Volume 12 Number 6 1984 Nucleic Acids Research

Use of a cDNA recombinant for the γ-subunit of mouse nerve growth factor to localize members of this multigene family near the TAM-1 locus on chromosome 7

Philip N.Howles, Douglas P.Dickinson, Laura L.DiCaprio, Mary Woodworth-Gutai* and Kenneth W.Gross

Department of Molecular Biology and *Department of Cell and Tumor Biology, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, NY 14263, USA

Received 12 December 1983; Revised and Accepted 24 February 1984

ABSTRACT

The γ-subunit of mouse 7S nerve growth factor (γ-NGF) is a member of a family of closely related serine proteases that includes kallikreins and tamases. We have isolated from a DBA/2J male submaxillary gland cDNA library a clone, pSM676, which codes for γ-NGF. Sequence analysis of the clone shows that it codes for the C-terminal 138 amino acids of the protein plus 23 bases of the 3'-nontranslated portion of the message. The predicted amino acid sequence agrees with that determined by Thomas et al. (1) for the γ-subunit of nerve growth factor from Swiss Webster mice except for the single, conservative substitution of glutamate for aspartate at amino acid 175. When used as a probe for Southern blot analysis, pSM676 hybridizes to at least twelve fragments of restricted mouse genomic DNA which correspond to several different serine protease genes. Using mouse-hamster hybrid cell lines and recombinant inbred strains of mice, we have demonstrated that all of the genes which show homology to pSM676 are located on mouse chromosome 7, clustered at or near the Tam-1 locus.

INTRODUCTION

Trypsin-like serine proteases are a family of enzymes involved in digestion, blood coagulation and the processing of kinins from their precursors (2). They are also thought to be involved in the processing or activation of other biologically important proteins such as nerve growth factor (NGF) and epidermal growth factor (EGF) (3-5). Different members of this family can be distinguished by electrophoretic mobility, utilization of different synthetic substrates such as tosyl arginine methyl ester (TAME), or by the efficiency with which they process kininogens to kinins (kallikrein activity) (6-9).

The male mouse submaxillary gland (SMG) is a rich source of several of these serine proteases as well as a number of physiologically important proteins such as renin, NGF and EGF. These proteins are synthesized in granular convoluted tubule cells and stored in secretory granules until α-adrenergic agents stimulate their release into the saliva. Curiously, production of these proteins appears to be modulated by androgens, either

directly or indirectly, such that levels in the SMG of adult males or testosterone treated females are several fold higher than in glands of adolescent males or untreated females. While some of these proteins have defined biological functions, the roles of others and the physiological significance of their presence in the SMG is not yet clear (10-12).

NGF is stored in the submaxillary gland as a 7S complex consisting of two molecules of the active polypeptide hormone (β-NGF), two molecules of a specific serine protease responsible for processing the precursor of the β-subunit (γ-NGF), and two molecules of a third, stabilizing subunit (α-NGF) which has some homology to the γ-subunit but whose function is not yet completely understood (13,14). Similarly, EGF is complexed with a specific serine protease (EGF-BP) in the submaxillary gland (15).

Despite the varied biological roles of these serine proteases, their synthesis in the SMG appears to be coordinated (10,12). Some of these protease activities, specifically EGF-BP (Loren Skow, pers. comm.), proteases A & E (16), and TAM-1 (17), have previously been shown to map at or near the Tam-1 locus on chromosome 7 in the mouse. We have chosen to study the genetic organization of the genes encoding these proteases with the aim of determining the size of this family and the extent to which the genes are clustered, and also with the goal of understanding how their coordinate and tissue specific expression is achieved.

MATERIALS AND METHODS

Materials

All chemicals and enzymes were purchased from commercial suppliers and used according to their recommendations. AMV reverse transcriptase was obtained from Dr. Joseph Beard, Life Sciences, Inc., St. Petersburg, Florida. Eight to ten week old male mice of recombinant inbred strains were purchased from The Jackson Laboratory, Bar Harbor, Maine. Other mice were obtained from West Seneca Labs, West Seneca, New York, or from departmental stocks. DNA from hamster and mouse-hamster hybrid cell lines was a gift from Dr. Susan Naylor, Dept. of Human Genetics, Roswell Park Memorial Institute, Buffalo, New York. The cell lines used were constructed by Dr. Peter Lalley, Oak Ridge National Laboratory, Oak Ridge, Tennessee (18). The *in vitro* translation extracts were a gift from Dr. William Held, Dept. of Cell & Tumor Biology, Roswell Park Memorial Institute. The dodecamer polynucleotide was purchased from Collaborative Research.

Fractionation and In Vitro Translation of RNA

Polyadenylated RNA was purified from SMG total RNA (19) and size fractionated on a 10% to 30% neutral sucrose gradient as previously described (20). RNA from individual fractions was translated in vitro using ^{35}S-methionine (21). The polypeptides synthesized were separated according to size by discontinuous SDS polyacrylamide gel electrophoresis (22) on 11% separating gels containing 6M urea and visualized by autoradiography of the dried gels. Those fractions enriched for RNA's stimulating the synthesis of polypeptides 25,000 to 35,000 daltons in size were pooled, ethanol precipitated, and used as templates for the preparation of cDNA probes.

