

Characterization by cDNA Cloning of the mRNA for Seminalplasmin, the Major Basic Protein of Bull Semen

SABINE WAGNER, JOHANNES FREUDENSTEIN, and KARL HEINZ SCHEIT

ABSTRACT

A cDNA library derived from poly(A)⁺RNA of bull seminal vesicle tissue was screened with synthetic DNA probes specific for seminalplasmin (SAP), the major basic protein of bull semen. From a number of positive clones, pBSV12, containing a 577-bp insert, was identified and sequenced. The derived amino acid sequence comprises the known amino acid sequence of SAP with an amino terminal representing a putative signal sequence; at the carboxyl terminus the sequence contains an additional lysine residue. Present experimental data do not distinguish between two potential SAP precursor molecules, each starting with a methionine residue and differing by 10 amino acid residues in the leader peptide. Comparative Northern analysis reveals a SAP-specific mRNA of 700 bp, which lacks RNA from bovine testis as well as from seminal vesicle tissue of a bull calf; hence, expression of the SAP gene appears to be under androgen and/or developmental control. Southern analysis indicates that one gene appears to specify SAP. SAP-like DNA sequences were detected in ovine and porcine genomic DNA.

INTRODUCTION

BULL SEMINAL PLASMA contains a highly basic protein of 47 amino acids of known sequence (Theil and Scheit, 1983; Lewis *et al.*, 1985; Sitaram *et al.*, 1986). This protein, originally discovered by Reddy and Bhargava (1979) and named seminalplasmin (SAP), possesses a number of interesting biological properties, including antimicrobial activity (Reddy and Bhargava, 1979; Scheit *et al.*, 1986), calmodulin-antagonist function (Comte *et al.*, 1986), inhibition of uptake of Ca²⁺ ions by epididymal bovine spermatozoa (Rufo *et al.*, 1982), and inhibition of lymphocyte proliferation (Derwenkus *et al.*, 1989).

Although SAP could be purified from the secretion of bull seminal vesicles (Scheit, 1986), an attempt to demonstrate the biosynthesis of SAP in this gland by immunohistochemical methods failed (Aumüller and Scheit, 1987). The reason for this failure might have been the use of an anti-SAP antiserum of low titer and specificity. Only recently, employing monospecific anti-SAP-IgGs (Krauchs *et al.*, 1990), evidence for the biosynthesis of SAP in bovine seminal vesicles was obtained by immunohistochemistry and cell-free translation of poly(A)⁺RNA (Wempe *et al.*, 1990). Here we report the sequence of a cDNA clone, de-

rived from bull seminal vesicle tissue, that encodes the precursor of SAP.

MATERIALS AND METHODS

Seminal vesicles from slaughtered bulls and a male calf (2 weeks old) were collected immediately after killing. The secretion was squeezed out from the glands and the tissue was washed in ice-cold water, rapidly frozen, and stored in liquid nitrogen. Total RNA and poly(A)⁺RNA (Kemme *et al.*, 1986; Kemme and Scheit, 1988) were prepared as described. cDNA synthesis followed the protocol of Gubler and Hoffman (1983) using a commercial cDNA synthesis kit (BRL). The procedure of Villa-Komaroff *et al.* (1978) was used for the construction of recombinant pBR322. Transformation of *Escherichia coli* RR1 was performed as published (Hanahan, 1983). The yield of transformation was 1.6×10^6 transformants/ μg poly(A)⁺RNA. Recombinants (6×10^4) were screened on replica nitrocellulose filters (Millipore HATF) with a ³²P-labeled hybridization probe (sp. act., 2×10^8 cpm/ μg DNA). Prehybridization and hybridization were performed according to Hanahan and Meselson (1983) at 60°C. A total of 64 positive clones were identified. The *Pst* I insert of pBSV12 was purified

on a 1% agarose gel and cut with *Hae* III, *Alu* I, *Nco* I, *Hpa* II, and *Sau* 3AI. The resulting individual mixtures were subcloned in M13mp19(RF). DNA sequencing was carried out by the dideoxy chain-termination method of Sanger *et al.* (1977). Oligonucleotides were synthesized using an Applied Biosystems DNA synthesizer model 318A.

