Biosystems).

Genotyping methods

The genotype of rs5778785 was determined by high resolution melting (HRM) analysis following the Type-it HRM PCR protocol (Qiagen, Madrid, Spain). The g.2085 T>C transition was genotyped by a HRM/razor probe technique. The genotype was verified by sequencing, with complete agreement. These two HRM analyses were carried out in Rotor-Gene Q 6000 (Qiagen) with the primers and probes shown in Suppl. Table 2 (available online at www.thrombosis-online.com).

Nine polymorphisms were genotyped by PCR-allelic specific restriction assays using primers and restriction enzymes described in **Table 1**, as previously described (14). For some polymorphisms, mutated primers were designed (see Suppl. Table 2, available online at www.thrombosis-online.com).

rs2227589 was genotyped by a Taqman SNP Genotyping Assay (C__16180170_20, Applied-Biosystem), as previously described (16).

The g.2143 C>G mutation was studied by PCR-allelic specific restriction assay using: 5'-GGAGTCCTTGATCACACAGCA-3' forward and 5'-GTCTTTGACTGTAACTACCAG-3' reverse primers and Bsa JI restriction enzyme (New England BioLabs, Barcelona, Spain) following manufacturer's specifications. Results were confirmed by sequencing.

Prediction of transcription factor binding sites

The search for sequences with potential transcriptional relevance affected by genetic variations identified in this study was performed by TESS and TFSEARCH programs, based on TRANSFAC databases (http://www.cbil.upenn.edu/cgi-bin/tess/tessand http://www.cbrc.jp/research/db/TFSEARCH).

Cloning and reporter assay

The 5' non transcribed portion of SERPINC1 from one individual heterozygous for the g.2143 C>G mutation was amplified using Expand High Fidelity DNA polymerase (Roche) to minimise nucleotides misincorporation with 5'-GGAGTCCTTGATCACACAGCA-3' (forward) and 5'-AATCTCGCAGAGGTTCCAGA-3' (reverse) primers. PCR products were ligated into a T/A cloning vector (pCR®2.1, Invitrogen, Barcelona, Spain). Clones with wildtype and mutant alleles were selected by PCR-allelic restriction assay with Bsa JI, and the inserts cloned into Kpn I and Sac I sites of luciferase pGL3 basic vector (Promega, Madrid, Spain). Compet-

Table 1: Plasma anti-FXa values of blood donors according to the genotype of SERPINC1 polymorphisms identified in five selected healthy subjects with extreme antithrombin levels and 14 patients with low antithrombin levels.

Polymorphism	Subject	Genotyping method	N*	Frequency		Anti-FXa activity (mean ± SD)	P°
				PGA European#	Our study		
g.2085 T>C (rs78972925)	H1	HRM/Razor probe & sequencing	131	No data	T:0.973	WT:96 ± 9	0.375
					C:0.027	Pol:99 ± 11	
rs2227588	P1 & P7	PCR-ASRA: Mnl I	126	A:0.875	A:0.921	WT:95 ± 8	0.342
				G:0.125	G:0.079	Pol:93 ± 9	
rs2227589	L1	Taqman probe	305	G:0.975	G:0.898	WT:97 ± 7	0.037
				A:0.043	A:0.102	Pol:95 ± 8	
rs2227590	L2	PCR-ASRA: Fok I	128	C:0.870	C:0.922	WT:95 ± 8	0.793
				T:0.130	T:0.078	Pol:94 ± 9	
rs2227596	H1	PCR-ASRA: Tsp 451	127	A:0.795	A:0.886	WT:94 ± 8	0.129
				G:0.205	G:0.114	Pol:97 ± 9	
rs2227603	H1	PCR-ASRA: Fok I	101	T:0.935	T:0.990	WT:94 ± 8	0.317
				G:0.065	G:0.010	Pol:100 ± 19	
rs941988	L1	PCR-ASRA: Hinf I	97	G:0.955	G:0.878	WT:97 ± 7	0.513
				A:0.045	A:0.122	Pol:95 ± 8	
rs2227607	H1	PCR-ASRA: Nla III	138	G:0.972	G:0.895	WT:95 ± 8	0.903
				T:0.028	T:0.105	Pol:95 ± 9	
rs5778785	H1	HRM & sequencing	106	No data	WT:0.896	WT:95 ± 8	0.838
					INS:0.104	Pol:95 ± 9	
rs2759328	L1	PCR-ASRA: Fau I	149	G:0.917	G:0.850	WT:95 ± 8	0.736
				A:0.083	A:0.150	Pol:94 ± 8	
rs677	L2	PCR-ASRA: Dde I	158	G:0.842	G:0.930	WT:95 ± 8	0.604
				C:0.158	C:0.070	Pol:94 ± 9	
rs2227616	H1	PCR-ASRA: Bst EII	183	A:0.935	A:0.880	WT:94 ± 7	0.923
				(-):0.065	(-):0.120	Pol:94 ± 8	

* N: Number of healthy blood donors studied. # Frequencies from http://www.ncbi.nlm.nih.gov/snp.° p-value obtained from the student-t statistical analysis following a dominant model. WT: Homozygous wild type genotype; Pol: Carriers of the polymorphism (heterozygous + homozygous). SD: standard deviation. rs941988 and rs2227592 were completely linked, according to the Haploview 4.1. The genotypes of all studied polymorphisms were in Hardy-Weinberg equilibrium.

ent cells were transformed, plasmid DNA was purified using a commercial kit (Plasmid Midi Kit, Qiagen), and they were sequenced in both directions to faithfully confirm the wild-type and mutant sequences.

The ability of each sequence to promote transcription of luciferase gene was transiently tested in human cell lines PLC-PRF-5 (human hepatome cell line) and HepG2 (human hepatocellular liver carcinoma cell line) both with antithrombin constitutive expression.

All plasmid DNA were quantified by Nanodrop (Thermo Scientific, Madrid, Spain). Eighteen hours before transfection, cells were plated at 80,000 per well of a 24-well plated (three replicates per clone for each cell line and in each experiment) and cultured in Dulbecco's Modified Eagle Medium (Invitrogen), at 37ºC with 5% CO2. Reporter plasmid (1,000 ng) along with 100 ng of Renilla luciferase control plasmid pRL-TK (Promega) were transfected by using Lipofectamine 2000 (Invitrogen) for PLC-PRF-5 cell line and GenJet reagent (SignaGen® Laboratories) for HepG2 cell line, according to the manufacturer's instructions. Luciferase assays were performed 48 h after transfection by using the Dual-GloTM Luciferase assay System (Promega). Promoter activity was normalised by dividing luciferase activity by Renilla luciferase activity for each transfected well. pGL3 control vector (Promega) was used as a transfection positive control.

