

The Protochordate *Ciona intestinalis* Has a Protein Like Full-Length Vertebrate Fibrinogen

Russell F. Doolittle

Departments of Chemistry & Biochemistry and Molecular Biology, University of California, San Diego, Calif., USA

Key Words

Fibrinogen · Fibrin · Protochordate · Sea squirt

Abstract

In a recent review, a putative fibrinogen-like protein in the protochordate *Ciona intestinalis* was noted. Unfortunately, computer-directed splicing had omitted several exons, mistakenly generating a single long polypeptide chain. In fact, 3 consecutive genes exist, the translated versions of which are homologous to individual vertebrate fibrinogen chains. The circulating form is likely a 6-chain covalent dimer, just as occurs in vertebrates.

Copyright © 2012 S. Karger AG, Basel

Introduction

In a recent minireview on the evolution of coagulation and inflammation [1], attention was called to a putative protein in the protochordate *Ciona intestinalis*, commonly called the sea squirt, that had many features characteristic of vertebrate fibrinogens. Although homologs of the carboxyl domains of β and γ chains of fibrinogen are widespread in the animal world [2–4], the cited protein was unique in that it also had sequences consistent with

‘coiled coils,’ as well as constellations of cysteines of the sort found in ‘disulfide rings’ known to bind 3 polypeptide chains symmetrically [5, 6]. The protein had initially been dismissed as just one of many fibrinogen-related domains during a search [7] of the draft genome of the *C. intestinalis* genome [8]. Subsequently, the highest scoring of those early candidates, found on scaffold 73, was back-searched against the GenBank nonredundant protein database [1], by which time it had been annotated by the NCBI (identification No. XP_002122759). The protein sequence had been generated by automated computational analysis from genomic sequence NW001955408 using the NCBI gene prediction tool GNOMON; it was said to have been supported by EST evidence and was predicted to be a single polypeptide chain composed of 1,160 amino acids.

However, an attempt to model the single-chain protein made it clear that the prediction was structurally improbable, especially in its central region. Accordingly, a manual reevaluation of the DNA sequence was undertaken, the results of which revealed 5 additional exons, including 2 encoding signal peptides. In the end, a 3-gene system was identified that encodes a protein with 3 nonidentical polypeptide chains, likely a covalent dimer, that is remarkably similar to vertebrate fibrinogen.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2012 S. Karger AG, Basel
1662–811X/12/0042–0219\$38.00/0

Accessible online at:
www.karger.com/jin

Dr. Russell F. Doolittle
Departments of Chemistry & Biochemistry and Molecular Biology
University of California, San Diego
La Jolla, CA 92093-0314 (USA)
Tel. +1 858 534 4417, E-Mail rdoolittle@ucsd.edu

stellation known as a 'disulfide ring.' In vertebrate fibrinogen, such disulfide rings set the boundaries for coiled coils. Further, one disulfide ring from each half molecule packs against the other to form short β sheets at the dimer interface [6]. Additionally, the vertebrate dimer depends on the β chain crossing over to the other half of the molecule to form a disulfide bond with the opposing α chain and vice versa. Equivalent cysteine residues occur in the putative sea squirt protein A and B chains, making it likely that the circulating protein is a 6-chained covalent dimer (fig. 1c).

Nonetheless, several anomalously positioned cysteine residues remain, including 2 in the A chain and 2 in the C chain in the region predicted to form coiled coils. It is possible that unusual splicing patterns could remove or replace some of these.

The sea squirt protein appears to be heavily glycosylated, with the A chain having 2 putative asparaginyl-sugar sites, the B chain 4, and the C chain 6, amounting to 12 all together or 24 in a 6-chained dimer. The molecular weight of the mature predicted protein would be approximately 330 kD, of which 53 kD would be carbohydrate. Residues consistent with calcium binding sites observed in vertebrate and invertebrate [12] fibrinogen-related domains appear intact.

Functional Aspects

None of the chains in the sea squirt protein has an amino terminal extension with a possible site for cleavage by thrombin. Nor do the carboxyl terminal regions have extensions with sites that might be cross-linked by factor XIII. On the other hand, all three carboxyl-terminal domains have sequences compatible with binding sites ('holes') of the sort that in vertebrate fibrinogen bind peptide knobs. In most fibrinogen-related proteins, these 'holes' bind sugars or sialic acids [12], and the same is likely true of the sea squirt protein. In this regard, 2 of the 3 sites lack key residues involved in the binding of arginine-containing peptides. There is a single RGD sequence in the C chain; in mammals RGD sequences are involved in fibrinogen binding to platelets.

Phylogenetic Relationships

Although two of the sea squirt polypeptide chains appear slightly more similar to vertebrate γ chains (fig. 1d), the distinction is slight and is offset by the disulfide arrangement of carboxyl domains of all three being characteristic of vertebrate fibrinogen β chains, which have an additional disulfide compared with γ chains.