For Northern transfer analysis, RNA was electrophoresed in a 1.5% agarose gel containing 2.2 M glyoxal and transferred to GeneScreen membranes (NEN). Hybridization of RNA blots strictly followed the protocol of Khandjian (1986).

Bovine and human genomic DNA was prepared from blood anticoagulated with EDTA. Genomic DNAs from pig, sheep, hamster, mouse, and turkey were gifts from Dr. M. Zimmer. For Southern blot analysis, DNA was digested with restriction enzymes, electrophoresed in 0.7% agarose, and transferred to nitrocellulose membranes (BAS 85, Schleicher & Schuell) essentially as described

(Southern, 1975). The blotting procedure was used with initial depurination by HCl treatment. Hybridization was performed at 60°C under stringent conditions following the protocol of Church and Gilbert (1984). DNA probes were ³²P-labeled by random priming to a specific activity of 4 × 10⁸ cpm/μg employing a commercial kit (Amersham).

DNA and protein sequence analysis as well as sequence comparisons were performed using the computer program of the University of Wisconsin Genetics Computer group (Devereux *et al.*, 1984).

RESULTS

For detection of SAP-specific mRNA in poly(A)⁺RNA from bull seminal vesicle tissue, Northern hybridization was carried out employing synthetic DNA probes. The probes were designed on the basis of the known amino

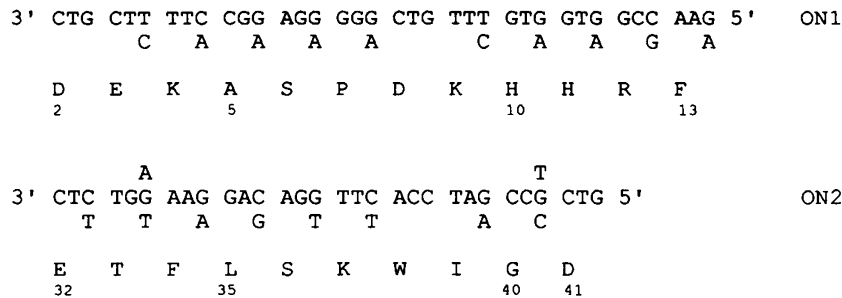


FIG. 1. Sequences of synthetic DNA hybridization probes. The ON1 sequence is derived from amino acids 2-13; the ON2 sequence derived from amino acid sequence 32-41. The DNA sequences are complementary to the corresponding SAP mRNA sequence.