Statistical analyses

Allele and genotype frequencies, deviations from Hardy-Weinberg expectations, haplotype analysis, association of haplotypes with anti-FXa activity and linkage disequilibrium analysis were calculated with the SNPstats software (15).

Statistical analysis was performed with the SPSS software (ver-

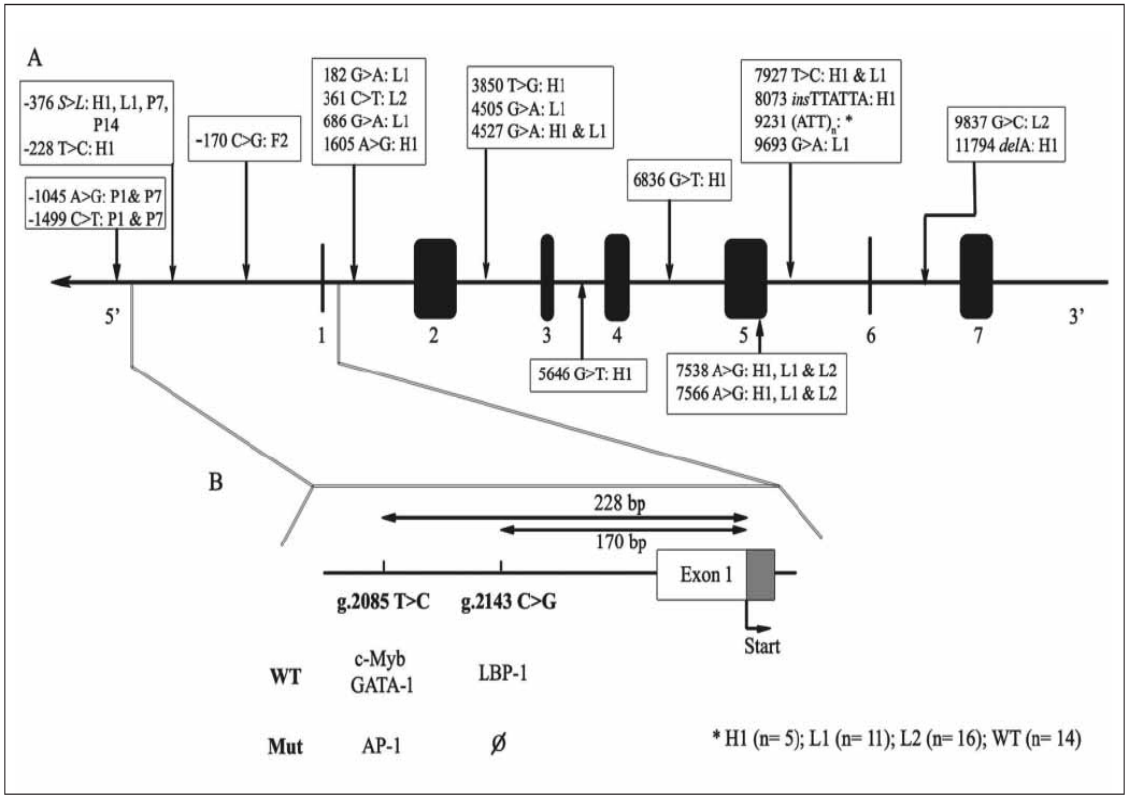


Figure 1: Genetic variations identified in the SERPINC1 gene. A) Localisation in the SERPINC1 locus of genetic variations identified in this study. The subjects carrying these variations are also indicated. S: small nucleotide series; L: large nucleotide series. B) Potential transcriptional binding sites affected by the g.2085 T>C and g.2143 C>G modifications located –228 and –170 bp from initiation codon respectively. Numbering is based on genomic DNA sequence, nucleotide +1 corresponds to the A of the ATG initiation codon of the Homo sapiens antithrombin gene (GenBank accession number NG_012462.1).

sion 15.0). Data are presented as mean ± standard deviation. Differences in plasma anti-FXa activity between genotypes were analysed by means of Student's t-test using a dominant model.

Results

Genetic study in healthy population

The anti-FXa activity observed in the 307 Spanish Caucasian healthy blood donors had a normal distribution with a mean value of 97 ± 8% (see **Suppl**. **Fig. 1**, available online at www.thrombosis-online.com). None of the subjects had antithrombin deficiency (<70% of the reference value).

We selected five subjects with extreme antithrombin levels within the normal range: three with the lowest levels (L1: 73.5%, L2: 74.5% and L3: 75.1%) and two with highest levels (H1: 114.4% and H2: 116.1%) (see **Suppl. Fig. 1**, available online at www.thrombosis-online.com). The anti-FXa activity of these subjects did not significantly change in at least two independent determinations with different samples. We sequenced 15,375 bp of the SERPINC1 gene, including the potential promoter, exons, introns and the 3' region (**Fig. 1A**). Comparison with the reference sequence revealed 18 polymorphisms and a new g.2085 T>C genetic change (**Fig. 1 and Suppl. Table 3**, available online at www.thrombosis-online.com). The last modification, identified in a subject with high antithrombin levels (H1), located in the potential promoter region, 228 bp upstream from the translation initiation codon (**Fig. 1A**). Interestingly, the g.2085 T>C transition would have potential transcriptional consequences, as it eliminates GATA-1 and c-Myb binding sites and creates a new site for AP-1 (**Fig. 1B**). The genotyping of this modification in 131 healthy subjects by HRM/Razor probe method (see **Suppl. Fig. 2**, available online at www.thrombosis-online.com) confirmed that this genetic change is a polymorphism. Actually, during this study, this change was already described as a SNP in the HapMap (rs78972925). We identified the g.2085 T>C transition in heterozygous state in seven out of 131 subjects (5%). However, this polymorphism only showed a trend to be associated with high anti-FXa activity (99 ± 11% in carriers vs. 96 ± 9% in non-carriers;p= 0.375) (**Table 1**).