Library Screening

The construction of the male DBA/2J SMG cDNA library has been described (23). The 729 clones obtained were screened by the colony filter method of Grunstein and Hogness (24) as modified by Pietras (25). For the selection of clones corresponding to androgen modulated sequences, cDNA probes were made from total male and female RNA and differential hybridization was carried out as described by Berger et al. (26). Other cDNA probes were prepared from appropriate fractions of mRNA using either oligo dT or the dodecamer 5'-d (GTTCCAGTGTTT)-3' as a primer for the reverse transcriptase.

Plasmid Isolation, Sequencing, and Message Recapture

Plasmid DNA was prepared by the cleared lysate method (27), and appropriate restriction fragments were sequenced by the method of Maxam and Gilbert (28).

Plasmid DNA was bound to nitrocellulose circles and used to select message from total male SMG mRNA as described previously (23). The selected mRNA's were eluted, ethanol precipitated and translated in vitro, and the products analyzed on polyacrylamide gels as described above.

Southern Analysis

High molecular weight genomic DNA was prepared from male mouse liver as described (23), digested with restriction endonucleases and electrophoresed through horizontal 0.8% agarose gels in 40 mM Tris acetate pH 7.8, 5 mM sodium acetate, 1 mM EDTA at 5 v/cm. The DNA was transferred to nitrocellulose according to Southern (29) and hybridized to nick-translated plasmid DNA (30) (2-3 x 10^8 dpm/µg) in 10% dextran sulfate, 4 X SSC (29), 5X Denhardt's (31), 0.1% SDS, 0.1% sodium pyrophosphate, and 0.1 mg/ml denatured salmon sperm DNA at 65°C. ^{32}P-labelled probe was

present at 1 x 10^6 dpm/ml. Volume was 25 to 50 ml. After 16 to 20 hours of hybridization, the nitrocellulose was washed at 65°C for 5 to 6 hours in several changes of 2 X SSC, 0.1% SDS, 0.1% sodium pyrophosphate. The filter was then exposed to Kodak XAR film at -70°C for 18 hours with one intensifying screen.

RESULTS

Characterization of Androgen Regulated Messages from Mouse SMG

When total RNA from adult male and female mouse submaxillary glands is translated *in vitro* and the products are separated on an SDS-urea-polyacrylamide gel, a number of 25,000 to 35,000 dalton polypeptides are seen whose levels are substantially elevated in the male and testosterone treated female (Fig. 1, lanes c to h). Prominent among these is a polypeptide of 28,000 daltons which is polymorphic among the three strains of mice presented (see arrow). Strain distribution patterns and recombinant inbred (RI) strain analysis of *in vitro* translation products strongly suggest that this polypeptide represents the Tam-1 gene product described by Skow (17). A number of other polypeptides in this region of the gel also appear to map to the Tam-1 locus (data not shown).

These androgen regulated messages are greatly enriched in the sucrose gradient fraction represented by Fig. 1, lane i. Besides those for TAM-1 and related serine proteases, this fraction is expected to contain the messages for all the subunits of NGF and EGF-BP based on the known sizes of the precursor polypeptides of these proteins (1,14,32).

Isolation and Characterization of a cDNA Clone for γ-NGF

As an initial screen for clones encoding serine proteases and, in particular, EGF-BP and NGF subunits, ^{32}P-labelled cDNA prepared from the sucrose gradient fraction described above was hybridized to a cDNA library made from male DBA/2J SMG mRNA (23). As the serine proteases and growth factors are known to be hormonally regulated (10-12,17), the library was also screened with ^{32}P-labelled cDNA prepared from adult male and female total RNA to identify clones corresponding to mRNAs whose levels are androgen modulated. Since we were also interested in obtaining a recombinant for β-NGF, positive clones were further screened with a cDNA probe prepared from the same gradient fraction described above but primed by the dodecamer 5'-d(GTTCCAGTGTTT)-3'. This polynucleotide is complementary to the codons for Lys-His-Trp-Asn (5'-AAACACUGGAAC-3'),