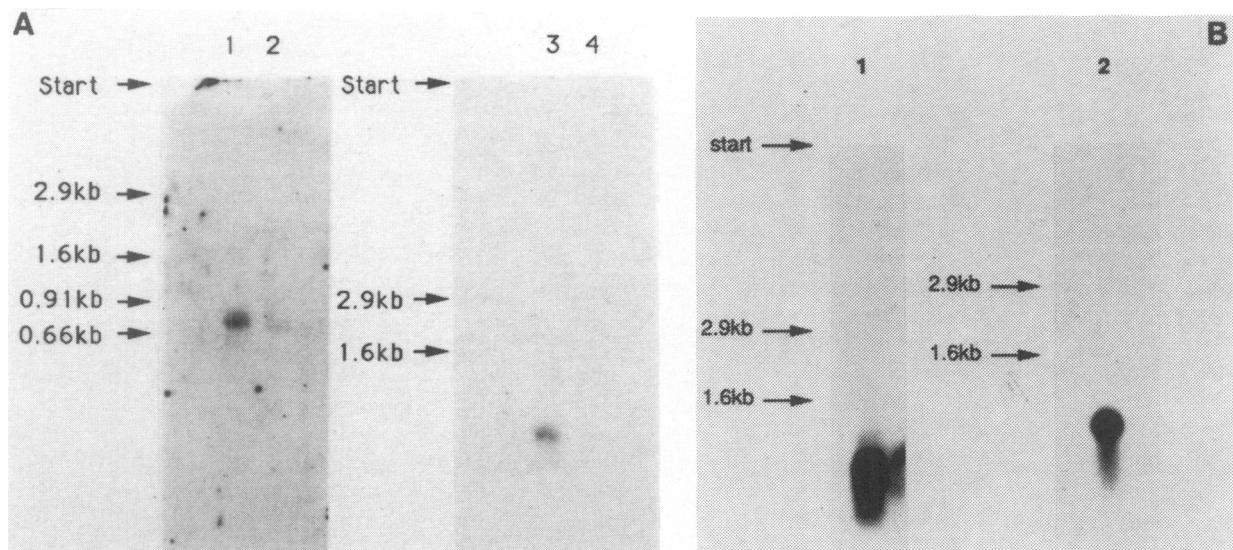


FIG. 2. Northern transfer analysis of bovine seminal vesicle and bovine testis poly(A)⁺RNA. A. Lane 1, 3 μg of poly(A)⁺RNA from seminal vesicle; lane 2, 10 μg of total seminal vesicle RNA; lane 3, 2 μg of poly(A)⁺RNA from seminal vesicle; lane 4, 3 μg of poly(A)⁺RNA from bovine testis. Hybridization was performed with ON2 (4.6 × 10⁸ cpm/μg). B. Lane 1, 3 μg of poly(A)⁺RNA from seminal vesicle hybridized with the *Pst* I insert of clone pBSV12; lane 2, poly(A)⁺RNA from seminal vesicle hybridized with the *Pst* I insert of pBSV16. The specific activity of hybridization probes was 3 × 10⁷ cpm/μg.

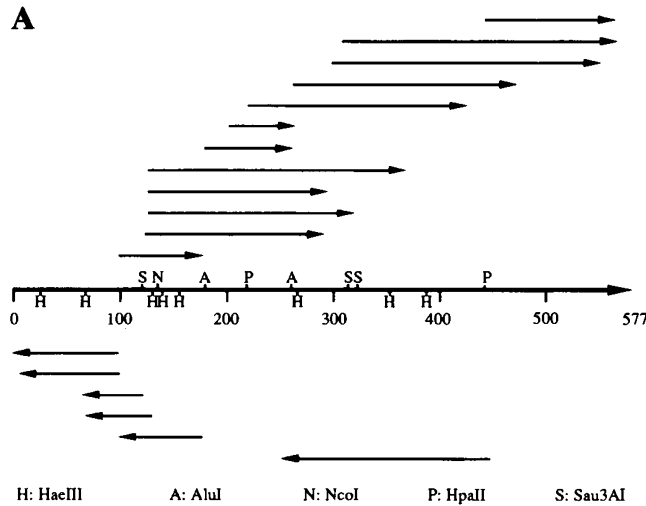


FIG. 3. cDNA sequence of pBSV12. A. Sequencing strategy and restriction map of pBSV12. Only restriction sites employed in sequencing are indicated. B. Nucleotide sequence of pBSV12 and deduced amino acid sequence of the SAP precursor. The amino acid sequence numbering starts with +1 for SAP. Putative signal sequences are underlined. The sequence regions corresponding to the hybridization probes ON1 and ON2 are also indicated.