We also studied the potential functional effect of the other polymorphisms identified in this study by correlating genotype with anti-FXa activity in healthy subjects. This study was restricted to those polymorphisms identified only in individuals with high or

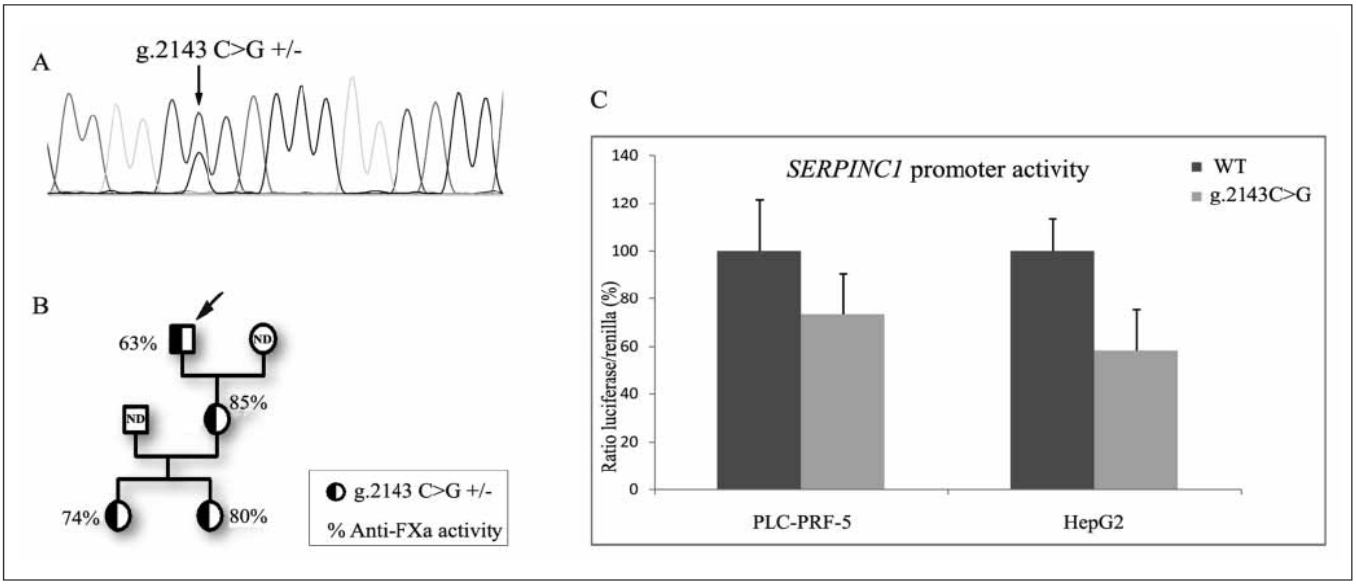


Figure 2: Identification and functional characterisation of the g.2143 C>G modification. A) Electropherogram of the mutation identified in Family 2. B) Family 2 pedigree. Proband is pointed by an arrow. Carriers of the g.2143 C>G modification are represented with the semi filled symbol. ND: not determined. Plasma anti-FXa activities (%) are also shown. C) Luciferase reporter assay of SERPINC1 promoter activity with the wild-type and mutated g.2143G allele in HepG2 and PLC-PRF-5 cell lines. Results were obtained from five independent experiments.

low antithrombin levels. As shown in Table 1, only rs2227589, located in intron 1, was significantly associated with antithrombin levels. Carriers of the polymorphic allele had a moderately reduced anti-FXa activity in plasma (16).

Haplotype analysis revealed 7 frequent haplotypes, although none of them seems to be significantly associated with anti-FXa activity (**Suppl. Figs. 3 and 4**, available online at www.thrombosis-online.com).

Finally, our sequence analysis revealed three new alleles of the rs57195053. This polymorphism was originally described as a deletion of an ATT triplet from the sequence of 14 tandem triplets on intron 5. We detected alleles with 5, 11 and 16 triplets. This result together with the description of 11 new polymorphisms affecting this location (rs72125649, rs71708020, rs71799889, rs71117551, rs71563245, rs66670532, rs72352633, rs10598596, rs72304917, rs67446132 and rs57195053), supports a single polymorphism with high allelic variability (at least from 5 to 16 ATT repeats).

Genetic analysis in thrombophilic patients with moderate antithrombin deficiency

The clinical features and anti-FXa activity of these patients are shown in **Suppl. Table 4** (available online at www.thrombosis-online.com). In these patients we only sequenced the coding region, flanking introns and the potential promoter region. We only detected three polymorphisms (**Suppl. Table 4**, available online at www.thrombosis-online.com). In addition to the rs3138521, two patients (P1 and P7) carried two linked polymorphisms (rs2227588 and rs61827938) located in the promoter region. However, none of these polymorphisms modified potential binding sites for transcription factors. Moreover, the functional effect of these polymorphisms was discarded after genotyping in 126 blood donors (**Table 1**).

Families with antithrombin deficiency carrying neither gross gene deletion nor mutation affecting exons or flanking regions

Family 1 had severe family history of thrombosis and a type I antithrombin deficiency. Three sisters with deep venous thrombosis and recurrent events showed 50–56% of functional antithrombin activity and similar antigen levels. Multiplex Ligation-dependent Probe Amplification (MLPA) analysis and sequencing of both strands of exons and flanking regions were done in two affected members, but no significant genetic change was discovered. Moreover, we did not identify mutations in the rest of the SERPINC1 gene, including 1,500 bp from the promoter region and 1,000 bp from 3' region.

In contrast, three members of Family 2 had a moderate type I antithrombin deficiency, although no one has developed thrombotic events (**Fig. 2A**). The single relevant genetic change identified in the proband of Family 2 was a point mutation g.2143 C>G, located 170 bp upstream from the translation initiation codon. In silico prediction suggested that this change could eliminate a potential binding site for the LBP-1 regulatory factor (**Fig. 1**).

