Vol. 167, No. 1, 1990
February 28, 1990

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 16-19

THE AMINO-TERMINAL SEQUENCE OF LOBSTER FIBRINOGEN REVEALS COMMON ANCESTRY WITH VITELLOGENINS

Russell F. Doolittle and Marcia Riley

Center for Molecular Genetics, M-034
University California, San Diego
La Jolla, CA 92093

Received January 8, 1990

SUMMARY The amino-terminal sequence of lobster "fibrinogen" was determined. A computer search of the sequence revealed that the lobster protein is homologous to vitellogenins (precursors of egg-yolk proteins).

Crustaceans have long been known to have an extracellular protein circulating in the hemolymph that clots *in vivo* under various conditions of trauma and gels *in vitro* upon the addition of cell and tissue extracts (1-4 *inter alia*). Extracts of the circulating cell called the hemocyte are particularly effective (4). Moreover, cell extracts from crustaceans as different as crabs and lobsters will effectively gel the "fibrinogen" of the other (6). From the very earliest studies, evidence has steadily accumulated that distinguishes the coagulation from the clotting process that occurs in vertebrates. Proteolysis is not involved in the crustacean event, the same amino-terminals being present before and after gelation (5). Rather, the basis of the coagulation is a transglutaminase-catalyzed reaction that results in the formation of intermolecular γ-glutamyl-ϵ-lysine crosslinks (5). Beyond that, the lobster fibrinogen molecule is composed of two apparently identical chains of approximately 200,000 molecular weight (7), whereas all vertebrate fibrinogens are made up of two each of three non-identical chains.

Given the clear differences between the crustacean and vertebrate fibrinogens, we have been pondering their separate evolutionary origins. In this regard, we recently resumed a longstanding inquiry into the nature of lobster fibrinogen. As a first step toward cloning its cDNA, we subjected a purified preparation to amino-terminal sequence determination. Fortuitously, a computer search of known protein sequences revealed a clue as to the evolutionary origins of this molecule.

MATERIALS AND METHODS

Lobster Fibrinogen. Spiny lobsters (*Panulirus interruptus*) were provided by the Collecting Facility at Scripps Institution of Oceanography. Blood was collected directly into cold trisodium citrate (final concentration = 0.02M) and cellular material immediately removed by centrifugation. Crude fibrinogen was prepared by isolelectric precipitation at pH 5.0 followed by gel filtration chromatography on agarose A1.5 in a low ionic strength buffer containing EDTA (4). Pure lobster "fibrinogen B" (4) was reprecipitated at pH 5.0 and stored as a paste.

Sequence Determination. Approximately 1 mg (ca 5 nanomoles) of lobster fibrinogen B was dissolved in water and dialyzed extensively to remove the last traces of buffer ingredients. A

0006-291X/90 $1.50

50 μ liter aliquot (ca 500 picomoles) of the solution was used for an extended run on an Applied Biosystems Gas Phase Sequencer. The operation was performed by the U.C.S.D. Microsequencing Facility.

Computer Programs. The amino-terminal sequence found for lobster fibrinogen was searched against our own sequence collection (NEWAT) and the National Biomedical Research Foundation protein database, version 21. The searching routine employed programs of our own devising, as described in Ref. 8. Subsequently, a multiple sequence alignment was constructed by the progressive alignment procedure (9).

RESULTS

The automatic sequencer run was continued through 37 cycles, although beginning with step 30 multiple amino acids were found. The amino-terminal residue was exclusively leucine, in line with the endgroup previously reported (5). Step 10 yielded approximately equal amounts of lysine (90 pmoles) and arginine (68 pmoles), and step 14 gave 68 pmoles of isoleucine and 56 pmoles of arginine. Similarly, step 20 had both proline (111 pmoles) and lysine (60 pmoles).