The approximate time of the various gene duplications can be gauged from the resemblances between the three

Ciona sequences and the β and γ chains of various vertebrate fibrinogens. As it happens the three *Ciona* polypeptides have approximately the same degree of resemblance to each other (33.7%) as they do to vertebrate fibrinogen β and γ chains (32.7%) and that vertebrate β and γ chains have to each other (34.6%). It seems clear that the gene duplications leading to the evolution of the three *Ciona* polypeptides and the β and γ chains of vertebrate fibrinogen occurred within the same general time frame.

Discussion

The absence of likely thrombin cleavage sites near any of the three amino termini suggests the putative fibrinogen-like protein in *C. intestinalis* is not transformed into an extracellular gel. This is consistent with longstanding observations that fibrin clots are not found in protochordates [13]. The lack of extended carboxyl terminal segments containing lysine and glutamine further implies that the protein is not a target for transglutaminase cross-linking.

However, urochordates like *C. intestinalis* do have circulating cells that are known to clump at sites of injury [13], and the protein could very well play a role in that process. In vertebrates, fibrinogen is well known to bind to platelets and various white cells quite independently of its ability to be changed into fibrin. Such interactions could be mediated either by the RGD sequence present in the C chain or by the binding of cryptic cell surface moieties, exposed by traumatic circumstances or infection, perhaps, to the molecule's well conserved 'holes'. Whatever its role, the protein is clearly derived from the same ancestral stock as vertebrate fibrinogen.

It has long been presumed that the prototypic vertebrate fibrinogen was composed of 3 identical chains, with globular carboxyl terminal regions for all three, in contrast to extant vertebrate fibrinogens which have only 2 globular domains at their termini. The conjecture was that the α chain suffered a genetic dislocation early in vertebrate evolution, the result of which was the mostly unfolded carboxyl terminal region that exists today [14]. Models of how that ancient protein would have been configured have been proposed [15].

References

- 1 Doolittle RF: Coagulation in vertebrates with a focus on evolution and inflammation. *J Innate Immun* 2011;3:9–16.
- 2 Doolittle RF: A detailed consideration of a principal domain of vertebrate fibrinogen and its relatives. *Protein Sci* 1992;1:1563–1577.
- 3 Wang X, Zhao Q, Christensen BM: Identification and characterization of the fibrinogen-like domain of fibrinogen-related proteins in the mosquito, *Anopheles gambiae*, and the fruitfly, *Drosophila melanogaster*, genomes. *BMC Genomics* 2005;6:114.
- 4 Zhang SM, Zeng Y, Loker ES: Expression profiling and binding properties of fibrinogen-related proteins (FREPs), plasma proteins from the schistosome snail host *Biomphalaria glabrata*. *Innate Immunol* 2008;14:175–189.
- 5 Doolittle RF: Structural aspects of the fibrinogen to fibrin conversion. *Adv Protein Chem* 1973;27:1–105.
- 6 Yang Z, Kollman JM, Pandi L, Doolittle RF: Crystal structure of native chicken fibrinogen at 2.7 Å resolution. *Biochemistry* 2001;40:12515–12523.
- 7 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 USA* 2003;100:7527–7532.
- 8 Dehal P, Satou Y, Campbell RK, et al: The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 2002;298:2157–2167.
- 9 Altschul SF, Madden T, 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;25:3389–3402.
- 10 Doolittle RF: Of URFs, ORFs: A Primer on How to Analyze Derived Amino Acid Sequences. Mill Valley, University Science Books, 1987.
- 11 Nagy A, Hegyi H, Farkas K, Tordai H, Kozma E, Banyai L, Patthy, L: Identification and correction of abnormal, incomplete and mispredicted proteins in public databases. *BMC Bioinformatics* 2008;9:353.
- 12 Kairies N, Beisel HG, Fuentes-Prior P, Tsuda R, Muta T, Iwanaga S, Bode W, Huber R, Kawabata S: The 2.0-Å crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems. *Proc Natl Acad Sci USA* 2001;98:13519–13524.
- 13 Rowley AF, Rhodes CP, Ratcliffe NA: Protochordate leucocytes: a review. *Zool J Linn Soc* 1984;80:283–295.
- 14 Doolittle RF, Watt KW, Cottrell BA, Strong DD, Riley M: The amino acid sequence of the α-chain of human fibrinogen. *Nature* 1979;280:464–468.
- 15 Doolittle RF, Spraggon G, Everse SJ: Evolution of vertebrate fibrin formation and the process of its dissolution; in Bock GR, Goode JA (eds): *Plasminogen-Related Growth Factors*, CIBA Foundation Symposium 212. New York, Wiley, 1997.