0013-7227/01/$03.00/0
Printed in U.S.A.

Endocrinology 142(10):4529–4539

Primate Epididymis-Specific Proteins: Characterization of ESC42, a Novel Protein Containing a Trefoil-Like Motif in Monkey and Human

QIANG LIU*, KATHERINE G. HAMIL, P. SIVASHANMUGAM, GAIL GROSSMAN, RAMA SOUNDARARAJAN†, A. JAGANNADHA RAO, RICHARD T. RICHARDSON, YONG-LIAN ZHANG, MICHAEL G. O'RAND, PETER PETRUSZ, FRANK S. FRENCH, AND SUSAN H. HALL

Departments of Pediatrics (Q.L., K.G.H., F.S.F., S.H.H.) and Cell and Developmental Biology (P.S., G.G., R.T.R., M.G.O., P.P.) and Laboratories for Reproductive Biology (Q.L., K.G.H., P.S., G.G., R.S., R.T.R., M.G.O., P.P., F.S.F., S.H.H.), University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599; State Key Laboratory of Molecular Biology (Q.L.,Y.-L.Z.), Institute of Biochemistry and Cell Biology, Shanghai Institute for Biological Science, Sinica Academia, Shanghai 200031, China; and Department of Biochemistry (A.J.R.), Indian Institute of Science, Bangalore, India 560012

Epididymal secreted proteins promote sperm maturation and fertilizing capacity by interacting with sperm during passage through the epididymis. Here we investigate the molecular basis of sperm maturation by isolating cDNA clones for novel epididymis-specific expressed sequences. Thirty-six novel cDNAs were isolated and sequenced from a subtracted *Macaca mulatta* epididymis library. The clones encode proteins with a range of motifs characteristic of protein-modifying enzymes, protease inhibitors, hydrophobic ligand-binding and transport proteins, extracellular matrix-interacting proteins, and transcription regulatory factors. The full length coding sequences were obtained for 11 clones representing a range of abundance levels. Expression of each is regionally localized and androgen regulated. The most abundant, ESC42, contains a cysteine-rich region similar to the signature binding domain of the trefoil family of motogenic wound repair proteins. The monkey and human proteins are nearly 90% identical. Immunohistochemical staining revealed that the protein is most abundant in the epithelium of the caput and is also present in the lumen and bound to sperm. The ESC42 gene, located on chromosome 20q11, contains two exons encoding two nearly identical predicted signal peptides and a third exon encoding the rest of the protein. (*Endocrinology* 142: 4529–4539, 2001)

SPERMATOZOA RELEASED FROM the seminiferous epithelium are carried in testicular fluid into the rete testis, where, although they appear highly developed morphologically, they lack forward motility and the ability to fertilize eggs (1, 2). Sperm proceed through the epididymal caput, corpus, and cauda, morphologically and biochemically distinct epididymal regions where they undergo sequential modifications (1–3) collectively known as sperm maturation. During the last 25 yr, analyses of regional differences in secreted epididymal proteins involved in sperm maturation have led to the discovery of a large number of epididymal proteins, primarily by three approaches. First, regionally secreted proteins in the epididymal lumen and extracted from the sperm membrane were analyzed by two-dimensional gel electrophoresis and immunodetection (4, 5). These methods yielded 146 epididymal proteins from adult boar epididymis (6) and 201 proteins from the stallion epididymal lumen (7), most of unknown function. Second, regionally localized expression of mRNAs for proteins of known function were analyzed. mRNAs for epididymal glutathione peroxidase and superoxide dismutase were detected primarily in caput and corpus, respectively, indicating that the need for antioxidant enzymes may vary along the length of the epididymal tubule (8). Third, subtractive hybridization cloning led to the identification of sequences expressed specifically in the human epididymis, including a disintegrin, a member of the metalloproteinase family of proteases (9), and a number of novel cDNAs (10). Nevertheless, the still fragmentary nature of our understanding of the molecular basis of sperm maturation suggests the involvement of numerous unknown proteins. The identification and definition of the structure, function, and regulation of epididymal sperm-modifying proteins are crucial to understanding male fertility.

Human epididymal studies are constrained by the impracticality of experimentation and by the advanced age of available tissue donors. To overcome these difficulties and to identify and investigate unknown epididymal proteins closely related to human proteins and involved in sperm function, we constructed and analyzed a rhesus monkey epididymis-specific cDNA library. The library contains 36 nonoverlapping epididymis-specific clones. Sequencing revealed encoded proteins homologous to various enzymes, protease inhibitors, and ligand-binding proteins, but more than half showed no relationship to any known proteins. The most abundant of these novel clones, epididymis-specific clone 42 (ESC42), was characterized by sequence and expression analyses. The protein contains a cysteine-rich region

Abbreviations: DAPI, 4′,6-Diamidino-2-phenylindole; pfu, plaque-forming units; TFF, trefoil factor.

similar to the signature motif of the trefoil family of motogenic proteins involved in wound healing.