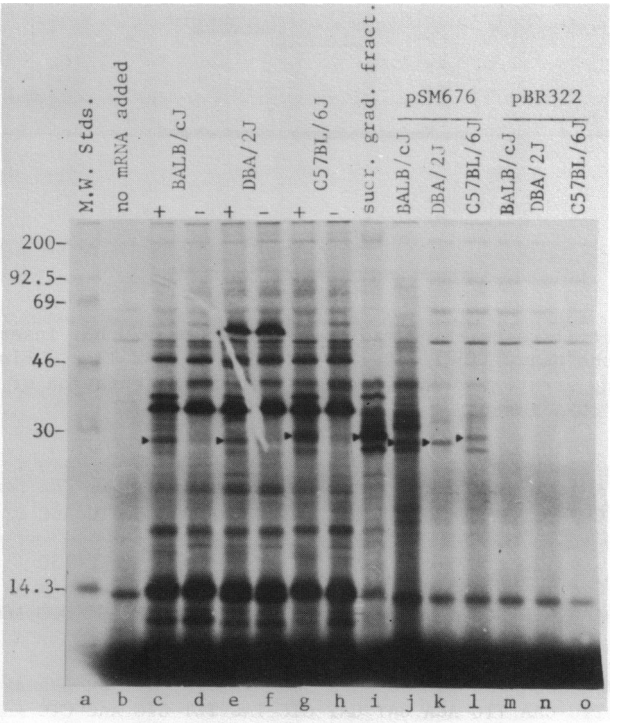

Figure 1. In vitro translation products of SMG mRNA from different sources. Lane a: molecular weight standards given in kilodaltons; Lane b: in vitro translation control without added message; Lanes c-h: polypeptides encoded by SMG mRNA from normal females (-) or testosterone-induced females (+) of the indicated strains; Lane i: polypeptides encoded by sucrose gradient fractionated SMG mRNA from AKR/J males; Lanes j-l: polypeptides encoded by SMG mRNA recaptured by pSM676; Lanes m-o: recapture controls using pBR322. The solid triangle marks the testosterone-induced polypeptide which correlates with the TAM-1 serine protease. Note, also, the testosterone-inducible family of polypeptides accentuated in the sucrose gradient fraction but visible, as well, in the lanes from total mRNA and pSM676 recaptured mRNA.

representing amino acids 74 to 77 of mature β-NGF (33). Of the clones which scored positive with all probes, pSM676, which contains a 520 base pair insert, was chosen for further analysis (Fig. 2).

Much to our surprise, instead of encoding β-NGF, pSM676 was found by DNA sequence analysis to encode 138 amino acids of the C-terminal portion of γ-NGF including the highly conserved active site of serine proteases (Fig. 3, asterisks) (34), and also contains 23 bases of the 3'-nontranslated region of the message (Fig. 3). Except for the

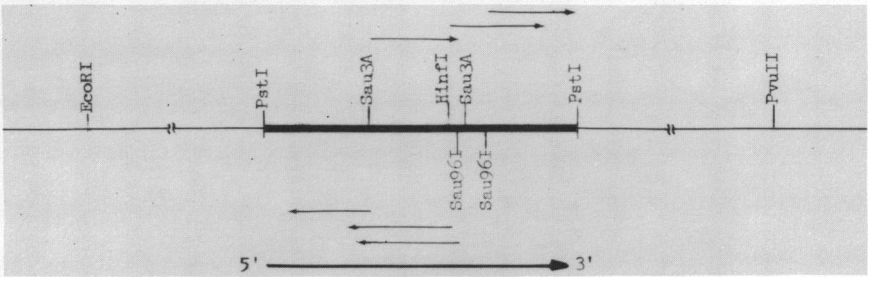

Figure 2. Partial restriction map of pSM676. The 520 bp. insert is indicated by the heavy line, pBR322 sequences by thin line. Fine arrows show sequencing strategy. Heavy arrow shows the orientation of the sense strand of the cDNA insert.

```
        100                                                 110
     Ser Lys Pro Ala Asp Ile Thr Asp Thr Val Lys Pro Ile Thr Leu Pro Thr Glu
5'...AGC AAG CCT GCT GAC ATC ACA GAT ACT GTG AAG CCC ATC ACC CTG CCC ACT GAG

                120                                                 130
     Glu Pro Lys Leu Gly Ser Thr Cys Leu Ala Ser Gly Trp Gly Ser Ile Thr Pro
     GAG CCC AAG CTG GGG AGC ACA TGC CTA GCC TCA GGC TGG GGC AGC ATT ACA CCC

                        140                                             150
     Thr Lys Phe Gln Phe Thr Asp Asp Leu Tyr Cys Val Asn Leu Lys Leu Leu Pro
     ACC AAA TTC CAA TTC ACA GAT GAT CTC TAC TGT GTG AAC CTC AAG CTC CTG CCT

                                160                                         170
     Asn Glu Asp Cys Ala Lys Ala His Ile Glu Lys Val Thr Asp Ala Met Leu Cys
     AAT GAG GAC TGT GCC AAA GCC CAC ATA GAG AAG GTG ACA GAT GCC ATG CTG TGT

                                        180
                                        #
     Ala Gly Glu Met Glu Gly Gly Lys Asp Thr Cys Lys Gly Asp Ser Gly Gly Pro
     GCA GGA GAG ATG GAA GGA GGC AAA GAC ACT TGC AAG GGT GAC TCA GGA GGC CCA
                                           *   *   *   *   *   *   *   *
        190                                                 200
     Leu Ile Cys Asp Gly Val Leu Gln Gly Ile Thr Ser Trp Gly His Thr Pro Cys
     CTG ATC TGT GAT GGT GTT CTC CAA GGT ATC ACA TCA TGG GGC CAT ACC CCA TGC
      *   *   *
                210                                                 220
     Gly Glu Pro Asp Met Pro Gly Val Tyr Thr Lys Leu Asn Lys Phe Thr Ser Trp
     GGT GAA CCT GAT ATG CCA GGC GTC TAC ACC AAA CTT AAT AAG TTT ACC TCC TGG

                        230
     Ile Lys Asp Thr Met Ala Lys Asn Pro END
     ATA AAA GAC ACT ATG GCA AAA AAC CCC TAA GTATCACAACTGTCTGTTCT(C40) -3'
```