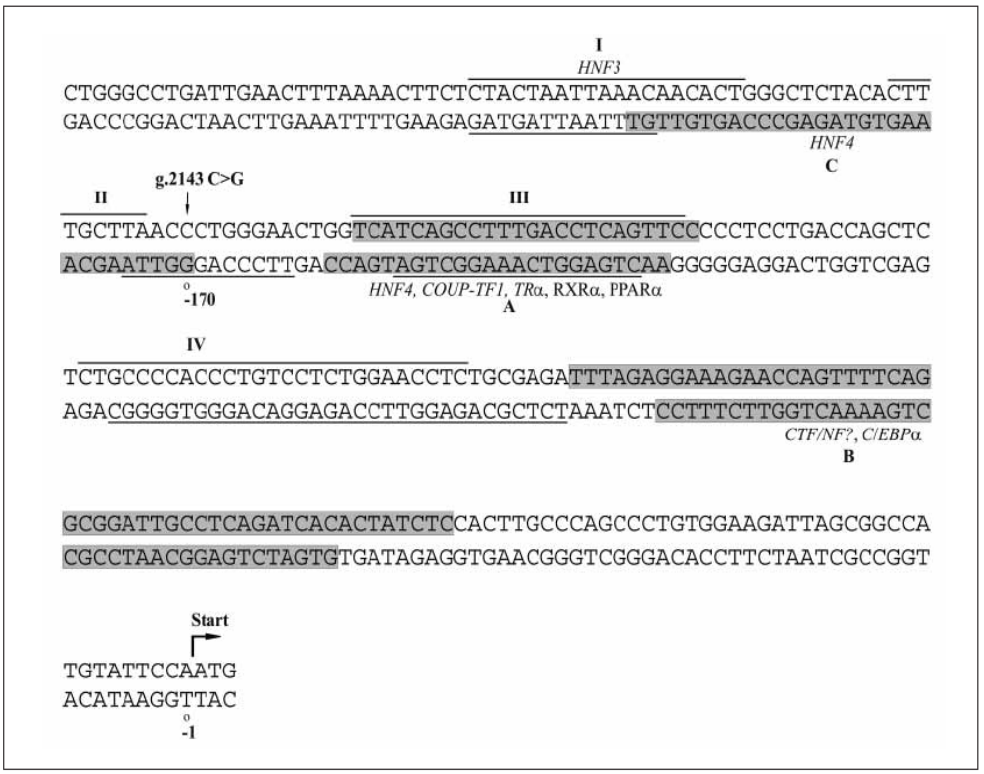


Figure 3: Localisation of g.2143 C>G in potential promoter regulatory regions of the SERPINC1 gene identified by DNase protection assay. Sequences highlighted in grey (A-C) were described by Fernandez-Rachubinski et al. (10) while regions underlined (I-IV) were described by Tremp et al. (10, 26). The interacting transcription factors identified are indicated below each sequence.

Then, we genotyped this transversion in 103 healthy blood donors by PCR-allelic specific restriction assay with Bsa JI, but no one carried it, supporting that this change is not a common polymorphism. In contrast, the mutation was present in the three membersof Family 2 with low antithrombin levels (Fig. 2A).

In order to further sustain the functional relevance of the g.2143 C>G transversion, luciferase reporter assays were performed. The plasmid containing the mutated promoter was associated with 30–50% reduction of luciferase activity in two different cell lines (HepG2 and PLC-PRF-5) (Fig. 2B).

Discussion

In the effort to find thrombotic risk factors, the identification of new elements able to affect the levels of key haemostatic molecules, such as antithrombin, is motivating. In addition to rare mutations causing severe antithrombin deficiency that significantly increase the risk of thrombosis (17), moderate reduction in antithrombin levels has recently been suggested to mildly increase the thrombotic risk (personal communication with Dr. I. Martinelli). Therefore, it would be of interest to identify new genetic variations associated with low antithrombin levels. This search has rendered excellent results in patients with congenital antithrombin deficiency, as the causative mutation is identified in a high percentage of cases (up to 90% according to recent large studies) (18). However, all genetic defects identified in these studies are gross gene deletionsor mutations affecting exons or flanking regions. The absence of genetic modifications associated with reduced levels of this key anticoagulant through regulatory mechanisms is surprising, probably reflecting that mutations affecting these regions are very rare or have small functional consequences. Alternatively, it can also be explained by the few studies that have tried to find them, probably because there are few regulatory regions identified in the SERPINC1 gene. So far, only one report from our group has reported a very mild effect associated to a common polymorphism in intron 1 (16), which also explains its negligible thrombotic consequences (19, 20). This study aimed to identify new genetic modifications associated with antithrombin levels through a regulatory mechanism, by sequencing the SERPINC1 gene in three groups of samples with different antithrombin levels. Moreover, the identification of
a functional change will also help to discover or authenticate regions with regulatory relevance in this gene.

The first approach was to dissect the genetic architecture of the SERPINC1 locus in healthy subjects with extreme antithrombin levels within the normal range (75% and 115%). The anti-FXa activity, the method widely used in thrombophilic test to detect antithrombin deficiency, shows a great variability with normal distribution (70–130%) (21). Factors such as sex, body mass index, oral contraceptives or race have been associated with antithrombin variability (22), but genetic factors must also play a relevant role according to the high heritability of this trait (h= 0.486) (23). We found a polymorphism that deserved our attention. The g.2085

What is known about this topic?

- Antithrombin is a key anticoagulant with wide variability in normal population. The high heritability of this trait sustains a role for genetic factors.
- Mutations in exons and flanking regions of SERPINC1, the gene encoding antithrombin, caused antithrombin deficiency, but only one polymorphism, rs2227589 located in intron 1, has mild functionalconsequences.
- DNase footprinting assays showed four potential regulatory regions in the 200 bp upstream the start codon.

What does this paper add?

- We identified the first mutation in the SERPINC1 promoter associated with antithrombin deficiency. This result recommends the analysis of this region in the molecular characterisation of subjects with antithrombin deficiency, particularly those with type I deficiency without mutation in the coding region of SERPINC1.
- Our study provides new evidences of the reduced genetic variability in SERPINC1 and the minor functional consequences of common polymorphisms.
- Our data support the existence of modulating genes that might beinvolved in antithrombin variability.

