Evolution of the fibrinogen γ′ chain: implications for the binding of factor XIII, thrombin and platelets

R. F. DOOLITTLE, S. HONG and D. WILCOX

Departments of Chemistry & Biochemistry and Molecular Biology, University of California, San Diego, La Jolla, CA, USA

The γ′ chain of human fibrin(ogen) has received much attention of late because of its apparent regulatory roles, both with regard to the binding of factor (F)XIII by fibrinogen [1,2] and the binding of thrombin by fibrin [3,4]. The γ′ chain is mostly a longer version of the γ chain, the result of alternative splicing [5,6], or in this case non-splicing. The result in humans is a γ chain 20 residues longer than the more usual spliced form. The extension is markedly anionic, in part because of two sulfated tyrosines [7] that have been implicated in binding to thrombin.

Interestingly, mouse fibrin does not bind thrombin with the same avidity as human fibrin, an observation attributed to its shorter γ′ extension. In a recent genetic engineering study, the mouse γ′ sequence was replaced by the human version, after which tight thrombin binding was observed [8]. The experiment raises the question, are humans (or primates) unique in their ability to regulate coagulation by tight binding of γ′ to thrombin? In an effort to answer this question, we conducted a bioinformatics survey in which we examined the genomes of 16 vertebrates in search of putative γ′ sequences. The *sine qua non* for having two versions of the γ chain is an intron near the 3′ terminus of the coding region. It has long been known that the lamprey γ-chain gene lacks the intron [9], as does the puffer fish [10].

Publicly accessible data bases were searched for lamprey, puffer fish, zebra fish, frog, chicken, platypus, opossum, dog, cat, mouse, rat, ox, horse, macaque, chimpanzee and human genes. These databases can be accessed through various portals, including http://www.ncbi.nlm.gov, http://genome.wustl.edu/, http://genome.ucsc.edu/cgi-bin/hgGateway/ and http://www.sanger.ac.uk/. Various kinds of BLAST searches [11] were conducted, TBLASTN in particular being used to search translated amino acid sequences against raw DNA data. The highly conserved region composed of γ-chain residues 380–410, which is present in both γ and γ′ chains, was mostly used as the probe, high scoring hits being examined for introns in the region and, when found, downstream DNA sequences scrutinized for splicing candidates.

We confirmed that the intron allowing alternative splicing leading to two forms of fibrinogen γ chains does not occur in any fish or frog. It exists in the green lizard gene, but the most probable candidate sequence for a spliced form was a solitary isoleucine residue; direct experiment will be needed to determine whether splicing really occurs. Of more interest, the unspliced form of the lizard gene has an RGD sequence at the same position as occurs in frog and lamprey (Fig. 1A,C).

The chicken gene also has the intron, but an almost immediate terminator codon leads to an unspliced form shorter than the spliced alternative (Fig. 1). Suspiciously, an RGD sequence occurs three residues after the terminator codon, followed by another terminator that coincides with lizard and frog (Fig. 1C).

The situation in the platypus (a monotreme) is similar. There is an intron at the expected location, but a mere 2-residue extension would occur in the absence of splicing. All other mammals studied (opossum, mouse, rat, dog, cat, bovids, horse and three primates) have extended γ′ sequences, the lengths of which range from eight to 20 residues, the longest being from primates and horse. The data strongly suggest that the horse gamma′ segment, which is the same length as is found in primates and has two tyrosines with flanking residues characteristic of sulfation, should bind thrombin just as tightly as does human fibrin.

Clearly, the regulation of clotting exercised by FXIII binding to fibrinogen and by thrombin binding to fibrin are phenomena that appeared relatively late in the evolution of the vertebrate coagulation process. There is another aspect to consider, however, the transition from an unspliced version of the γ chain to a spliced version must have had an impact on the way fibrinogen binds to platelets. The gamma-A chain, the result of splicing, is well known to employ its carboxy-terminus in the binding of platelets [12]. In humans, the unspliced form (gamma′) must lack that property. Early diverging vertebrates (fish and amphibians) lack the intron and have only a single kind of carboxy-terminus. The fact that some of these, but not all, have an RGD sequence at that position (Fig. 1B), a motif known to be integral to platelet–fibrinogen interactions, strongly suggests the carboxy terminus is used for binding thrombocytes (the