Figure 3. Nucleotide sequence of pSM676 sense strand and corresponding amino acid sequence. # indicates the single difference between the amino acid sequence encoded by pSM676 and that reported by Thomas et al. (1). Asterisks mark the highly conserved amino acid sequence forming the active site of serine proteases. Numbers refer to the amino acid sequence for γ-NGF reported by Thomas et al. (1).

conservative change from aspartate to glutamate at position 175, the amino acid sequence encoded by pSM676 is identical to that reported by Thomas et al. (1) for γ-NGF. Since the amino acid sequence was determined from Swiss Webster mice and the nucleotide sequence from DBA/2J, this single difference may simply reflect a polymorphism between mouse strains.

Analysis of mRNAs Recaptured by pSM676

pSM676 was used to hybrid-select message from poly A+ RNA of adult male SMG. When the recaptured messages were translated in vitro and the products analyzed on polyacrylamide gels, we found that pSM676 selected messages coding for at least four to six different polypeptides, depending on mouse strain and gel resolution (Fig. 1, lanes j-l), including the Tam-1 protein and a group of polypeptides which appear to map to the Tam-1 locus as well (see above). This result demonstrated that while pSM676 codes specifically for γ-NGF, it is highly homologous to a number of other SMG mRNAs, at least one of which encodes a serine protease that has been mapped to the Tam-1 locus.

Southern Analysis of Recombinant Inbred Lines of Mice

Because other androgen modulated SMG serine proteases, including EGF-BP (Skow, pers. comm.), have been shown to map to the vicinity of the Tam-1 locus on chromosome 7, we wished to determine whether or not γ-NGF also maps to this region. When ^{32}P-labelled pSM676 was used to probe Southern blots of Eco RI digested DBA/2J and C57BL/6J genomic DNA under conditions of moderate stringency (.7 M NaCl, 65^{o}C, 10% dextran sulfate), it hybridized to at least twelve fragments, most of which are common to both strains (Fig. 4A, lanes 1 and 2). This complex pattern is due in part to hybridization to genes encoding serine proteases other than γ-NGF as would be predicted by the message recapture data presented above.

There are, however, five distinct differences between the hybridization patterns of DBA/2J and that of C57BL/6J. Since these strains are progenitors for a set of recombinant inbred strains of mice (BXD RI's) (35), DNA from this RI strain set was used to map these restriction fragment polymorphisms (RFP's). A BXD F_1 animal shows a hybridization pattern with pSM676 which combines that of each parent (Fig. 4A, lane 3) while each RI strain shows the pattern of one parent or the other. Comparison of the patterns for the 12 BXD RI strains shown, established that each of the RFP's mapped to the Tam-1 locus (Fig. 4A, lanes 4 to 9 and 13 to 18; Fig. 4B). Similar analyses using other restriction enzymes gave the same result (data not shown).