T>C change identified in a subject with high antithrombin levels could potentially modify the binding of different transcription factors (**Fig. 1B**). However, our study revealed that this polymorphism had no significant functional effect. The low allelic frequency of this polymorphism (0.027) encourages confirming these results in further studies including more subjects. Furthermore, despite its small size, our study provides evidence of a reduced genetic variability in SERPINC1 and the minor functional consequences of common polymorphisms, particularly those affecting coding regions. This can be explained by the extraordinary functional and conformational sensitivity of this molecule to even minor modifications that would result in a classical antithrombin deficiency (24). These data, together with the high heritability of antithrombin levels (23) support the existence of modulating genes that might be involved in antithrombin variability, and could play a role in the course or severity of thrombotic disorders (25). These genes might encode for transcriptional factors, molecules involved in folding or secretion, enzymes implicated in post-translational modifications, proteases, or miRNAs. An effort to identify these potential modulating genes of antithrombin has to be done.

However, we cannot exclude the presence of regulatory elements located up or down stream the region studied. Actually, the identification of a family (Family 1), with type I antithrombin deficiency and severe thrombosis who had neither gross deletion nor mutation in the 15,925 bp of the SERPINC1 gene strongly suggests the existence of remote regulatory regions.

Finally, we identified the first mutation in the promoter region associated with antithrombin deficiency. Certainly, the g.2143 C>G change identified in Family 2 does not abolish the transcription of the mutated allele, as demonstrated by the antithrombin levels identified in carriers, but significantly impaired it. This functional effect was further confirmed by reporter assays in different cell lines. Moreover, our result verify the regulatory relevance of the region of at least 200 bp upstream the ATG start codon. The g.2143 C>G modification is located in one of the four protected regions identified in two independent reports (10, 26). Thus, it is located at the end of region C proposed by Fernandez-Rachubinski et al.(10), and in the middle of region II according to the nomenclature of Tremp et al. (26) (F**ig. 3**). Interestingly, it has been suggested that the hepatic nuclear factor 4 (HNF4) interacts with this region (10) (**Fig. 3**). Therefore, our results support that this region, and probably the whole promoter region, should be included in the molecular analysis of subjects with antithrombin deficiency, particularly in those subjects with type I deficiency and without mutation in the coding region of SERPINC1.

Acknowledgements

We acknowledge L. Velázquez and N. Bohdan for their excellent technical assistance.

Conflict to interest

None declared.

References

1. Abildgaard U. Antithrombin--early prophecies and present challenges. Thromb Haemost 2007; 98: 97–104.
2. Rau JC, Beaulieu LM, Huntington JA, et al. Serpins in thrombosis, hemostasis and fibrinolysis. J Thromb Haemost 2007; 5 (Suppl 1): 102–115.
3. Bayston TA, Lane DA. Antithrombin: molecular basis of deficiency. Thromb Haemost 1997; 78: 339–343.
4. Garcia de Frutos P. Mechanisms of thrombophilia. Thromb Haemost 2007; 98: 485–487.
5. Lane DA, Bayston T, Olds RJ, et al. Antithrombin mutation database: 2nd (1997) update. For the Plasma Coagulation Inhibitors Subcommittee of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. Thromb Haemost 1997; 77: 197–211.
6. Olds RJ, Lane DA, Chowdhury V, et al. Complete nucleotide sequence of the antithrombin gene: evidence for homologous recombination causing thrombophilia. Biochemistry 1993; 32: 4216–4224.
7. Hernandez-Espinosa D, Ordonez A, Vicente V, et al. Factors with conformational effects on haemostatic serpins: implications in thrombosis. Thromb Haemost 2007; 98: 557–563.
8. Bock SC, Levitan DJ. Characterization of an unusual DNA length polymorphism 5' to the human antithrombin III gene. Nucleic Acids Res 1983; 11: 8569–8582.
9. Prochownik EV. Relationship between an enhancer element in the human antithrombin III gene and an immunoglobulin light-chain gene enhancer. Nature 1985; 316: 845–848.
10. Fernandez-Rachubinski FA, Weiner JH, Blajchman MA. Regions flanking exon 1 regulate constitutive expression of the human antithrombin gene. J Biol Chem 1996; 271: 29502–29512.
11. Ochoa A, Brunel F, Mendelzon D, et al. Different liver nuclear proteins binds to similar DNA sequences in the 5' flanking regions of three hepatic genes. Nucleic Acids Res 1989; 17: 119–133.
12. Sheffield WP, Wu YI, Blajchman MA. Antithrombin: Structure and Function. In: Molecular Basis of Thrombosis and Hemostasis.New York: Marcel Dekker, Inc.; 1995. pp. 355–377.
13. Corral J, Huntington JA, Gonzalez-Conejero R, et al. Mutations in the shutter region of antithrombin result in formation of disulfide-linked dimers and severe venous thrombosis. J Thromb Haemost 2004; 2: 931–939.
14. Corral J, Aznar J, Gonzalez-Conejero R, et al. Homozygous deficiency of heparin cofactor II: relevance of P17 glutamate residue in serpins, relationship with con-

formational diseases, and role in thrombosis. Circulation 2004; 110: 1303–1307.
15. Sole X, Guino E, Valls J, et al. SNPStats: a web tool for the analysis of association studies. Bioinformatics 2006; 22: 1928–1929.
16. Anton AI, Teruel R, Corral J, et al. Functional consequences of the prothrombotic SERPINC1 rs2227589 polymorphism on antithrombin levels. Haematologica 2009; 94: 589–592.
17. Maclean PS, Tait RC. Hereditary and acquired antithrombin deficiency: epidemiology, pathogenesis and treatment options. Drugs 2007; 67: 1429–1440.
18. Luxembourg B, Delev D, Geisen C, et al. Molecular basis of antithrombin deficiency. Thromb Haemost 2011; 105: 635–646.
19. Bezemer ID, Bare LA, Doggen CJ, et al. Gene variants associated with deep vein thrombosis. J Am Med Assoc 2008; 299: 1306–1314.
20. Austin H, De SC, Lally C, et al. New gene variants associated with venous thrombosis: a replication study in White and Black Americans. J Thromb Haemost 2011; 9: 489–495.
21. Tait RC, Walker ID, Islam SI, et al. Influence of demographic factors on antithrombin III activity in a healthy population. Br J Haematol 1993; 84: 476–480.
22. Conlan MG, Folsom AR, Finch A, et al. Antithrombin III: associations with age, race, sex and cardiovascular disease risk factors. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. Thromb Haemost 1994; 72: 551–556.
23. Souto JC, Almasy L, Borrell M, et al. Genetic determinants of hemostasis phenotypes in Spanish families. Circulation 2000; 101: 1546–1551.
24. Corral J, Vicente V, Carrell RW. Thrombosis as a conformational disease. Haematologica 2005; 90: 238–246.
25. Westrick RJ, Ginsburg D. Modifier genes for disorders of thrombosis and hemostasis. J Thromb Haemost 2009; 7 (Suppl 1): 132–135.
26. Tremp GL, Duchange N, Branellec D, et al. A 700-bp fragment of the human antithrombin III promoter is sufficient to confer high, tissue-specific expression on human apolipoprotein A-II in transgenic mice. Gene 1995; 156: 199–205.

