ARTICLE

The 11 kb *FGA* deletion responsible for congenital afibrinogenaemia is mediated by a short direct repeat in the fibrinogen gene cluster

Marguerite Neerman-Arbez, Stylianos E Antonarakis, Ariane Honsberger and Michael A Morris

Division of Medical Genetics, University Medical School and University Hospital, Geneva, Switzerland

Congenital afibrinogenaemia is an autosomal recessive disorder characterised by the complete absence of detectable fibrinogen. We previously identified the first known causative mutations for this disorder in a non-consanguineous Swiss family. The four affected male individuals (two brothers and their first two cousins) were shown to have homozygous deletions of approximately 11 kb of the fibrinogen alpha chain (*FGA*) gene. Haplotype data suggested that the deletions occurred on three distinct ancestral chromosomes, implying that the *FGA* region of the fibrinogen locus is susceptible to deletion by a common mechanism, but the sequences responsible for the recombination remained to be identified. Here, we report the detailed characterisation of the deletion by nucleotide sequence analysis of all three deletion junctions and comparison with normal sequences. We found that all three deletions were identical to the base-pair and probably resulted from non-homologous (illegitimate) recombination. The centromeric and telomeric deletion junctions featured both a 7 bp direct repeat, AACTTTT, situated in *FGA* intron 1 and in the *FGA-FGB* intergenic sequence and a number of inverted repeats which could be involved in the generation of secondary structures. Analysis with closely linked flanking polymorphic markers revealed the existence of at least two haplotypes, further suggesting independent origins of the deletions in this family.

Keywords: recombination, genetic; mutation, gene deletion; haplotypes; haemostasis; blood coagulation disorder

Introduction

Congenital afibrinogenaemia (Mendelian Inheritance in Man, 202400) is a rare, autosomal recessive disorder

Correspondence: Marguerite Neerman-Arbez, Centre Médical Universitaire, 1 rue Michel-Servet, CH-1211 Geneva, Switzerland. Tel: +44 22 702 56 92; Fax: +41 22 702 57 06; E-mail: Marguerite.Arbez@medecine.unige.ch or Michael.Morris@medecine.unige.ch

Received 29 April 1999; revised 15 July 1999; accepted 3 August 1999

characterised by the complete absence of detectable fibrinogen and was originally described in 1920.¹ To date some 150 families with this disorder have been reported² with approximately 50% of cases present in consanguineous pedigrees.³ Despite the fact that functional assays of clot formation are infinitely prolonged, the coagulation defect is surprisingly no more severe than in the haemophilias A and B, varying from severe to mild. Uncontrolled bleeding after birth from the umbilical cord is common, and spontaneous intra-

cerebral bleeding and splenic rupture can occur throughout life. Patients generally respond well to replacement therapy with purified human fibrinogen.

The fibrinogen gene family is comprised of three genes coding for fibrinogen gamma (*FGG*), alpha (*FGA*) and beta (*FGB*), clustered in a region of approximately 50 kb on chromosome 4q28–q31.⁴ We have previously studied this region by microsatellite analysis, PCR amplification, and Southern blotting in a non-consanguineous Swiss family with four affected individuals (two brothers and their first cousins), identifying the first known causative mutations for congenital afibrinogenemia.⁵ All affected individuals were found to be homozygous for a deletion of approximately 11 kb of DNA which eliminates the majority of the *FGA* gene. The nature of the mutation identified confirmed that humans can survive despite totally lacking the capacity to synthesise active fibrinogen, as can the fibrinogen knock-out mice⁶ generated by disruption of *FGA* exon 1. Interestingly, haplotype data implied that the deletions occurred as separate historical events, on three different ancestral chromosomes, suggesting a common molecular mechanism leading to recurrent mutation.⁵

We wished to identify the molecular mechanism involved in the generation of the deletions responsible for congenital afibrinogenemia. Although the centromeric breakpoints of all deletions were known to occur within *FGA* intron 1, the exact sequence of the deletion junctions was unknown since the telomeric end of the deletions, as determined by the size of the deletion measured in Southern blots, was situated in the previously uncharacterised intergenic *FGA–FGB* region.