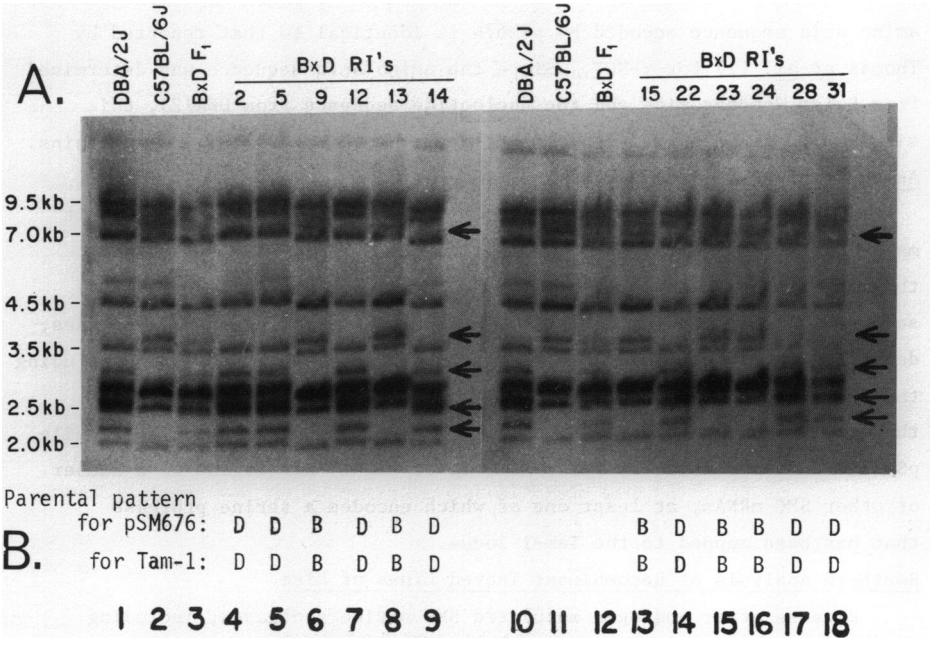

Figure 4. Southern blot of Eco RI cut, mouse genomic DNA probed with pSM676.
A. Arrows mark bands which are polymorphic between the parental strains (C57BL/6J (B) and DBA/2J (D)). Each BXD RI strain shows the entire pattern of one parent or the other, while a BXD F_1 animal shows a pattern that combines that of both parents.
B. First row of letters indicates the parental pattern exhibited by each RI strain when probed with pSM676. Second row letters indicate the segregation of the Tam-1 locus in these strains (17). Concordance is 100%.

Southern Analysis of Mouse-Hamster Cell Hybrids

Since these results did not allow us to genetically map fragments for which no polymorphism is detected, we used DNA from mouse-hamster hybrid cell lines to determine if these other fragments came from genes clustered on a single chromosome or dispersed throughout the genome. The set of cell lines employed was constructed from BALB/cJ spleen cells and the E36 hamster cell line (18).

DNA from BALB/cJ mice, from hamster cell line E36, and from the mouse-hamster hybrid cell lines indicated was restricted with Eco RI, fractionated, blotted and hybridized as before (Fig. 5). The BALB/cJ hybridization pattern is the same as that for C57BL/6J (compare Fig. 5, lane 1 with Fig. 4A, lane 2). The hamster DNA (Fig. 5, lane 2) shows

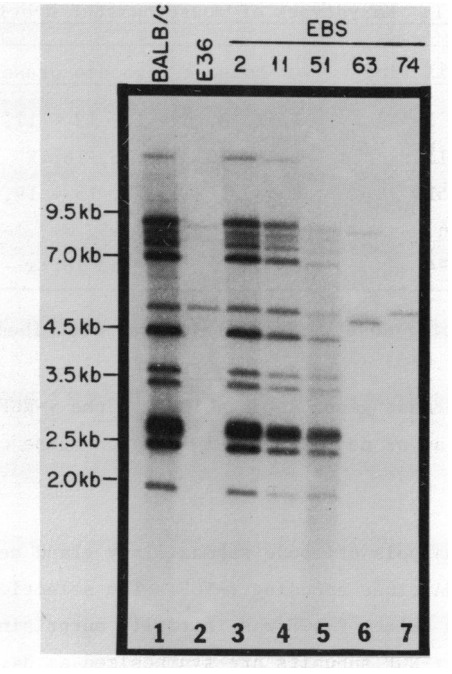

Figure 5. Southern blot of Eco RI cut DNA from mouse-hamster hybrid cell lines. Mouse and hamster parentals are in lanes 1 and 2. Hybrids are in lanes 3-7. Mouse pattern of hybridization with pSM676 is either entirely present or entirely absent. See Table 1 for chromosome constitution of the hybrids. The hamster DNA seems to be polymorphic, as well, as seen in lines 6 and 7 (see text also).

only two fragments hybridizing with pSM676 while the different hybrid cell lines show either all of the fragments characteristic of mouse DNA (Fig. 5, lanes 3 to 5) or none of them (Fig. 5, lanes 6 and 7). Polymorphisms appears to have arisen in the hamster DNA in hybrid line EBS63 and EBS74 as judged by the size shift of the 5 kb fragment and loss of the 9 kb fragment, respectively. As can be seen from Table 1, the presence or absence of the mouse DNA pattern correlates with the presence or absence of mouse chromosome 7 in the hybrid cell line. These data indicate that all serine protease genes detectable with pSM676 are located on mouse chromosome 7. Similar results, using a different serine protease probe, have recently been reported by Mason et al. (36).

These results, taken together with our results from the RI strain analysis and the message recapture experiment, suggest that γ-NGF, as

Table I. Karyotypes of mouse-hamster hybrids.

Hybrid cell line	Mouse chromosome present[1]
EBS 2	1, 2, 3, 6 - 10, 12 - 17, 19, X
EBS 11	1, 7, 10, 12, 15, 16, X
EBS 51	2, 3, 5 - 9, 12, 15 - 19, X
EBS 63	1, 10, 12, 13, 15, X
EBS 74	2, 3, 12, 14, 15, 16, X

[1]Karyotyping done by Peter Lalley as described (18).

well as all serine protease genes detectable with the γ-NGF cDNA probe, pSM676, are clustered at or near the Tam-1 locus on mouse chromosome 7.