Congenital disorder of glycosylation (PMM2-CDG) in a patient with antithrombin deficiency and severe thrombophilia

M. E. DE LA MORENA-BARRIO, *T. S. SEVIVAS, < I. MARTINEZ-MARTINEZ, *A. MIÑANO, *V. VICENTE, *J. JAEKEN and J. CORRAL*

*Centro Regional de Hemodonacio´n, University of Murcia, Regional Campus of International Excellence Campus Mare Nostrum, Murcia, Spain; <Department of Hematology, Centro Hospitalar de Coimbra, Portugal; and Center for Metabolic Diseases, Universitair Ziekenhuis Gasthuisberg, Leuven, Belgium

To cite this article: de la Morena-Barrio ME, Sevivas TS, Martínez-Martínez I, Miñano A, Vicente V, Jaeken J, Corral J. Congenital disorder of glycosylation (PMM2-CDG) in a patient with antithrombin deficiency and severe thrombophilia. J Thromb Haemost 2012; 10: 2625–7.

Antithrombin deficiency was the first congenital thrombophilic factor identified [1]. Since then, protein C and protein S deficiencies, as well as specific mutations in the factor V and prothrombin genes, have joined antithrombin deficiency as genetic defects predisposing patients to venous thrombosis. However, it is unlikely that new common polymorphisms could play a key role in venous thrombosis [2]. Therefore, rare mutations affecting different elements of the hemostatic system are emerging candidates [3,4]. The analysis of antithrombin in a thrombophilic patient with antithrombin deficiency has allowed us to identify a particular thrombophilic defect that, affecting the N-glycosylation pathway, indirectly disturbs this key anticoagulant serpin.

We report on an 18-year-old male with recurrent left-leg venous thrombosis. The first event happened at 11 years of age. Oral anticoagulation with warfarin was established for 5 months. A second event occurred at the same location at 12 years. Oral anticoagulation with warfarin was re-established. Surprisingly, a new thrombotic event took place 6 months later, under stable oral anticoagulant therapy with an International Normalized Ratio (INR) of 2–3. A fourth thrombotic event was diagnosed at 17 years, again under oral anticoagulant therapy. Currently, the patient is being treated with higher warfarin doses (INR 2.5–3.5).

No familial history of thrombosis was reported. Thrombophilic analysis revealed no prothrombotic polymorphisms (FV Leiden or prothrombin G20210A). No antiphospholipid antibodies were detected. Protein C and protein S levels were within the normal range. The only thrombophilic defect identified was an antithrombin deficiency (anti-FXa activity of 67%). Antigen levels of antithrombin were also low (60%), suggesting a spontaneous heterozygous type I antithrombin deficiency. However, the sequencing of the promoter, as well as the coding and flanking regions of the SERPINC1 gene, [5] revealed no mutation. Large gene deletions were excluded by multiplex ligation-dependent probe amplification (MLPA) studies. Interestingly, electrophoretic analysis of plasma antithrombin performed under either denaturing or native conditions [6] revealed the presence of an abnormal antithrombin (Fig. 1A). The increased mobility in SDS suggested abnormal post-translational modifications. We evaluated impaired N-glycosylation, the only post-translational modification described in antithrombin that also affects the function and clearance of this anticoagulant serpin [7]. N-glycan removal with N-glycosidase F [6] equalized the mobility of both the major and the minor antithrombin species, suggesting an N-glycosylation defect for 20% of the antithrombin molecules in the plasma of the proband (Fig. 1B). This abnormal antithrombin was not the b-glycoform as, among other findings, this defect was not specific for antithrombin. Thus, smaller forms of a_1-antitrypsin were also detected in the plasma of this patient (Fig. 1C). These biochemical and clinical data suggested a congenital disorder of glycosylation (CDG), a category that comprises a broad group of autoso-mal recessive disorders [8]. Actually, the antithrombin and a_1-antitrypsin electrophoretic profiles were very similar to those shown by a patient with the commonest CDG, PMM2-CDG, with mutations in PMM2, the gene encoding phosphoman-nomutase 2. This is a cytosolic enzyme that catalyzes the conversion of mannose 6-phosphate to mannose 1-phosphate, a substrate for GDP-mannose synthesis, required in N-glycosylation [8] (Fig. 1C). HPLC analysis of transferrin glycoforms pointed to a CDG-I diagnosis by the presence of asialotransferrin and a significant increase in the amount of disialotransferrin (Fig. 1D). We identified two mutations in PMM2: one of maternal origin in exon 5, and the other in exon 8, responsible for the p.R141H and p.C241S substitu-tions, respectively. Unfortunately, we do not have samples from the father. The same combination has been reported previously in a PMM2-CDG patient with a mild clinical phenotype [9]. In addition, the proband had mental retarda-tion without ataxia, strabismus, a thyroid nodule, and kyphosis. The digestive, hepatic, renal and cardiovascular systems were normal.