B

```

1           G16-AGGAGGCCACAGCCCCACCTGTAGAAGTCCAGCCCTTCGGACTC 60
61 CTCTGGCCTTCCATCTCCTGACCACCTGCGCTCTAGAGTGTGCTATGATGGCGGGGCGC 120
           S V A M M A G R
           -32
121 AGATCGTGGCCCGCCATGGCCACAGTGTGCTGGCCCTGCTTGTCTGCCTGGGAGAGCTG 180
    R S W P A M A T V L L A L L V C L G E L
    -25 -22 -20 -15 -10
181 GTCGACTCCAAACCCCAGCCTTCAGATGAAAAAGCCTCACCGGATAAGCATCATCGTTTC 240
    V D S K P Q P S D E K A S P D K H H R F
    -5 -1 +1 |----- ON1 -----|
241 TCGCTGAGCCGCTATGCCAAGCTGGCCAATCGACTGCGAATCCTAAACTCTTAGAAACT 300
    S L S R Y A K L A N R L A N P K L L E T
    |-----
301 TTCCTCTCCAAAATGGATCGGTGATCGCGGGAATCGCTCAGTCAAGTAGCGCCAGAGGAG 360
    F L S K W I G D R G N R S V K *
    ----- ON2 -----|
361 CCTAACTGTGGTGAAGACCGCTACGGCCTCCAGAGAGAAATATGCCAAGACACCGACCCC 420
421 ATTTCCATGTTTGCCTGACCCCGGAACCTGGATTCTTCTCCAGGCTGGCTATGTCTGA 480
481 GGAGTGTGTCCCTTCCCTGGGCTGGAGGGCTGTAATGGTCCTTCCTTGTCCCTAAAATA 540
541 AAGTGCAAATTTCT-C23 577
    
```

acid sequence (Theil and Scheit, 1983; Lewis *et al.*, 1985; Sitaram *et al.*, 1986). The amino acid sequences extending from residues 31-40 (ON1) and 2-13 (ON2) of the SAP sequence were chosen. A 30-nucleotide-long DNA (ON1) as well as a 36-nucleotide-long DNA (ON2), complementary to the corresponding SAP mRNA, were synthesized (Fig. 1). The selection of codons for the synthetic DNA probes was based on the codon usage derived from bovine mRNAs (Aota *et al.*, 1988). A cDNA library derived from bull seminal vesicle poly(A)⁺DNA (6×10^4 recombinant clones) was screened by colony hybridization employing the radioactively labeled probe ON2. Hybridization at 60°C, following the protocol of Hanahan and Meselson (1983), furnished 64 positive clones with cDNA inserts of

lengths between 400 and 650 bp. Plasmid DNA from the purified positive clones was isolated and subjected to hybridization with the synthetic probes ON1 and ON2. Plasmid DNAs derived from clones carrying the complete coding region should hybridize to both ON1 and ON2, thus yielding higher signal intensities. Indeed, we observed three categories of hybridization intensities: low, medium, and high. Northern analysis of seminal vesicle poly(A)⁺RNA carried out with the synthetic probe ON2 using standard conditions at 42°C in 50% formamide, indicated the presence of a SAP-specific mRNA species of approximately 700 bp (Fig. 2); experiments with inserts of cDNA clones pBSV12 and pBSV16 led to similar results.

cDNA clone pBSV12, belonging to the strongly hybrid-

izing group and carrying the longest *Pst* I insert of ~600 bp, was sequenced. The sequencing strategy of clone pBSV12 is outlined in Fig. 3A. The nucleotide sequence obtained contains an open reading frame extending from nucleotide 97 to nucleotide 345 (Fig. 3B). The amino acid sequence derived from this open reading frame comprises the complete sequence of SAP. Surprisingly, the known amino acid sequence of SAP, which terminates with valine, contains an additional carboxy-terminal lysine residue in the derived amino acid sequence. Upstream of the amino-terminal serine residue of SAP, the derived amino acid sequence possesses 35 amino acid residues. Two methionine residues located at positions -22 and -32 might be translational starts of putative SAP precursor species. The corresponding methionine codons are part of the sequences GCCATGG and ATGATGG, respectively, which fulfill most of the criteria for a consensus sequence controlling translating efficiency of mammalian mRNAs (Kozak, 1984a,b, 1987; Lütke *et al.*, 1987). Although the cDNA sequence of pBSV12 contains the polyadenylation signal AATAAA, starting at nucleotide 539 of the sequence, a poly(A)⁺ tail is missing.