Correspondence: Russell F. Doolittle, Department of Chemistry & Biochemistry, University of California, San Diego, La Jolla, CA 92093-0314, USA.
Tel.: +1 858 534 4417; fax: +1 858 534 4985.
E-mail: rdoolittle@ucsd.edu

```
A                                                    B                                            splice
                                                                                                   ▼
Ox         ..K Q V G E H H V E I E Y D                 Ox         ..G E G Q Q H Q L G G A K Q   A G D V
Horse      ..K Q V G P E H R G K I E H E Y D T E Y D E  Horse      ..G E G Q Q F H L G A A K Q   A G D V

Dog        ..K Q V G P E H H V E I H Y D               Dog        ..G E G Q Q H L G G A K Q   A G D I
Cat        ..K Q V G P E H H V E I M T K S F Y P E N   Cat        ..G E G Q Q H L G G A K Q   V G D I

Rat        ..K Q V S V E H E V D V E Y P               Rat        ..G E G Q Q H M G G S K Q   V G D M
Mouse      ..K Q V S V D H E V E I E Y                 Mouse      ..G E G Q Q H M G G S K Q   A G D V

Macaque    ..K Q V R Q E H P V E T E Y D S F Y P E D D S  Macaque  ..G E G Q Q H L G G A K Q   A G D V
Chimp      ..K Q V R P E H P V E T E Y D S L Y P E D D L  Chimp    ..G E G Q Q H L G G A K Q   A G D V
Human      ..K Q V R P E H P A E T E Y D S L Y P E D D L  Human    ..G E G Q Q H L G G A K Q   A G D V

Opossum    ..K Q V G S D H T L K                       Opossum    ..V D G Q Q H M G A S K Q   - G D V
Platypus   ..K Q V G                                   Platypus   ..G E G Q Q H H L G G S K Q   V G D A

Chicken    ..K Q V X P L D R G D T                     Chicken    ..S I D G Q Q H S G G L K Q   V G D S
Lizard     ..K E V L G N R G D I                       Lizard     ..S V S G Q Q I - G G V K E   I

        C
           HumanS      ..M K I I P F N R L T I G E - - - G Q Q H H L G G A K Q A G D V
           ChickenS    ..M K I I P F N R L S I D G - - - - - Q Q H S G G L K Q V G D S
           ChickenU    ..M K I I P F N R L S I D G - - Q Q H S G G L K Q X P L R G D T
           LizardU     ..M K I I P F S R L S V S G - - Q Q I G G V K E V L G N R G D I
           FrogU       ..M K I I P F N R Y G A E G Q Q T L G G S K K S D F E N R G D F
           ZebrafishU  ..M K I I P I N R I G A T - - - Q Q Q E S G S K Q F G G I G D I
           CarpU       ..M K I I P I N R I S A A T - - G S Q F T D A K Q F G G L G D I
           PufferfishU ..M K L I P L S R L A A G G - - - - Q T T G V K E F G G L G - L
           LampreyU    ..M K L L P M G R D L S G H - - - - G G Q Q Q S K G N S R G D N
```

Fig. 1. (A) Carboxy-terminal segments of fibrinogen gamma′ sequences from various vertebrates. Tyrosines that may be sulfated are emboldened. Note that in the case of the chicken the sequence shown here has an early terminator (X) after which an RGD sequence occurs (underscored) before encountering another terminator. (B) Carboxy-terminal segments of gamma A sequences, showing the results of splicing leading to gamma A chain. Note that the residues to the left of the splicing site also occur in the γ′ chains. (C) Alignment of carboxy-terminal segments of γ chains from human and chicken (spliced form, chickenS; unspliced, chickenU), and unspliced forms of green lizard (lizardU), frog and assorted fish γ chains. Lysines (K) putatively involved in γ−γ cross-linking are emboldened. Note the underscored RGD motif near the ends of the (unspliced) lizard, frog and lamprey chains (underscored).

infra-mammalian equivalent of platelets). Indeed, the lizard, frog and some fish all have an RGD motif very near the carboxy-terminus of unspliced forms of the γ chain (Fig. 1C).