Correspondence: Dr Javier Corral, University of Murcia, Centro Regional de Hemodonacio´n, Ronda de Garay S/N, Murcia 30003, Spain.
Tel.: +34 968341990; fax: +34 968261914.
E-mail: javier.corral@carm.es

DOI: 10.1111/jth.12031.
Received 4 September 2012, accepted 15 October 2012

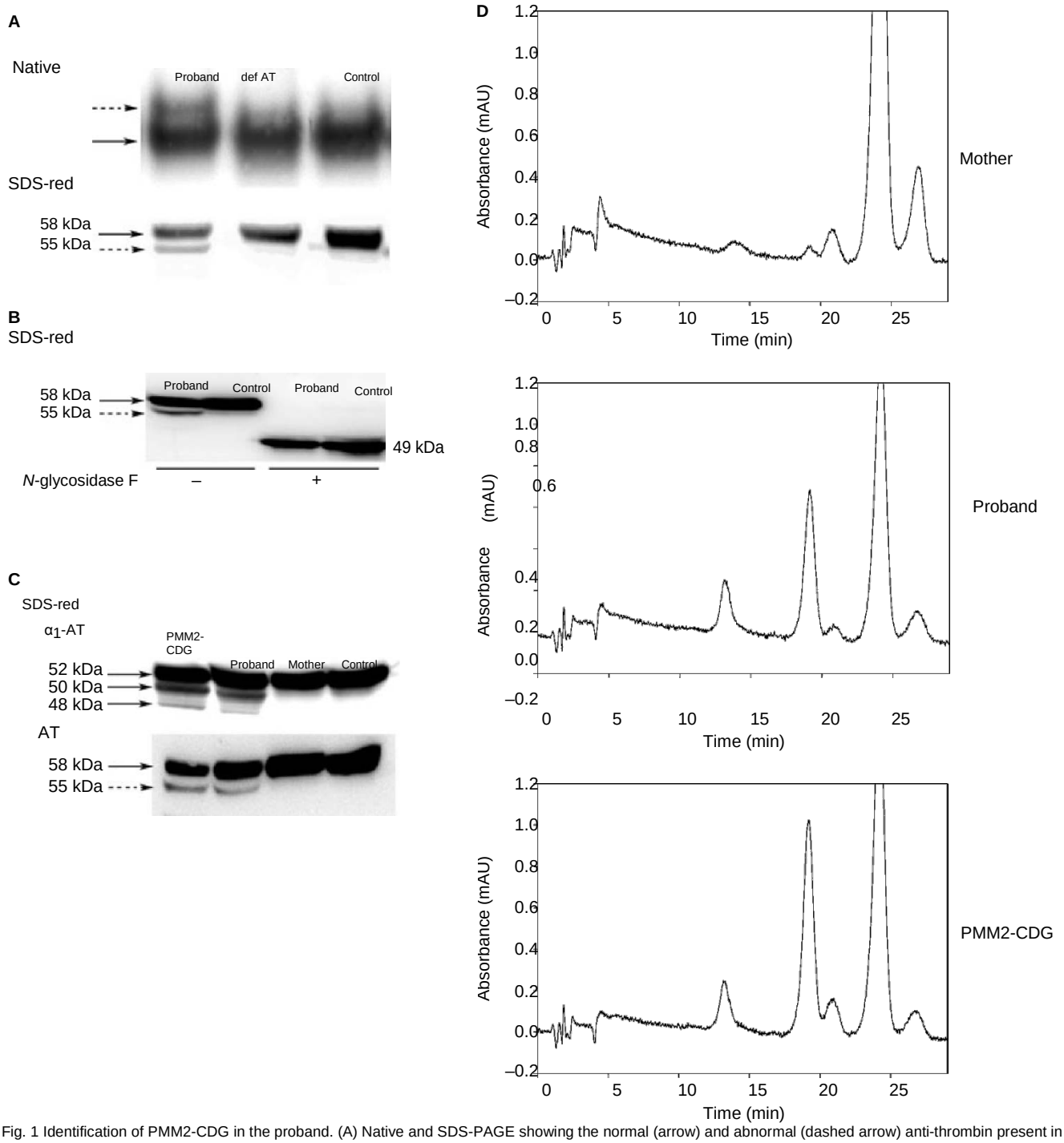


Fig. 1 Identification of PMM2-CDG in the proband. (A) Native and SDS-PAGE showing the normal (arrow) and abnormal (dashed arrow) anti-thrombin present in plasma of the proband, a patient with congenital type I antithrombin deficiency (AT def) (p.R149X) and a control. (B) SDS-PAGE and Western blot of ntithrombin from plasma untreated (–) or treated (+) with N-glycosidase F. (C) SDS-PAGE pattern of a_1-antitrypsin (a_1-AT) and antithrombin (AT) in the proband, his mother, a control, and a PMM2-CDG patient (p.F119L and p.R141H). Arrows mark the abnormal glycoforms of these molecules. (D) HPLC profile of transferrin glycoforms of the proband, his mother, and a PMM2-CDG patient (p.F119L and p.R141H).

Up to 55% of PMM2-CDG patients show acute events such as ₃stroke-like₃ episodes and ischemic cerebral vascular accidents, and at least 10 cases with either venous or arterial thrombosis have been described [10]. Acute vascular events occurred in patients younger than 15 years, especially during fever and prolonged immobilization [10]. These accidents can be recurrent in the same patient [10]. Only a few studies have evaluated hemostatic factors [10,11] and platelets [12] in PMM2-CDG with the aim of identifying the mechanism(s) involved in the development of these vascular events. PMM2-CDG patients consistently have antithrombin deficiency [10,11], with abnormal hypoglycosylated antithrombin mole-cules [10]. This might contribute to the increased risk of thrombotic events described in these patients. However, there

are no differences in antithrombin levels between patients with and without vascular events [10].

This is the first report on a PMM2-CDG patient with documented recurrent venous thrombosis. Moreover, the patient suffered from thrombotic events without a clear triggering factor, such as fever, catheterization, or immobilization. Finally, two recurrent thrombotic events occurred under oral anticoagulant therapy. This unusual clinical severity suggests that hemostatic or vascular defects other than those described in PMM2-CDG patients (or additional unknown thrombophilic anomalies) might be present in this patient. This deserves further study.

This case report also highlights another important conclusion concerning the diagnostic procedure in patients with antithrombin deficiency. Direct sequencing and MLPA anal-ysis of SERPINC1 in patients with thrombosis and antithrom-bin deficiency is highly effective. A recent study including 234 cases with antithrombin deficiency found a high mutation detection rate for SERPINC1 (83.5%) [13]. Our study suggests that a diagnosis of PMM2-CDG might be suspected among cases with antithrombin deficiency and no mutations in SERPINC1, particularly in those cases without a family history of thrombosis, as well as in cases with a lower molecular weight plasma antithrombin.

Disclosure of conflict of interests

This study was supported by 04515/GERM/06, PI12/00657 (ISCIII and FEDER), and RECAVA RD06/0014/0039 (ISCIII and FEDER). M. E. de la Morena-Barrio is a holder of a predoctoral research grant from ISCIII (FI09/00190). I. Martínez-Martínez is a researcher from Fundación para la Formación e Investigación Sanitarias de la Región de Murcia.