Here we report the cloning of the *FGA–FGB* intergenic region following long-range PCR, and the sequencing of the deletion junctions for all mutated chromosomes. We found that all three deletions were identical to the base-pair and probably resulted from non-homologous (illegitimate) recombination, mediated by direct and possibly indirect repeats in the breakpoint region.

Methods

Patients

The four adult male patients with congenital afibrinogenemia and their family have been described in detail by Neerman-Arbez *et al.*⁵

Long-range PCR Amplification and Cloning of the *FGA–FGB* Intergenic Sequence

A first long-range PCR amplification with primers *FGAN1* and *FGBN1* (Table 1) was performed from genomic DNA from affected individuals in a total volume of 100 µl with ExTaq (Takara, distributed by Axon Lab, Baden Dättwil, Switzerland) according to the manufacturer's protocol, with betaine (Qiagen Q, Basel, Switzerland) solution included in the reaction mix. The amplification programme was: 94°C, 1 min; followed by 98°C, 20 s; 68°C, 20 min (14 cycles) and 98°C, 20 s; 68°C, 20 min + 15 s/cycle (16 cycles). A final elongation step at 72°C for 10 min was performed. Nested forward primers were designed from the sequence for *FGA* (Genbank accession number M64982) and nested reverse primers from *FGB* (Genbank M64983); for primer sequences see Table 1.

For the nested amplification, 20 µl of the first reaction was used in a total volume of 100 µl as before, with primers *FGAN2* and *FGBN2*. The amplified fragment obtained from affected individuals (homozygous for the deletion) or carriers (heterozygous for the deletion) was approximately 11.5 kb (Figure 1). This product was cloned using the TOPO-XL cloning method (Invitrogen, Groningen, The Netherlands).

To obtain the centromeric portion of the *FGA–FGB* intergenic sequence, we performed a single-round PCR amplification from genomic DNA from heterozygous individuals using a forward primer in *FGA* exon 5 (*FGA* × 5L) and a reverse primer designed from the novel intergenic sequence, *FGbreak1R*. After denaturation at 94°C, the amplification programme consisted of 10 touchdown cycles of 30 s denaturation at 94°C, 30 s annealing between 60 and 50°C and 10 min elongation at 72°C followed by 20 cycles (30 s at 94°C, 30 s at 50°C, 10 min at 72°C). The enzyme used was ExTaq (TaKaRa). The amplified product from the normal, non-deleted allele, approximately 7 kb, was also cloned in the TOPO-XL vector (Invitrogen, Groningen, The Netherlands).

PCR Amplification and Sequencing of the Deletion Junctions

PCR amplification of the deletion junction from all patients and from carrier individuals was performed with primers *FGA* × 1L and *FGbreak1R*. Sequences were obtained with the primers used for the amplification with a semi-automated sequencer (ABI 377; Perkin-Elmer Applied Biosystems),

Table 1 List of oligonucleotide primers

Primer	Sequence 5'–3'
<i>FGAN1</i>	GGAGCCAGCCCCACCCCTTAGAAAA
<i>FGAN2</i>	TCCATGAGGATCGTCTGCCTAGTT
<i>FGBN1</i>	TAGCCAGCTTACCAGGATGGGACC
<i>FGBN2</i>	GCCAAACAAATACCAGATCTCACTG
<i>FGA</i> × 5L	GATCTTGTCTGAGGGTCATGC
<i>FGA</i> × 1L	CAGCCCCACCCCTTAGAAAAAG
<i>FGAi1AL</i>	GATGGATGAGAGTGTACAAC
<i>FGAi1AR</i>	CTACTGTAGAGCTCTTTCTCC
<i>FGAi1BR</i>	GCATTATAATGCACCAAGGC
<i>FGAi1DL</i>	GCATCTGTCTTCAGTCTTACC
<i>FGbreak1R</i>	ACCACCATCAATAGGCTTGG
<i>FGbreak2R</i>	TATTATAGCATATGGGTGTG
<i>FGbreak4R</i>	GGGTGTGATAATATACATTC

using standard protocols. The novel 1094bp sequence surrounding the breakpoint in the *FGA-FGB* intergenic sequence has been submitted to Genbank (accession number AF145725).