DISCUSSION

As part of our analysis of mouse submaxillary gland serine proteases, we have isolated a cDNA clone encoding γ-NGF. The selection of this clone by a probe designed to select β-NGF was initially surprising. However, in the mouse SMG, β- and γ-NGF subunits are synthesized as 34,000 and 28,500 dalton precursor polypeptides, respectively (32,13,14), corresponding to messages which would have cofractionated under our sucrose gradient conditions. Furthermore, both of these polypeptides are known to be modulated by testosterone (11). Thus, our first two screens of the cDNA library would have selected clones for both of these gene products.

One possible reason why the β-NGF "specific" primer used in the third screen selected the γ-NGF cDNA clone is that the β- and γ-subunits have regions of homology, and the β-NGF cDNA probe cross-hybridized with the γ-NGF clone. However, a computer search of the β-NGF sequence reported by Scott *et al.* (32), using both strands of pSM676 revealed no significant homology. Another explanation is that the β-NGF primer was able to hybridize sufficiently to the γ-NGF mRNA to prime γ-NGF cDNA synthesis. Precedent for such aberrant priming exists in the literature (37), and there is a region (nucleotides 240 to 251) where the β-NGF primer shows 66% homology to the non-sense strand of pSM676. At present, we favor the notion that the selection of a γ-NGF cDNA clone by a β-NGF primer was fortuitous.

γ-NGF is a member of a family of highly conserved serine proteases which includes EGF-BP (15), the glandular kallikreins (2), γ-renin (8),

and the tamases (17). A comparison of the amino acid sequences for some of these proteases is shown in Figure 6. Table II lists percents of homology to γ-NGF at the amino acid and nucleotide level. These proteins

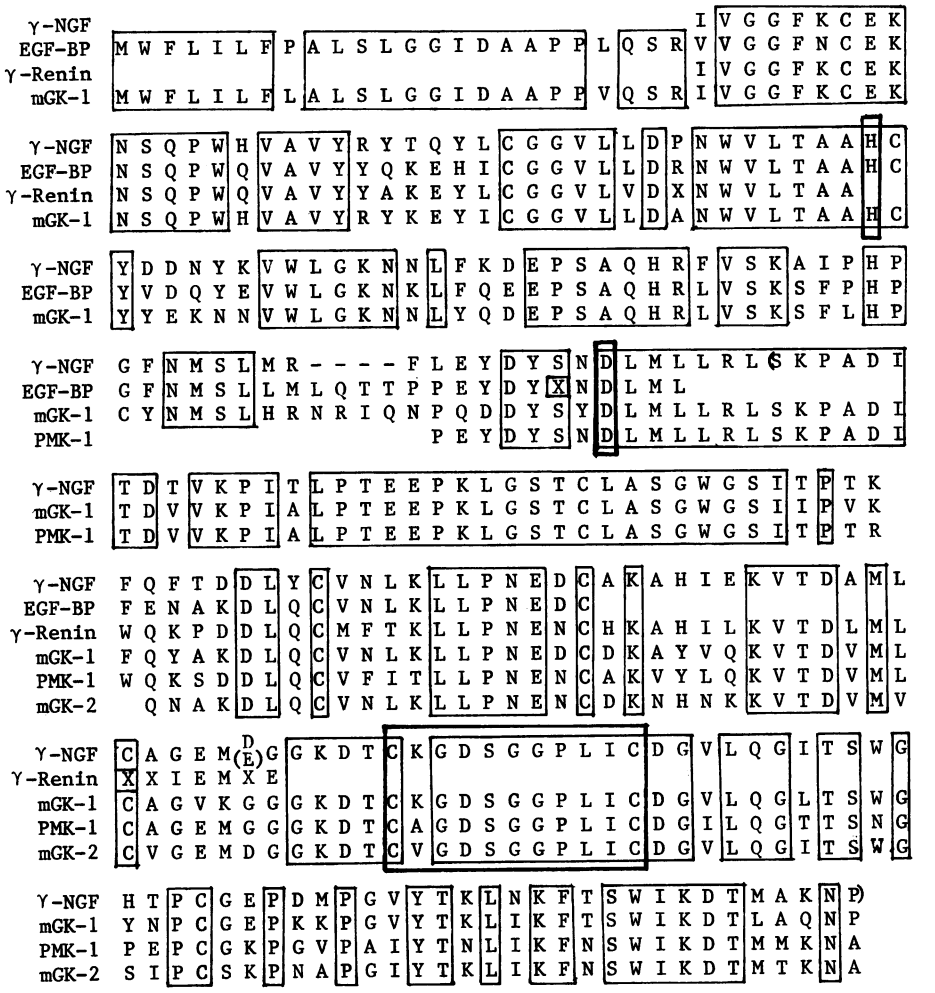

Figure 6. Alignment of amino acid sequences of mouse serine proteases. γ-NGF is from Thomas et al. (1), EGF-BP is from Ronne et al. (14) and Silverman (38) γ-renin is an SMG serine protease with renin-like activity reported by Poe et al. (8), mGK-1 and mGK-2 are from putative coding sequences on a genomic clone from Mason et al. (36), and pMK-1 is from an SMG serine protease cDNA clone reported by Richards et al. (39). Parentheses enclose pSM676 encoded amino acids of γ-NGF. Boxes enclose regions of homology. Heavily boxed residues comprise the active site.