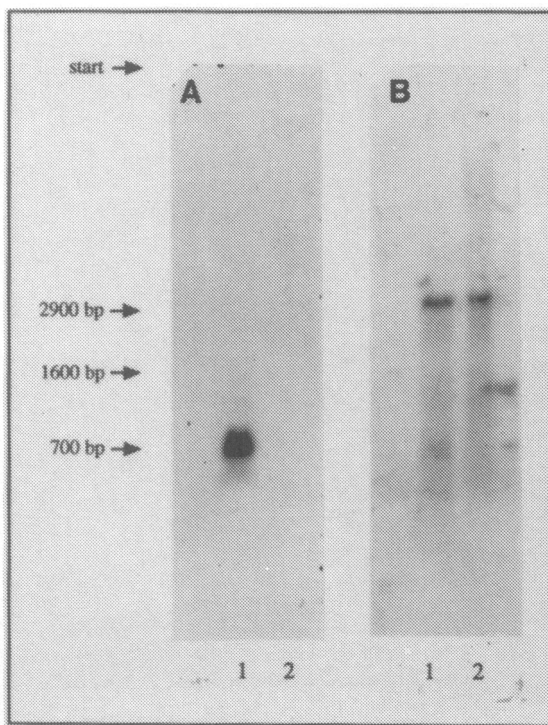


FIG. 4. Northern analysis of total RNA from mature bull and calf seminal vesicles. A. Lane 1, 10 μ g of total RNA from seminal vesicle of a mature bull; lane 2, total RNA from seminal vesicle of a male calf (2 weeks old). The insert of cDNA clone pBSV12 (sp. act., 2.7×10^8 cpm/ μ g) was employed as hybridization probe. B. Lane 1, 10 μ g of total RNA from seminal vesicle from a mature bull; lane 2, 10 μ g of total RNA of seminal vesicle of male calf. The cDNA insert of the human elongation factor 2 clone pHGR81 was used for hybridization (sp. act., 3.4×10^8 cpm/ μ g).

A comparative Northern analysis employing total RNA from seminal vesicle tissue of mature bull and of a male calf (2 weeks old) was performed. Only total RNA from mature bull gives the expected signal for the SAP-specific mRNA of 700 bp (Fig. 4). As a control, both RNA preparations were hybridized with a DNA fragment comprising the coding region for the human elongation factor 2 (Rapp *et al.*, 1989). Both RNA species reacted positively, indicating the presence of bovine elongation factor 2 mRNA (~3,500 bp) (Fig. 4). Hence, expression of the SAP gene may be androgen and/or developmentally controlled.

Male and female bovine genomic DNA was analyzed by Southern hybridization. After digestion with *Eco* RI, *Hind* III, *Pst* I, and *Pvu* II, enzymes known not to cut within the coding region of the cDNA insert of pBSV12, hybridization yielded single radiolabeled bands with each digest. Neither differences in the size of restriction fragments nor in the intensities of hybridization were seen between the fe-

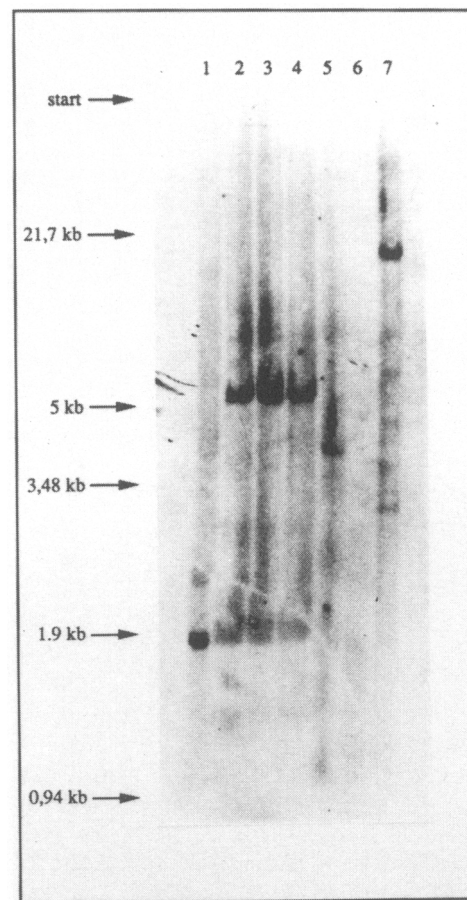


FIG. 5. Southern blot analysis of the gene for SAP. Bovine genomic DNA was digested with *Taq* I (lane 1, male individual A), *Eco* RI (lane 2), *Eco* RI (lane 3, male individual B), *Eco* RI (lane 4, female), *Pst* I (lane 5, female), *Hind* III (lane 6, female), and *Pvu* II (lane 7, female). A 230-bp fragment comprising the coding region was prepared by PCR and employed as hybridization probe (sp. act., 5×10^8 cpm/ μ g). Hybridization was performed under stringent conditions.