Secondary Structure Prediction

The sequences around the deletion junctions were analysed for secondary structures with the mfold program (<http://mfold2.wustl.edu/~mfold/dna/form1.cgi>).⁷

Deletion Carrier Screening

A total of 468 anonymised DNA samples isolated from buccal smears from unrelated individuals from the Geneva region were analysed for the presence of the deletion junction by PCR amplification with primers FGA × 1L and FGbreak1R and for the control normal band with a second reverse primer FGAi1AR. A nested PCR was performed with primers FGAi1AL and FGbreak2R (for the deletion) and FGAi1BR (for the normal band). The PCR products were then dot-blotted and hybridised with a ³²P end-labelled internal oligonucleotide (FGAi1DL for the normal product, FGbreak3R for the deletion), under standard conditions.

Chromosome 4 Microsatellite Analysis

DNA extractions and microsatellite analysis were performed as previously described.⁸ The new markers used for haplotype analysis were *D4S3021*, *D4S2999*, *D4S2631* and *D4S3016*. The markers were located to either side of the *FGA* gene, covering a total area of 5 centimorgans (Marshfield sex-averaged genetic map, <http://www.marshmed.org>). The locus order and genetic distances in Kosambi centimorgans were as follows: *D4S2962* – 1.6 – *D4S3021* – 0.0 – *fibrinogen (FGB)* – 0.6 – *D4S2999* – 0.2 – *D4S2934* – 2.5 – *D4S2631/D4S3016*.

Results

The fibrinogen locus is comprised of single copies of three genes coding for fibrinogen gamma (*FGG*), alpha (*FGA*) and beta (*FGB*), clustered in a region of approximately 50 kb on chromosome 4q28–q31,⁴ (Figure 1). The sequences of the three fibrinogen genes, and that of the intergenic region between *FGG* and *FGA* have been previously determined (Genbank accession numbers: alpha M64982; beta M64983; gamma M10014; *FGG-FGA* intergenic U36478). In contrast the sequence for the intergenic region between *FGA* and *FGB*, the site of the telomeric deletion break-point was unknown. The cloning and partial sequencing of this portion of chromosome 4q was therefore an essential prerequisite to the sequencing of the deletion junctions.

The strategy used to clone the *FGA-FGB* intergenic region is summarised in Figure 1. We used long-range PCR to amplify this region from the genomic DNA of individuals from the Swiss family with congenital afibrinogenaemia, using forward primers in *FGA* exon 1 and reverse primers in *FGB* exon 7 (note the opposite orientations of the *FGA* and *FGB* genes). All products amplified from carrier individuals and afibrinogenaemia patients gave a single band of approximately 11.5 kb. No larger products were seen in amplifications from heterozygote deletion carriers: the

Figure 1 Schematic representation of the human fibrinogen gene cluster on 4q28–31 and the strategy used to clone the *FGA-FGB* intergenic sequence. The large white arrows represent the gamma, alpha and beta chain genes and their 5' to 3' orientation. The small black arrows represent the oligonucleotide primers used for the PCR amplification of **a**) the sequence telomeric to the breakpoint (which is shown as a cross) from the deleted allele (see Methods), and **b**) the sequence centromeric to the breakpoint from the non-deleted allele. Oligonucleotides: 1,2: FGA1 and N2; 3,4: FGBN1 and N2; 5: FGA × 5L; 6: FGbreak1R

normal allele, which would yield a fragment of over 20 kb, was not amplified.

In order to determine the normal sequence at the telomeric end of the deleted region, we performed a single PCR amplification from genomic DNA of heterozygous individuals using a forward primer in *FGA* exon5 and a reverse primer designed from the novel intergenic sequence, FGbreak1R (Figure 1, oligonucleotides 5 and 6).