Recently, a set of crystal structures was published for the platelet integrin $\alpha_{IIb}\beta_3$ complexed with synthetic peptides corresponding to the carboxy-terminus of the human gamma A chain, a 10-mer (LGGAKQAGDV) and a 12-mer with two additional residues at the amino terminus [13]. These authors also took note of the fact that the lamprey γ chain has the sequence RGD very near the carboxy-terminus, an observation that had been remarked upon previously [14]. Knowing that the $\alpha_{IIb}\beta_3$ integrin often binds to regions with the RGD motif, they also solved two additional structures of $\alpha_{IIb}\beta_3$ complexed with two peptides in which the AGDV terminus was changed to RGDV. Comparisons of the four structures showed that the basis for the integrin binding the gamma A carboxy-terminus depends on the side chain of Lys-406, six residues from the carboxy-terminus, providing the positive charge ordinarily contributed by the arginine in RGD. In humans and all mammals, the same lysine is known to be essential for the formation of γ−γ dimers by FXIII. Whether or not lysine residues further from the carboxy terminus, as is the case in some fish that do not have RGD near their carboxy-termini (puffer fish, zebra fish, carp in Fig. 1C), could play the same role remains to be determined. It should be noted that the puffer fish lacks the penultimate aspartic acid, also thought to be essential for binding to $\alpha_{IIb}\beta_3$.

In summary, the fibrinogen gamma′ sequence is directly descended from a primitive unspliced version of the gamma chain gene. The primitive form has an RGD sequence that almost certainly was used for bridging thrombocytes. After the introduction of the splicing site, the thrombocyte-binding function was transferred to the spliced version, a proximal lysine in the parent form being used as a substitute for the arginine in RGD [13]. Freed from its earlier obligation, the (unspliced) gamma′ chain gene was free to acquire new functions, including the binding of FXIII and thrombin.

Disclosure of Conflict of Interests

The authors state that they have no conflict of interest.

References

1 Siebenlist KR, Meh DA, Mosesson MW. Plasma factor XIII binds specifically to fibrinogen molecules containing γ′ chains. *Biochemistry* 1996; **35**: 10448–53.

2 Moaddel M, Farrell DH, Daugherty MA, Fried MG. Interactions of human fibrinogens with factor XIII: roles of calcium and the gamma′ peptide. *Biochemistry* 2000; **39**: 6698–705.

3 Meh DA, Siebenlist KR, Brennan SO, Holyst T, Mosesson MW. The amino acid sequence in fibrin responsible for high affinity thrombin binding. *Thromb Haemost* 2001; **85**: 470–4.

4 Lovely RS, Moaddel M, Farrell DH. Fibrinogen γ′ chain binds to thrombin exosite II. *J Thromb Haemost* 2003; **1**: 124–31.

5 Crabtree GR, Kant JA. Organization of the rat gamma-fibrinogen gene; alternative mRNA splice patterns produce the gamma A and gamma B (gamma′) chains of fibrinogen. *Cell* 1982; **31**: 159–66.

6 Chung DW, Davie EW. γ and γ′ chains of human γ-fibrinogen are produced by alternative mRNA processing. *Biochemistry* 1984; **23**: 4232–6.

7 Farrell D, Mulvihill ER, Huang SM, Chung DW, Davie EW. Recombinant human fibrinogen and sulfation of the gamma′ chain. *Biochemistry* 1991; **30**: 9414–20.

8 Mosesson MW, Cooley BC, Hernandez I, Diorio JP, Weiler H. Thrombosis risk modification in transgenic mice containing the human fibrinogen thrombin-binding gamma′ chain sequence. *J Thromb Haemost* 2009; **7**: 102–10.

9 Pan Y. *Studies on Fibrinogen Evolution. Part I: Intron Characterization of Lamprey Fibrinogen Genes*. PhD Dissertation, San Diego, CA: University of California, 1992.

10 Jiang Y, Doolittle RF. The evolution of vertebrate blood coagulation as viewed from a comparison of puffer fish and sea squirt genomes. *Proc Natl Acad Sci U S A* 2003; **100**: 7527–32.

11 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389–402.

12 Hawiger J. Adhesive ends of fibrinogen and its anti-adhesive peptides: the end of a sage? *Semin Hematol* 1995; **32**: 99–109.

13 Springer TA, Zhu J, Xiao T. Structural basis for distinctive recognition of fibrinogen γC peptide by the platelet integrin $\alpha_{IIb}\beta_3$. *J Cell Biol* 2008; **182**: 791–800.

14 Doolittle RF. The structure and evolution of vertebrate fibrinogen: a comparison of the lamprey and mammalian proteins. *Adv Exp Med Biol* 1990; **281**: 25–37.