Materials and Methods

Preparation of RNA from rhesus monkey tissues

Total RNA was purified by the method of Chirgwin *et al.* (11). *Macaca mulatta* (rhesus monkey) epididymis, testis, liver, and brain (Covance Research Products, Alice, TX) were pulverized in liquid nitrogen and homogenized in guanidine thiocyanate (Fluka Chemical Co., Milwaukee, WI). Debris were removed by centrifugation at 10,000 rpm at 10 C, and supernatants were layered over 5.7 M CsCl cushions (biochemical grade; Gallard and Schlesinger, Carl Place, NY) and centrifuged at 35,000 rpm at 25 C for 16–20 h in a Beckman Coulter, Inc. (Palo Alto, CA) SW41 rotor. Polyadenylated [poly(A)$^+$] mRNA was purified from total RNA by standard methods (12) using oligo(dT) cellulose type 2 (Becton Dickinson and Co., Collaborative Biochemical Products, Bedford, MA).

Preparation of subtracted rhesus monkey epididymis cDNA library

The subtracted cDNA library was prepared using the PCR-Select cDNA Subtraction kit (CLONTECH Laboratories, Inc., Palo Alto, CA) according to the manufacturer's recommendations. Double stranded cDNAs were synthesized from poly(A)$^+$ mRNA from rhesus monkey epididymis (tester), rhesus monkey testis (driver), and human skeletal muscle (control) using avian myeloblastosis virus reverse transcriptase for the first strand and T4 DNA polymerase for the second strand. *Rsa*I-digested epididymal cDNA (tester) was divided into two portions. Each portion was ligated to a different adaptor, denatured, and hybridized to excess testis cDNA (driver). The remaining population of single stranded cDNAs was enriched for epididymis-specific expression with equalized representation of abundant and rare sequences. The two portions of epididymal cDNA were hybridized to each other. Duplexes containing both adapters were amplified by PCR and cloned into pGEM-T Easy vector (Promega Corp., Madison, WI). Blue/white screening in XL-1 Blue cells revealed that the library contains 667 insert-containing clones.

The insert of each of the 667 clones was amplified by PCR using 1 μl of each bacterial culture as a template. PCR products (3 μl each) were arranged in arrays on Zeta-probe membranes (Bio-Rad Laboratories, Inc., Hercules, CA) using a Bio-Dot SF blotting apparatus (Bio-Rad Laboratories, Inc.) according to the manufacturer's protocol. On each blot, 3 μl of glycerol 3-phosphate dehydrogenase PCR product was loaded into one slot as a control. The cDNA probes were prepared using rhesus monkey epididymis, testis, brain, and liver poly(A)$^+$ RNA, [α-^{32}P]dCTP, and SuperScript II reverse transcriptase (Life Technologies, Inc., Rockville, MD). The 511 clones that tested positive for epididymis expression and negative for testis were hybridized to monkey epididymis, liver, and brain cDNAs for a second cycle of screening. Clones hybridizing to testis, brain, or liver cDNAs were eliminated, leaving a total of 424 clones.

Sequencing

To minimize repeated sequencing of identical clones, a hybridization sequencing strategy was used. Initially, 10 clones with different insert lengths were sequenced. Slot blots containing the 424 epididymis-specific clones were hybridized with a combined probe containing these 10 clones. Clones that hybridized to the combined probe were thus duplicates and were eliminated from further sequencing. From among those clones that failed to hybridize to the original 10, a second set of 10 clones with different insert lengths was sequenced and hybridized to the remaining unknown clones. This screening procedure was repeated until all of the different clones were sequenced.

Miniprep DNA was prepared by standard alkaline lysis of 5-ml overnight culture of each clone. Plasmid was further purified with ribonuclease A digestion and precipitation with PEG 8000. Plasmid DNA was subjected to sequencing at the University of North Carolina, Chapel Hill, Automated DNA Sequencing Facility using an ABI PRISM model 377 DNA sequencer (PE Applied Biosystems, Foster City, CA) and the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq Æ DNA Polymerase FS. Primers were synthesized on an automated PE Applied Biosystems DNA synthesizer (model 394) using standard cyanoethyl phosphoranidite chemistry.

Construction of rhesus monkey nonsubtracted epididymis cDNA library

The library was prepared using the ZAP cDNA synthesis kit and the ZAP cDNA Gigapack II packaging kit (Stratagene, La Jolla, CA) according to the manufacturer's recommendations. cDNAs greater than 500 bp in length were inserted into UniZap XR and transformed into XL1-Blue MRF′. The primary library contained 2.0×10^6 plaque-forming units (pfu) with an average insert size of 1.6 kb. The library was amplified once by standard methods to a titer of 3.5×10^9 pfu/μl.

