Protein Science (1992), *1*, 1225–1226. Cambridge University Press. Printed in the USA.
Copyright © 1992 The Protein Society

FOR THE RECORD

A common domain within the proenzyme regions of the *Drosophila* snake and easter proteins and *Tachypleus* proclotting enzyme defines a new subfamily of serine proteases

CYNTHIA L. SMITH AND ROBERT DELOTTO
Program in Molecular Biology, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, and Graduate Program in Molecular Biology, Cornell University Graduate School of Medical Sciences, New York, New York 10021

(RECEIVED June 24, 1992; REVISED MANUSCRIPT RECEIVED July 6, 1992)

Several of the genes required maternally in *Drosophila melanogaster* for the correct establishment of dorsal-ventral polarity in the developing embryo appear to encode serine proteases. These include the genes *snake* and *easter* (DeLotto & Spierer, 1986; Chasen & Anderson, 1989). The activities of the proteins encoded by these genes are located in the perivitelline space of the embryo and have been proposed to function in this compartment as part of a signal transduction process (Stein et al., 1991; Stein & Nuesslein-Volhard, 1992). The identification of snake and easter as serine proteases was based upon homologies to the carboxy-terminal (catalytic) portion of these proteins and other members of the serine protease superfamily. To date, no convincing homologies have been found within the proenzyme regions of these proteins and any other protein sequence. We examined the proenzyme regions of snake and easter and found that they are homologous to one another as well as to another invertebrate serine protease from the horseshoe crab, *Tachypleus tridentus.*

Discussion

We compared the distribution of cysteine residues of the snake and easter amino acid sequences and found that 14 cysteine residues can be aligned. Although seven of these cysteines lie within the catalytic region of these proteins, an additional seven of the cysteines lie within the proenzyme portion of the molecule. Furthermore, these 14 cysteine residues are also shared among snake, easter, and the invertebrate serine protease, proclotting enzyme from horseshoe crab (Muta et al., 1990). Proclotting enzyme is a serine protease zymogen of the endotoxin sensitive hemolymph coagulation system of limulus (*T. tridentus*). It is a component of an enzymatic cascade that catalyzes the conversion of coagulogen to coagulin, an insoluble gel, and is considered to be the prothrombin counterpart in the hemolymph coagulation system (Miyata et al., 1984). The proclotting enzyme proenzyme polypeptide (light chain) has a unique disulfide knotted domain that has been proposed to function as a recognition site for factor B, factor G, upstream activators, or coagulogen, the substrate. Proclotting enzyme has both O-linked and N-linked glycosylation at a number of defined residues within the light chain.

An alignment of the protein sequences of proclotting enzyme, snake, and easter is illustrated in Figure 1. Muta et al. (1990) have determined the complete disulfide bridge structure of proclotting enzyme and found three intrachain disulfide bridges within the proenzyme polypeptide as illustrated. An additional disulfide bridge exists between the proenzyme light chain and the catalytic chain between C_{89} and C_{219}, so that upon activation of proclotting enzyme light and heavy chains remain covalently linked (Nakamura et al., 1985). The amino-terminus of proclotting enzyme, Q_1, exists in the form of a blocked pyroglutamic acid when the enzyme is isolated from hemocytes. A reasonable similarity exists in the sequence at the amino-termini of proclotting enzyme and easter. The alignment reveals conservation in all three proteins of the six cysteines involved in the disulfide knotted domain as well as the two involved in the intrachain

Reprint requests to: Robert DeLotto, Program in Molecular Biology, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021.