male and the male animals (Fig. 5). In the case of *Taq* I, which cuts twice within the coding region of the cDNA insert of pBSV12, two hybridization signals of similar size in the 1.9-kb range were detected by Southern analysis (Fig. 5). We tentatively infer from these experiments that there is one gene for SAP per haploid bovine genome.

Employing Southern hybridization, we attempted to detect SAP-related sequences in other species. Using a DNA fragment comprising the coding sequence of the cDNA insert of pBSV25 as a hybridization probe, no signals were found in genomic DNA of mouse, turkey, Chinese hamster, and monkey. However, besides a strong signal for bovine, significant hybridization signals were obtained for ovine and porcine genomic DNAs, respectively (Fig. 6).

DISCUSSION

The open reading frame of the cDNA insert extending from position 97 to 345 contains methionine codons at nucleotide positions 109 and 136 which are located within sequences that fulfill criteria of translational consensus

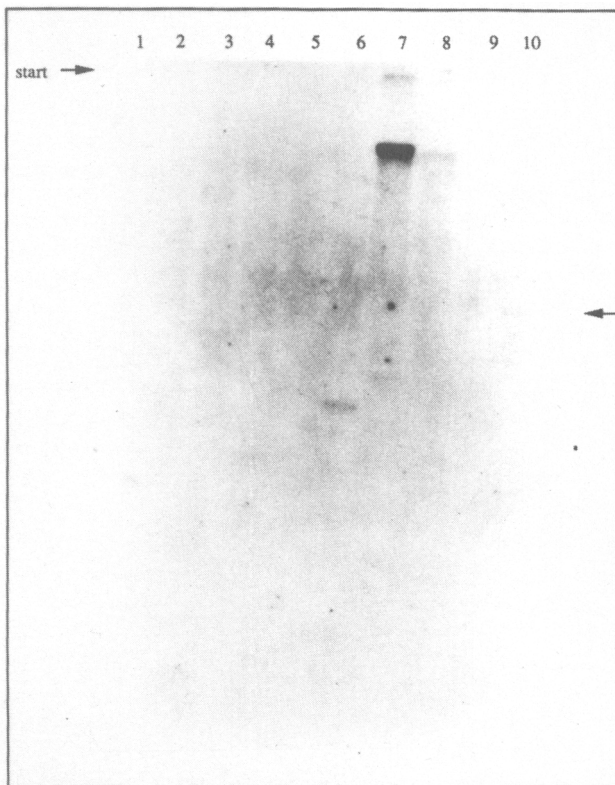


FIG. 6. Comparative Southern blot analysis. Genomic DNA (10 μ g each) was digested with *Eco* RI: lane 1, λ -DNA; lane 2, turkey; lane 3, mouse; lane 4, Chinese hamster; lane 5, Syrian hamster; lane 6, sheep; lane 7, bovine; lane 8, porcine; lane 9, monkey; and lane 10, human. The hybridization probe employed was the same as in Fig. 5. Hybridization was performed under stringent conditions. The arrow indicates a weak signal in lane 10.