Acknowledgements

We are indebted to the patient and his family for their collaboration in this study.

References

1 Egeberg O. Inherited antithrombin deficiency causing thrombophilia. Thromb Diath Haemorrh 1965; 13: 516–30.

2 Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, de Andrade M. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. J Thromb Haemost 2012; 10: 1521–31.

3 Miyawaki Y, Suzuki A, Fujita J, Maki A, Okuyama E, Murata M, Takagi A, Murate T, Kunishima S, Sakai M, Okamoto K, Matsushita T, Naoe T, Saito H, Kojima T. Thrombosis from a prothrombin mutation conveying antithrombin resistance. N Engl J Med 2012; 366: 2390–6.

4 Simioni P, Tormene D, Tognin G, Gavasso S, Bulato C, Iacobelli NP, Finn JD, Spiezia L, Radu C, Arruda VR. X-linked thrombophilia with a mutant factor IX (factor IX Padua). N Engl J Med 2009; 361: 1671–5.

5 de la Morena-Barrio ME, Anton AI, Martinez-Martinez I, Padilla J, Minano A, Navarro-Fernandez J, Aguila S, Lopez MF, Fontcuberta J, Vicente V, Corral J. Regulatory regions of SERPINC1 gene: identification of the first mutation associated with antithrombin deficiency. Thromb Haemost 2012; 107: 430–7.

6 Martinez-Martinez I, Johnson DJ, Yamasaki M, Navarro-Fernandez J, Ordonez A, Vicente V, Huntington JA, Corral J. Type II antithrombin deficiency caused by a large in-frame insertion. Structural, functional and pathological relevance. J Thromb Haemost 2012; 10: 1859–66.

7 McCoy AJ, Pei XY, Skinner R, Abrahams JP, Carrell RW. Structure of beta-antithrombin and the effect of glycosylation on antithrombin₃s heparin affinity and activity. J Mol Biol 2003; 326: 823–33.

8 Jaeken J. Congenital disorders of glycosylation (CDG): it₃s (nearly) all in it. J Inherit Metab Dis 2011; 34: 853–8.

9 Gru¨newald S, Schollen E, Van Schaftingen E, Jaeken J, Matthijs G. High residual activity of PMM2 in patients₃ fibroblasts: possible pitfall in the diagnosis of CDG-Ia (phosphomannomutase deficiency). Am J Hum Genet 2001; 68: 347–54.

11 Arnoux JB, Boddaert N, Valayannopoulos V, Romano S, Bahi-Buisson N, Desguerre I, de Keyzer Y, Munnich A, Brunelle F, Seta N, Dautzenberg MD, de Lonlay P. Risk assessment of acute vascular events in congenital disorder of glycosylation type Ia. Mol Genet Metab 2008; 93: 444–9.

12 Young G, Driscoll MC. Coagulation abnormalities in the carbohydrate-deficient glycoprotein syndrome: case report and review of the literature. Am J Hematol 1999; 60: 66–9.

13 Van Geet C, Jaeken J, Freson K, Lenaerts T, Arnout J, Vermylen J, Hoylaerts MF. Congenital disorders of glycosylation type Ia and IIa are associated with different primary haemostatic complications. J Inherit Metab Dis 2001; 24: 477–92.

14 Caspers M, Pavlova A, Driesen J, Harbrecht U, Klamroth R, Kadar J, Fischer R, Kemkes-Matthes B, Oldenburg J. Deficiencies of antithrombin, protein C and protein S – practical experience in genetic analysis of a large patient cohort. Thromb Haemost 2012; 108: 247–57.

Appendix II

Diffusion of other scientific results

This appendix includes other results that have been published in peer-reviewed journals.

1. López-Contreras AJ, Sánchez-Laorden BL, Ramos-Molina B, **de la Morena ME,** Cremades A, Peñafiel R. Subcellular localization of antizyme inhibitor 2 (AZIN2) in mammailian cells: Influence of intrinsic sequences and interaction with antizymes. J. **Cell Biochem**. 107; 732-740; 2009. **Impact Factor: 3.122.**

2. Martínez-Martínez I, Ordóñez A, Pedersen S, **de la Morena-Barrio ME**, Navarro-Fernández J, Kristensen SR, Miñano A, Padilla J, Vicente V, Corral J. Heparin affinity of factor VIIa: Implications on the physiological inhibition by antithrombin and clearance of recombinant factor VIIa. **Thrombosis Research**. 127 (2); 154-160; 2011. **Impact Factor: 2.372.**

3. Guerrero JA, Teruel R, Martínez C, Arcas I, Martínez-Martínez I, **de la Morena-Barrio ME**, Vicente V, Corral J. Protective role of antithrombin in mouse models of liver injury**. J Hepatol**. 2012 Nov; 57(5):980-6. **Impact factor: 9.264**.

4. Martínez-Martínez I, Navarro-Fernández J, Østergaard A, Gutiérrez-Gallego R, Padilla J, Bohdan N, Miñano A, Pascual C, Martínez C, **de la Morena-Barrio ME**, Aguila S, Pedersen S, Kristensen SR, Vicente V, Corral J. Amelioration of the severity of heparin-binding antithrombin mutations by posttranslational mosaicism. **Blood.** 2012 Jul 26;120(4):900-4. **Impact factor: 9.898.**

5. Martínez-Martínez I, Navarro-Fernández J, Aguila S, Miñano A, Bohdan N, **de La Morena-Barrio ME**, Ordóñez A, Martínez C, Vicente V, Corral J. The infective polymerization of conformationally unstable antithrombin mutants may play a role in the clinical severity of antithrombin deficiency. **Mol Med.** 2012 Jul 18;18(1):762-70. **Impact factor: 3.76.**

6. Águila S, Martínez-Martínez I, Collado M, Llamas P, Antón AI, Martínez-Redondo C, Padilla J, Miñano A, **de la Morena-Barrio ME**, García-Avello A, Vicente V, Corral J. Compound heterozygosity involving Antithrombin Cambridge II (p.Arg416Ser) in antithrombin deficiency. **Thromb. Haemost.** 2013 Jan 17;109(3). [Epub ahead of print]. **Impact Factor: 5.044.**