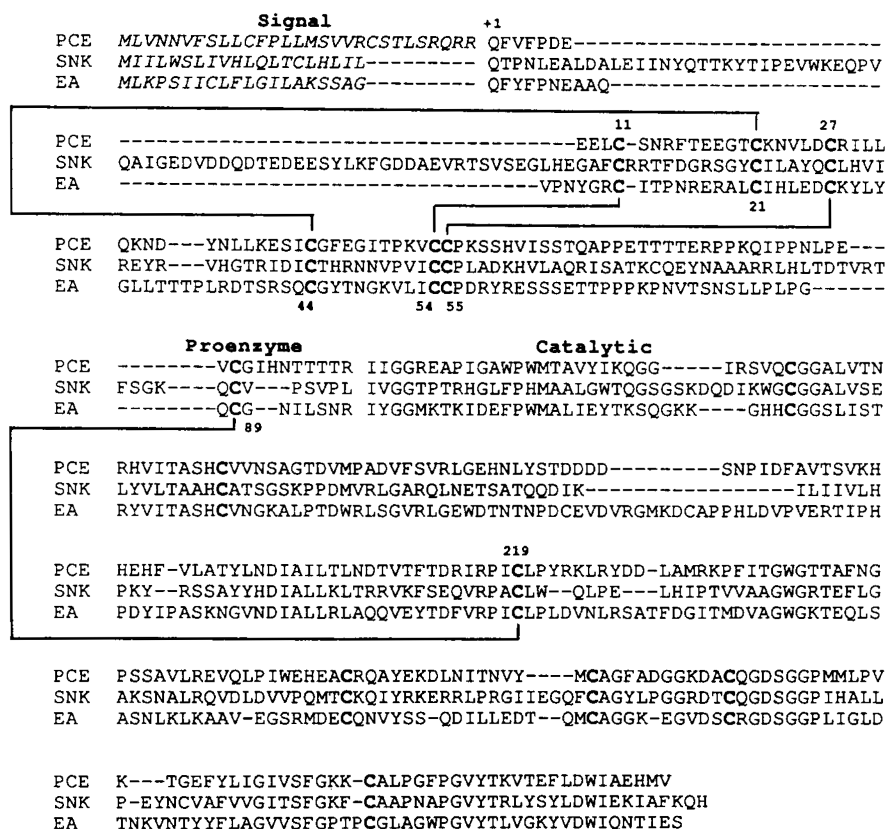

Fig. 1. Alignment of the predicted amino acid sequences of proclotting enzyme (PCE), snake, and easter. Signal peptide of PCE and predicted signal peptides for snake (SNK) and easter (EA) are in italics. Numbering is based upon proclotting enzyme, as described by Muta et al. (1990). Conserved cysteine residues are in bold, and disulfide bonds of PCE are indicated by lines. Protein alignments are derived from Genbank accession numbers M58366, X04513 (revised), and J03154.

disulfide bridge. The conservation of the cysteine residues in the alignment suggests that these proteins share a common functional domain. They appear to be closely related members of a new subfamily of invertebrate regulatory serine proteases. The primary difference between snake and the other two proteins is the addition of an acidic domain (Asp- and Glu-rich) in its amino-terminus.

Several specific predictions follow from the protein sequence alignment: (1) Because both C_{89} and C_{219} are conserved in snake and easter, the light and heavy chains of these two proteases probably remain covalently attached via an intrachain disulfide bond after activation. (2) Q_{21} of easter is probably the amino-terminus of the proenzyme and may exist in the form of a blocked pyroglutamate residue in vivo. (3) All three proteins share a protein domain, called a disulfide knot, which has been proposed to function as a receptor for the upstream activator or the substrate of proclotting enzyme. The conservation in structure of this domain presumably reflects a similarity of function, suggesting that it has a similar role in snake and easter. (4) Proclotting enzyme is extensively glycosylated within a Ser and Thr rich region between C_{55} and P_{88} (Muta et al., 1990). The corresponding portions of snake and particularly easter are also Ser- and Thr-rich and therefore may be similarly glycosylated in vivo. We suggest that future biochemical characterization of snake and easter should address these predictions.

The homology described here provides an evolutionary link between two serine proteases that are involved in pattern formation in the early embryo and a member of the primitive coagulation system of the horseshoe crab. It suggests that a specialized response to trauma and infection may have evolved from a common ancestor to proteins that are presently required for organization of the body plan of an insect embryo.

References

Chasen, R. & Anderson, K.V. (1989). The role of *easter*, an apparent serine protease, in organizing the dorsal-ventral pattern of the *Drosophila* embryo. *Cell 56*, 391-400.

DeLotto, R. & Spierer, P. (1986). A gene required for the specification of dorsal-ventral patterning in *Drosophila* appears to encode a serine protease. *Nature 323*, 688-692.

Miyata, T., Hiranaga, M., Umezu, M., & Iwanaga, S. (1984). Amino acid sequence of the coagulogen from *Limulus polyphemus* hemocytes. *J. Biol. Chem. 259*, 8924-8933.

Muta, T., Hasimoto, R., Miyata, T., Nishimura, H., Toh, Y., & Iwanaga, S. (1990). Proclotting enzyme from horseshoe crab hemocytes. *J. Biol. Chem. 265*, 22426-22433.

Nakamura, T., Morita, T., & Iwanaga, S. (1985). Intracellular proclotting enzyme in limulus (*Tachypleus tridentus*) hemocytes: Its purification and properties. *J. Biochem. 97*, 1561-1574.

Stein, D. & Nuesslein-Volhard, C. (1992). Multiple extracellular activities in *Drosophila* egg perivitelline fluid are required for establishment of embryonic dorsal-ventral polarity. *Cell 68*, 429-440.

Stein, D., Roth, S., Vogelsang, E., & Nuesslein-Volhard, C. (1991). The polarity of the dorsoventral axis in the *Drosophila* embryo is defined by an extracellular signal. *Cell 65*, 725-735.