A computer search of the sequence identified vitellogenin sequences from *Caenorhabditis elegans* as being 30% identical. Only four other candidate sequences were identified, all at lower scores: human factor V, an opacity protein from *Neisseria gonorrhoeae*, a virus genome polyprotein and the FIXC protein from *Rhizobium meliloti*. Closer scrutiny showed the latter resemblances to be spurious.

In contrast, the first 30 amino acids of lobster fibrinogen were readily aligned with the amino-terminal regions of vitellogenins from both invertebrates and vertebrates (Fig. 1). In this regard, the lobster sequence is as similar to the *C. elegans* vitellogenin sequences as the latter are to vertebrate vitellogenin sequences in this region.

DISCUSSION

Vitellogenins are large molecular weight proteins found in the blood plasma or hemolymph of egg-laying animals, both vertebrate and invertebrate. They are the precursors of the major egg-yolk proteins and are synthesized in various tissues: the liver in vertebrates (10), for example, the intestine in nematodes (11) and the fat body in insects (12). The proteins

```
                          *  *              *  *
Chicken          ..F N S R R S Y L Y N Y E G S M L N G L Q D R S L G K A G V R..
Xenopus A        ..F S E S K T S V Y N Y E A V I L N G F P E S G L S R A G I K..
Xenopus B        ..F G E S K P Y V Y N Y E G I I L N G I P E N G L A R S G I I..
C. elegans 1     ..F E P K I D Y H Y K F D G L V L S G L P T A S S E L S Q S R..
C. elegans 5     ..F S P K S E Y V Y K F D G L L L S G L P T R S S D A S Q T L..
                    .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Lobster            L Q P K L E Y Q Y K Y H G I V A L G I P S Y K T Q F Y D A H..
                             (R)    (R)       (K)              (R)
```

Figure 1. Alignment of lobster fibrinogen amino-terminal sequence with corresponding regions of vitellogenins from chicken (21), *Xenopus* (22) and *C. elegans* (23). The chicken, *Xenopus* and *C. elegans* sequences are translated from DNA sequences and are preceded in each case by putative signal peptides. Asterisks (*) denote positions in which four or five of the five comparison sequences are identical to the lobster sequence. Dots (·) denote positions where at least one of the other sequences is identical to the lobster sequence.

are secreted into the general circulation (plasma or hemolymph) and taken up preferentially by the ovary. Processing in the way of partial proteolysis occurs in the oocyte, after which these proteins are usually referred to as vitellins or lipovitellins (13).

There are a number of reasons why vitellogenins are logical candidates for the evolutionary antecedents of crustacean fibrinogen. For one, the two proteins have approximately the same molecular weights and are secreted into the hemolymph in abundance where they then circulate as dimers. Moreover, both proteins are glycoproteins that bind substantial amounts of lipid (4,10). Indeed, it is not at all difficult to envision how the liberation of a cellular transglutaminase in response to injury could crosslink a vitellogenin homologue in a way that would lead to a polymeric gel network.

In most egg-laying species, vitellogenin synthesis is limited to adult females (10,12). In some, however, including *C. elegans*, hermaphroditic animals synthesize these proteins (11), and in sea urchins, both sexes synthesize large amounts of vitellogenins (19). That being the case, adapting production of a vitellogenin homologue to both sexes does not appear to be a major obstacle.

Interestingly, vitellogenins have been shown to contain a 250-residue segment that is homologous with a repeated region of von Willebrand factor, a vertebrate protein tangentially involved in vertebrate blood clotting (16). The von Willebrand factor is, like many vertebrate plasma proteins, a mosaic of various repeated units whose evolution appears to be the result of exon shuffling (17), and it is the sections referred to as "D domains" that exhibit similarity to a region of vitellogenin. In this regard, also, it is interesting that vertebrate vitellogenins differ from invertebrate vitellogenins in that the latter lack the serine-rich domain denoted phosvitin (14,18).