Figure 2 shows the sequence of the afibrinogenaemia deletion junction, with the corresponding normal sequences from *FGA* intron 1 (the centromeric breakpoint) and from the *FGA-FGB* intergenic sequence (the telomeric breakpoint). This deletion junction sequence was identical in all the affected individuals and in their heterozygous parents (this was confirmed with three different primer pairs, in independent PCR and sequencing experiments; data not shown).

The breakpoint is situated in a stretch of four Ts located in both *FGA* intron1 and on the same DNA strand in the *FGA-FGB* intergenic sequence. The regions bordering the deletion breakpoint contain a number of sequence elements that have been implicated in the generation of human gene deletions (Figure 2a). Firstly, the breakpoint is located in a 7-bp direct repeat, AACTTTT; one copy of the repeat persists on the deleted alleles, but with the deletion of a single T. Secondly, there is an imperfect inverted copy of this repeat (6 out of 7 nucleotides) immediately telomeric of the deletion junction. Thirdly, the breakpoints are surrounded by further inverted repeats: analysis with the computer program mfold identified one centromeric and two telomeric predicted stem-loop structures (Figure 2b).

Previous haplotype data obtained with markers flanking the fibrinogen gene cluster suggested an

Figure 2a Sequence of the afibrinogenaemia deletion junction. The normal centromeric sequence is from *FGA* intron 1 (top line) and the normal telomeric sequence from the *FGA-FGB* intergenic region (bottom line). Homologies are indicated by shading. The 7 bp direct repeat is in bold. The deleted sequence (middle line) has a deletion of one T. The inverted repeat is indicated by asterisks, and its centre of symmetry by a vertical line. **b** Secondary structure predictions around the centromeric and telomeric deletion junction regions. The grey bars indicate the position of the breakpoints. Structures were predicted with the mfold program; the calculated free energies were -3.9 and -5.8 kcal/mol respectively (<http://mfold2.wustl.edu/~mfold/dna/form1.cgi>)⁷

Figure 3 Pedigree of the Swiss congenital afibrinogenemia family and haplotypes of the region surrounding the fibrinogen gene cluster. The markers (from top to bottom) and genetic distances in centimorgans are: D4S2962 – 1.6 – D4S3021 – 0.0 – fibrinogen (FGB) – 0.6 – D4S2999 – 0.2 – D4S2934 – 2.5 – D4S2631/D4S3016. FGAi3 is the tetranucleotide repeat in intron 3 of FGA, as described in reference 5. The FGA deletion (shown as Δ) is associated with three different haplotypes

with the four most closely linked markers, two to each side of the fibrinogen gene locus, strongly suggest the former hypothesis: for example, the affected brothers in the left half of Figure 3 are heterozygous at each of these four loci. At least two recombination events would therefore be necessary to derive these haplotypes from a common origin, between *FGA* and very tightly linked markers ($\theta = 0.00$ and $\theta = 0.56$ cM according to the Marshfield map). Less probably, the separate haplotypes could also have evolved from a common origin by four distinct mutations of the dinucleotide repeats. However, if this deletion were indeed a recurrent mutation, one would expect the disorder to be more common, whereas congenital afibrinogenaemia is extremely rare; in this respect, we studied nearly 1000 alleles from the local population, and found no further deleted alleles. It will not be possible to distinguish between these two hypotheses without widespread studies in afibrinogenaemia patients from various populations.

We have examined three afibrinogenaemia alleles of apparently independent origin, according to their haplotypes, revealing a deletion mediated by direct and perhaps by indirect repeats. This implies the existence of a relative 'hotspot' for deletion in the fibrinogen gene cluster. It is interesting to note that the first studies of the fibrinogen gene cluster in afibrinogenaemia patients by Southern blotting detected no deletions,¹⁶ implying that further mutation mechanisms exist. It will consequently be of great interest to evaluate the frequency of the mutation amongst further afibrinogenaemia patients.