sequences (Kozak, 1984a,b, 1987; Lütke *et al.*, 1987). According to Kozak (1984b) in most mammalian mRNAs the first A/G XX ATG of a long open reading frame is used as translational start. This would imply that the putative signal peptide for SAP comprises the amino-terminal extension of amino acid residues up to methionine -32 of the amino acid sequence; the corresponding M_r of the putative precursor would be 8,845 daltons. The second methionine codon within a consensus sequence at nucleotide position 136 leads to an amino-terminal extension of 22 amino acid residues. Hydrophobicity analysis of both amino-terminal extensions employing the algorithm of Kyte-Doolittle (data not shown) clearly favors the peptide fragment up to methionine -22 as a hydrophobic signal sequence, according to Austen (1979). However the latter SAP precursor possesses a M_r of 7,832 daltons, which does not agree with the size of the precursor ($M_r \sim 10$ kD) as determined by cell-free translation (Wempe *et al.*, 1990). However, SAP itself displayed an unusually high apparent M_r value of ~ 8 kD by NaDodSO₄ polyacrylamide gel electrophoresis compared to the true M_r of 5,470 daltons.

The amino acid immediately preceding the mature SAP peptide is a proline residue. This situation is rather uncommon for the linkage between signal peptides and the mature secreted proteins. However, a similar situation is encountered in the case of the precursor for elastin (Raju and Anwar, 1987).

The amino acid sequence derived from the cDNA is extended by an extra carboxy-terminal lysine residue. SAP, isolated from seminal vesicle secretion or bull seminal plasma, normally does not contain a carboxy-terminal lysine residue. Whether SAP prior to secretion undergoes both amino- as well as carboxy-terminal processing, thereby also losing the carboxy-terminal lysine extension, or whether the carboxy-terminal lysine is removed after secretion by specific carboxypeptidases present in seminal vesicles secretion has to await further investigation.

REFERENCES

- AOTA, S.-i., GOJOBORI, T., ISIBASHI, F., MARUYAMA, T., and IKEMURA, T. (1988). Codon usage tabulated from the GenBank Genetic Sequence Data. *Nucleic Acids Res.* **16**, 315-402.
- AUMÜLLER, G., and SCHEIT, K.H. (1987). Immunohistochemistry of secretory proteins in the bull seminal vesicle. *J. Anat.* **150**, 43-48.
- AUMÜLLER, G., VESPER, M., SEITZ, J., KEMME, M., and SCHEIT, K.H. (1988). Binding of a major secretory protein from bull seminal vesicles to bovine spermatozoa. *Cell Tissue Res.* **252**, 377-384.
- AUSTEN, B.M. (1979). Predicted secondary structures of amino-terminal extension sequences of secreted proteins. *FEBS Lett.* **103**, 308-313.
- CHURCH, G., and GILBERT, W. (1984). A genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**, 1991.
- CLEMENS, M.J. (1985). In *Transcription and Translation. A Practical Approach*. E.B. Hames and S.J. Higgins, eds. (IRL Press, Oxford) pp. 252-255.