Insect Clotting Protein. Insects are the other major class of mandibular arthropods, and it is natural to inquire whether their clotting systems are similarly based. It has been shown that hemolymph clotting in insects is, for the most part but not entirely, inhibited by the same agents that inhibit crustacean clotting (19). The major gel-forming protein has not yet been characterized. It is important that the heritage of this protein be established in order to pinpoint when the critical vitellogenin duplication leading to a clottable homologue occurred. Although many insects have large molecular weight vitellogenins that seem comparable to those found in *C. elegans*, a set of smaller ones has been characterized in *Drosophila* that appear unrelated to the large type (20). It is not impossible that a similar diversity could exist among the clotting proteins.

In summary, crustacean fibrinogen, as exemplified here by lobsters, appears to be descended from the more widely distributed vitellogenins. It is likely that the major insect hemolymph protein involved in clotting will prove to be similar.

ACKNOWLEDGMENT

This work was supported by N.I.H. Grants HL26873 and GM34434.

REFERENCES

1. Glavind, J. (1948) Studies on the Coagulation of Crustacean Blood. Arnold Busck, Copenhagen
2. Morrison, P. R. & Morrison, K. C. (1952) Biol. Bull. 103, 395-406
3. Duchateau, G. & Florkin, M. (1954) Bull. Soc. Chim. Biol. 36, 295-305

4. Fuller, G. M. & Doolittle, R. F. (1971) Biochemistry 10, 1305-1311
5. Fuller, G. M. & Doolittle, R. F. (1971) Biochemistry 10, 1311-1315
6. Ghidalia, W., Vendrely, R., Montmory, C., Coirault, Y. & Brouard, M. (1981) J. Comp. Physiol. 142, 473-478
7. Doolittle, R. F. & Fuller, G. M. (1972) Biochim. Biophys. Acta 263, 805-809
8. Doolittle, R. F. (1987) Of URFs and ORFs: A Primer On How to Analyze Derived Amino Acid Sequences. University Science Press, Mill Valley, CA
9. Feng, D. F. & Doolittle, R. F. (1987) J. Mol. Evol. 25, 351-360
10. Wahli, W., Dawid, I. B., Ryffel, G. U. & Weber, R. (1981) Science 212, 298-304
11. Kimble, J. & Sharrock, W. J. (1983) Develop. Biol. 96, 189-196
12. Bownes, M. (1986) Annual Rev. Entomol. 31, 507-531
13. Wyatt, G. R. & Pan, M. L. (1978) Annual Rev. Biochem. 47, 779-817
14. Spieth, J., Denison, K., Kirtland, S., Cane, J. & Blumenthal, T. (1985) Nucleic Acids Res. 13, 5283-5298
15. Shyu, A. B., Raff, R. A. & Blumenthal, T. (1986) Proc. Natl. Acad. Sci., U.S.A. 83, 3865-3869
16. Baker, M. E. (1988) Biochem. J. 256, 1059-1061
17. Bonthron, D. T., Handin, R. I., Kaufman, R. J., Wasley, L. C., Orr, E. C. Mitsock, L. M., Ewenstein, B., Loscalzo, J., Ginsburg, D. & Orkin, S. H. (1986) Nature 324, 270-273
18. Nardelli, D., Gerber-Huber, S., van het Schip, A. D., Gruber, M., AB, G. & Whali, W. (1987) Biochemistry 26, 6397-6402
19. Bohn, H. & Barwig, B. (1984) J. Comp. Physiol. B 154, 457-467
20. Yan, Y. L., Kunert, C. J. & Postlethwait, J. H. (1987) Nucleic Acids Res. 15, 67-85
21. Burch, J. B. E. (1984) Nucleic Acids Res. 12, 1117-1135
22. Walker, P., Brown-Luedi, M., Germond, J.-E., Wahli, W., Meijlink, F. C. P. W., van het Schip, A. D., Roelink, H., Gruber, M. & AB, G. (1983) EMBO J. 2, 2271-2279
23. Spieth, J., Denison, K., Zucker, E. & Blumenthal, T. (1985) Nucleic Acids Res. 13, 7129-7138