Acknowledgements

We thank all members of the family for their continued enthusiasm and support, without which this study would never have been possible. We are particularly grateful to Dr Philippe de Moerloose for his expert assistance and advice, as well as to Dr Anni Schönborner, Dr Colette Rossier and Isabelle Bouchardy for expert technical help. MNA is the recipient of grants from the Roche Research Foundation (98-180) and from the Swiss National Science Foundation (Marie-Heim Vögtlin subsidy). This study was supported by grant 31-055848.98 from the Swiss National Science Foundation, from grant HL38165-14 from the NIH, by the ZLB Central Laboratory, Berne, Switzerland and by the University and Cantonal Hospital of Geneva. We thank Hewlett-Packard (Switzerland) for their kind support by donation of computer equipment.

References

- 1 Rabe F, Salomon E: Ueber-faserstoffmangel im Blute bei einem Falle von Hämophilie. *Arch Int Med* 1920; **95**: 2-14.
- 2 Martinez J: Congenital dysfibrinogenemia. *Curr Opin Hemat* 1997; **4**: 357-365.
- 3 Crabtree GR: The molecular biology of fibrinogen. In: Stamatoyannopoulos *et al* (eds). *The Molecular Basis of Blood Diseases*. W.B. Saunders Co., Philadelphia, PA, USA, 1983, chap 17.
- 4 Kant JA, Fornace AJ Jr, Saxe D, Simon MI, McBride OW, Crabtree GR: Evolution and organization of the fibrinogen locus on chromosome 4: gene duplication accompanied by transposition and inversion. *Proc Natl Acad Sci USA* 1985; **82**: 2344-2348.
- 5 Neerman-Arbez M, Honsberger A, Antonarakis SE, Morris MA: Deletion of the fibrinogen alpha-chain gene (*FGA*) causes congenital afibrinogenemia. *J Clin Invest* 1999; **103**: 215-218.
- 6 Suh TT, Holmback K, Jensen NJ *et al*: Resolution of spontaneous bleeding events but failure of pregnancy in fibrinogen-deficient mice. *Genes Dev* 1995; **9**: 2020-2033.
- 7 SantaLucia J Jr: A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* 1998; **95**: 1460-1465.
- 8 Neerman-Arbez M, Antonarakis SE, Blouin JL *et al*: The locus for Combined Factor V-Factor VIII Deficiency (F5F8D) Maps to 18q21, between D18S849 and D18S1103. *Am J Hum Genet* 1997; **61**: 143-150.
- 9 Cooper DN, Krawczak M: Human Gene Mutation. BIOS Scientific: Oxford, 1993.
- 10 Ottolenghi S, Giglioni B: The deletion in a type of delta 0-beta 0-thalassaemia begins in an inverted Alu repeat. *Nature* 1982; **300**: 770-771.
- 11 Lehrman MA, Schneider WJ, Sudhof TC, Brown MS, Goldstein JL, Russell DW: Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* 1985; **227**: 140-146.
- 12 Lakich D, Kazazian HH Jr, Antonarakis SE, Gitschier J: Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet* 1993; **5**: 236-241.
- 13 Rumsby G, Fielder AH, Hague WM, Honour JW: Heterogeneity in the gene locus for steroid 21-hydroxylase deficiency. *J Med Genet* 1988; **25**: 596-599.
- 14 Phillips JA, Vik TA, Scott AF *et al*: Unequal crossing-over: a common basis of single alpha-globin genes in Asians and American blacks with hemoglobin-H disease. *Blood* 1980; **55**: 1066-1069.
- 15 Efstratiadis A, Posakony JW, Maniatis T *et al*: The structure and evolution of the human β -globin gene family. *Cell* 1980; **21**: 653-668.
- 16 Uzan G, Courtois G, Besmond C *et al*: Analysis of fibrinogen genes in patients with congenital afibrinogenemia. *Biochem Biophys Res Commun* 1984; **120**: 376-383.