- COMTE, M., MALNOE, A., and COX, J.A. (1986). Affinity purification of seminalplasmin and characterization of its interaction with calmodulin. *Biochem. J.* **240**, 567-573.
- DERWENSKUS, K.H., SPRINZL, M., and SCHEIT, K.H. (1989). Inhibition of cell proliferation by basic proteins from bull seminal plasma. *Biol. Chem. Hoppe-Seyler* **370**, 285-292.
- DEVEREUX, J., HAEBERLI, P., and SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the Vax. *Nucleic Acids Res.* **12**, 387-395.
- GUBLER, U., and HOFFMAN, B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **15**, 263-269.
- HANAHAH, D. (1983). Studies on transformation of *E. coli* with plasmids. *J. Mol. Biol.* **166**, 557-580.
- HANAHAH, D., and MESELSON, M. (1983). Plasmid screening at high colony density. *Methods Enzymol.* **100**, 333-343.
- KEMME, M., MADIRAJU, M.V.V.S., KRAUHS, E., ZIMMER, M., and SCHEIT, K.H. (1986). The major protein of bull seminal plasma is a secretory product of seminal vesicle. *Biochim. Biophys. Acta* **884**, 282-290.
- KEMME, M., and SCHEIT, K.H. (1988). Cloning and sequence analysis of a cDNA from seminal vesicle tissue encoding the precursor of the major protein of bull semen. *DNA* **7**, 595-599.
- KHANDJIAN, E.W. (1986). UV-crosslinking of RNA to nylon membrane enhances hybridization signals. *Mol. Biol. Rep.* **11**, 107-115.
- KOZAK, M. (1984a). Point mutations close to the aug initiator codon affect the efficiency of translation of rat preproinsulin *in vivo*. *Nature* **308**, 241-246.
- KOZAK, M. (1984b). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857-872.
- KRAUHS, E., PREUSS, K.D., and SCHEIT, K.H. (1990). Functional properties of peptides derived from seminalplasmin: Binding to monospecific anti-seminalplasmin IgGs and calmodulin. *Biol. Chem. Hoppe-Seyler* **371**, 111-116.
- LAEMMLI, U.K. (1970). Cleavage of structural proteins through the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- LEWIS, R.V., AUGUSTIN, J.S., KRUGGEL, W., and LARDY, H.A. (1985). The structure of caltrin, the calcium transport inhibitor of bovine seminal plasma. *Proc. Natl. Acad. Sci. USA* **82**, 6490-6491.
- LÜTKE, H.A., CHOW, K.C., MICKEL, F.S., MOSS, K.A., KERN, H.F., and SCHEELE, G.A. (1987). Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6**, 43-48.
- PALMIERI, M., CARSANA, A., FURIA, A., and LIBONATI, M. (1985). Sequence analysis of a cloned cDNA coding for bovine seminal ribonuclease. *Eur. J. Biochem.* **152**, 275-277.
- RAJU, K., and ANWAR, R.A. (1987). Primary structures of bovine elastin a, b, c deduced from the sequences of cDNA clones. *J. Biol. Chem.* **262**, 5755-5762.
- REDDY, E.S.P., and BHARGAVA, P.M. (1979). Seminalplasmin, an antimicrobial protein from bovine seminal plasma which acts in *E. coli* by specific inhibition of rRNA synthesis. *Nature* **279**, 725-728.
- RUFO, G.A., SINGH, J.P., BABCOCK, D.F., and LARDY, H.A. (1982). Purification and characterization of a calcium transport inhibitor protein from bovine seminal plasma. *J. Biol. Chem.* **257**, 4627-4632.
- SANGER, F., NICKLEIN, S., and COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
- SITARAM, N., KUMARI, V.K., and BHARGAVA, P.M. (1986). Seminal plasmin and caltrin are the same protein. *FEBS Lett.* **201**, 233-236.
- SCHEIT, K.H., ZIMMER, M., RAO, N.S., PRASAD, K.S.N., and BHARGAVA, P.M. (1986). Antibacterial activity of seminalplasmin, a basic protein from bovine seminal plasma. *Microbios* **45**, 105-112.
- SCHEIT, K.H. (1986). The major proteins of bull seminal plasma. *Biol. Chem. Hoppe-Seyler* **367**, 1173-1180.
- SOUTHERN, E. (1975). Detection of specific sequences among DNA fragments separated by gelelectrophoresis. *J. Mol. Biol.* **98**, 503-517.
- THEIL, R., and SCHEIT, K.H. (1983). Amino acid sequence of seminalplasmin, an antimicrobial protein from bull semen. *EMBO J.* **2**, 1159-1163.
- VILLA-KOMAROFF, L., EFSTRADIADIS, A., BROOME, S., LOMEDICO, P., TIZARD, R., NAKER, S.P., CHICK, W.L., and GILBERT, W. (1978). A bacterial clone synthesizing proinsulin. *Proc. Natl. Acad. Sci. USA* **75**, 3727-3731.
- WEMPE, F., WAGNER, S., v.d. KAMMER, H., KRAUHS, E., and SCHEIT, K.H. (1990). Seminalplasmin, the major basic protein of bull seminal plasma is a secretory protein of seminal vesicles. *Biochim. Biophys. Acta* (in press).

Address reprint requests to:

Dr. Karl Heinz Scheit
Max-Planck-Institut für biophysikalische Chemie
Göttingen, West Germany

Received for publication March 14, 1990, and in revised form May 11, 1990.