Table II. Homology of other mouse serine protease to γ-NGF.

	% Homology to γ-NGF	
Serine Protease	Amino Acid	Nucleotide
EGF-BP	76	NA[1]
γ-Renin	75	NA[2]
pMK-1	77	83
mGK-1	80	89
mGK-2	75	84

[1]Not available. Regions of nucleotide sequence do not overlap.
[2]Not available. No nucleotide sequence available.

exhibit a remarkable degree of homology, especially at the nucleotide level, where degeneracy of the genetic code would predict less homology than at the amino acid level. Long stretches of nearly perfect homology can be seen, even in regions away from the active site common to all serine proteases (see heavily boxed region). The significance of this degree of homology is underscored by our results when pSM676 was used to select message from SMG RNA and by the complexity of the hybridization patterns seen when pSM676 was used to probe Southern blots. The extent of this homology is also demonstrated by the fact that under more stringent hybridization conditions (70% formamide, 0.5M NaCl, 45°C) the Southern blot hybridization patterns remained unchanged (data not shown).

The relative decrease in homology in residues adjacent to the active site and considerations of the structure of the serine protease molecule have led to the suggestion that the differences in sequence and coding potenial of the genes may be functionally related to the observed substrate preference of the different enzymes, and the specificity with which they interact with other polypeptides to form higher order complexes (as in the case of EGF-BP and γ-NGF) (4,34,36).

Our Southern analyses of mouse-hamster hybrid cell lines and recombinant inbred lines of mice, along with our message selection data, suggest that γ-NGF and all serine proteases detectable with pSM676 are clustered at or near the Tam-1 locus on mouse chromosome 7 as has previously been reported for other serine proteases (7,16). Our conclusion is further supported by the results of Mason et al. (36) who isolated a genomic clone containing putative coding sequences for

two kallikreins separated by less than 4,000 base pairs. They also reported results similar to ours using a different serine protease probe for Southern analysis of mouse-hamster hybrid cell lines.

Curiously, pSM676 hybridizes to only one or a few restriction fragments from hamster DNA (see Fig. 5). A similar, although less dramatic, result is apparent between mouse and rat DNA in a report using a different mouse serine protease probe (36). These data raise interesting questions regarding the evolution and function of the members of this multigene family.

One explanation for the limited number of sequences detected in hamster DNA is that the many sequences seen to hybridize to pSM676 in mouse DNA are the result of a mouse specific amplification of some of these genes that postdates the evolutionary divergence of this species from other rodents. The fact that these sequences appear to be highly conserved (note Table II) and tightly clustered in one region of the mouse genome may support this conclusion. Given potential differences in homology, probable differences in the distribution of exons among restriction fragments and overlapping fragment sizes, it is difficult to evaluate the precise number of homologous genes in each organism by the type of analysis described here. However, in the case of the hamster, the size polymorphism between cell lines observed for one of the homologous fragments suggests a minimal number of genes is being detected (see Fig. 5).

An alternative explanation for the limited cross-hybridization between pSM676 and hamster DNA is that there has been considerable divergence of sequences for members of this subfamily of serine proteases during the evolution of the hamster. However, this raises difficulties in interpreting the patterns of hybridization observed without invoking mechanisms acting differentially in the two species to maintain homology among the members of the family in the mouse but not in the hamster.

The biological significance of an amplification of the serine protease genes in the mouse is unclear. Since the hamster and the more closely related rat appear to have relatively fewer copies of the serine protease genes, any unique vital function performed by the products of the extra gene copies seen in the mouse would seem to be specific to that organism. However, the existence of apparent null alleles for Tam-1, PRT-5 and EGF-BP (16,17) and the fact that homozygotes carrying

these alleles are viable suggest that at least these members of the family are not essential. Finally, given conflicting reports in the literature (40,41), it would be of interest to know whether β-NGF is obligatorily found in association with γ-NGF at other sites of synthesis such as the guinea pig prostate (42).

Thus, we are presently entertaining the possibility that certain members of the serine protease gene family may have arisen from a mouse specific amplification event and that some of the gene products encoded may be of uncertain physiological relevance. An examination of various wild and inbred strains of mice and other closely related rodents using pSM676 should lead to a more complete understanding of the origin, organization and expression of these genes.

ACKNOWLEDGEMENTS

We wish to thank Dr. William Held for providing the in vitro translation system, Dr. Sue Naylor for providing the hamster and mouse-hamster cell line DNA's, Dr. Nina Piccini and Mr. John Knopf for assistance in preparation of the cDNA library, Loren Skow for his many helpful discussions with us, William Philbrick for performing the computer search of nucleotide sequences and for critically reviewing this manuscript, Ken Abel for technical assistance, and Cindy Bell and Nancy Holdsworth for preparing this manuscript. This work was supported by grants GM19521 and GM30248.

REFERENCES

1. Thomas, K.A., Baglan, N.C. and Bradshaw, R.A. (1981) J. Biol. Chem. 256, 9156-9166.
2. Schachter, M. (1980) Pharm. Rev. 31, 1-17.
3. Wilson, W.H. and Shooter, E.M. (1979) J. Biol. Chem. 254, 6002-6009.
4. Berger, E.A. and Shooter, E.M. (1977) Proc. Natl. Acad. Sci. USA 74, 3647-3651.
5. Frey, P., Forand, R., Maciag, T. and Shooter, E.M. (1979) Proc. Natl. Acad. Sci. USA 76, 6294-6298.
6. Boesman, M., Levy, M. and Schenkein, J. (1976) Arch. Biochem. Biophys. 175, 463-476.
7. Schaller, E. and vonDeimling, O. (1979) Anal. Biochem. 93, 251-256.
8. Poe, M., Wu, J.K., Forance, J.R., Rodkey, J.A., Benner, C.D. and Hoogsteen, K. (1983) J. Biol. Chem. 258, 2209-2216.
9. Bothwell, M.A., Wilson, W.H. and Shooter, E.M. (1979) J. Biol. Chem. 254, 7287-7294.
10. Bhoola, K.D., Dorey, G. and Jones, C.W. (1973) J. Physiol. 235, 503-522.
11. Hirata, Y. and Orth, D.N. (1979) Endocrinology 105, 1382-1387.
12. Chretien, M. (1977) Int. Rev. Cytol. 50, 333-396.
13. Thomas, K.A., Silverman, R.E., Jeng, T., Baglan, N.C. and Bradshaw, R.A. (1981) J. Biol. Chem. 256, 9147-9155.

14. Ronne, H., Lundgren, S., Severinsson, L., Rask, L. and Peterson, P.A. (1983) EMBO J. 2, 1561-1564.
15. Taylor, J.M., Mitchell, W.M. and Cohen, S. (1974) J. Biol. Chem. 249, 2188-2194.
16. Otto, J. and vonDeimling, O. (1981) Biochem. Genet. 19, 431-444.
17. Skow, L.C. (1978) Genetics 90, 713-724.
18. Lalley, P.A., Franke, U. and Minna, J.D. (1978) Proc. Natl. Acad. Sci. USA 75, 2382-2386.
19. Toole, J.J., Hastie, N.D. and Held, W.A. (1979) Cell 17, 441-448.
20. Dobberstein, B., Garoff, H., Warren, G. and Robinson, P.J. (1979) Cell 17, 759-769.
21. Held, W.A., West, K. and Gallagher, J.F. (1977) J. Biol. Chem. 252, 8489-8497.
22. Laemmli, U.K. (1970) Nature 227, 680-685.
23. Piccini, N., Knopf, J.L. and Gross, K.W. (1982) Cell 30, 205-213.
24. Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.
25. Pietras, D.F. (1981) Ph.D. Thesis, State University of New York at Buffalo.
26. Berger, F.G., Gross, K.W. and Watson, G. (1981) J. Biol. Chem. 256, 7006-7013.
27. Katz, L., Kinsbury, D. and Helinski, D.R. (1973) J. Bacteriol. 114, 577-591.
28. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
29. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
30. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Proc. Natl. Acad. Sci. USA 72, 1184-1188.
31. Denhardt, D.T. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.
32. Scott, J., Selby, M., Urdea, M., Quiroga, M., Bell, G.I. and Rutter, W.J. (1983) Nature 302, 538-540.
33. Angeletti, R.H., Hermodson, M.A. and Bradshaw, R.A. (1973) Biochemistry 12, 100-115.
34. Dayhoff, M.O., Hunt, L.T., McLaughlin, P.J. and Barker, W.C. (1972) in Atlas of Protein Sequence and Structure, Dayhoff, M.O., ed., Vol. 5, pp. D99-D103.
35. Taylor, B.A. (1978) in Origins of Inbred Mice, Morse, H.C., ed., pp. 423-438, Academic Press, New York.
36. Mason, A.J., Evans, B.A., Cox, P.R., Shine, J. and Richards, R.I. (1983) Nature 303, 300-307.
37. Panabieres, F., Fort, P., Piechaczyk, M., Blanchard, J.M. and Jeanteur, P. (1982) Gene 19, 321-326.
38. Silverman, R.E. (1977) Ph.D. Thesis, Washington University, St. Louis.
39. Richards, R.I., Catanzaro, D.F., Mason, A.J., Morris, B.J., Baxter, J.D. and Shine, J. (1982) J. Biol. Chem. 257, 2758-2761.
40. Chapman, Carey A., Banks, B.E.C., Carstairs, J.R., Pearce, F.L. and Vernon, C.A. (1979) FEBS Lett. 105, 341-344.
41. Rubin, J.S. and Bradshaw, R.A. (1981) J. Neurosci. Res. 6, 451-464.
42. Harper, G.P. and Thoenen, H. (1980) J. Neurochem. 34, 893-903.