Screening of rhesus monkey epididymis cDNA library for full length clones

Duplicate 0.45-μm NitroPure (Osmonics, Inc., Minnetonka, MN) nitrocellulose transfer membranes were placed on agarose plates containing 5×10^4 plaques to allow the transfer of the phage particles and hybridized to radiolabeled subtracted library inserts. Positive clones were purified and confirmed by insert amplification by PCR with universal primer M13R or M13F and a gene-specific primer based on the sequence of each of the different subtracted clones. *In vivo* excision of recombinant pBluescript plasmids used host cells XL1-Blue MRF′ and SOLR cells (Stratagene) and ExAssist helper phage (1×10^6 pfu/μl). Well isolated single clones were picked for overnight Luria-Bertani culture and plasmid prepared for sequencing.

Northern blot hybridization

Northern hybridizations were performed as described previously (13). Briefly, 10 μg of each RNA sample was glyoxalated at 50 C for 1 h. The RNA samples were separated on 1% agarose gels in 10 mM sodium phosphate (pH 7.0) at 6 V/cm with buffer recirculation. The RNA was transferred to nylon membrane (Biotrans, neutral; ICN, Costa Mesa, CA) by capillary action in 10× standard saline citrate overnight and fixed by UV treatment on the Autocrosslink setting in the StrataLinker UV crosslinker (120,000 μJ of UV energy decreasing to 0 during 30 sec) (Stratagene). RNAs were stained and evaluated for equality of loading and for degradation by soaking membranes in 5% acetic acid for 5 min, covering with acidic methylene blue dye (0.4% methylene blue, 0.5 M sodium acetate, pH 5.2) for 5 min, and rinsing with water (14). Positions of 18S, 28S, and RNA molecular mass standards were marked with a pencil, and the membrane was allowed to air dry.

Peptide synthesis

A monkey ESC42 C-terminal peptide (QGTQTSPPNVHHTC) was synthesized using a Rainin Instrument Co., Inc. (Woburn, MA) multiple peptide synthesizer using fluoroenylmethyloxycarbonyl chemistry in the University of North Carolina Program in Molecular Biology Protein Chemistry Facility. The peptide was purified by HPLC and conjugated to keyhole limpet hemocyanin. The C-terminal cysteine was added to facilitate coupling. Antibodies were raised in rabbits 5497 and 5498 at Bethyl Laboratories, Inc. (Montgomery, TX). An affinity column was prepared by attaching 2 mg of this antigen peptide to SulfoLink gel (Pierce Chemical Co., Rockford, IL). Ten milliliters of antiserum was passed over the column, and bound antibody was eluted in low pH according to the column manufacturer's recommendations.

Tissue sources

Human epididymides for immunohistochemistry and Northern blot analyses were obtained from prostate cancer patients ranging in age from 58 to 83 yr. The epididymides were trimmed of fat and connective tissue and dissected into caput, corpus, and caudal regions before freezing or fixation.

For analysis of androgen regulation, male rhesus monkeys of similar age, weight, and testicular size underwent subcapsular orchiectomy (15) or sham operation. One orchiectomized monkey was immediately in-

jected im with T enanthate 30 mg/kg body weight (400 mg total), and the other was injected with vehicle. Epididymides and remaining testes were removed 6 d later and frozen in liquid nitrogen. Serum samples for T RIA were taken just before surgery on d 0 and 6. All animals used in these studies were maintained in accordance with the NIH Guide for the Care and Use of Laboratory Animals. The protocol follows accepted veterinary medical practice and was approved by the University of North Carolina Animal Care and Use Committee. The animals were given analgesics and were monitored closely after surgery.

Rhesus monkeys 10–12 yr of age with proven breeding history (Covance Research Products and Dr. Catherine VandeVoort, CA Regional Primate Center, Davis, CA) provided tissues for Northern analysis and immunostaining. Tissues for immunohistochemistry were fixed in Bouin's solution (75 ml of saturated picric acid, 5 ml of glacial acetic acid, 25 ml of 37% formaldehyde) promptly after excision. Surplus human testes and epididymides were made available by Dr. James L. Mohler (Department of Urology/Surgery, University of North Carolina, Chapel Hill). Other human tissues were obtained from the Tissue Procurement Core Facility of the Lineberger Comprehensive Cancer Center (University of North Carolina, Chapel Hill). Human tissues are not accompanied by identifying information and cannot be traced to the donor.

Subcloning cDNAs into pSG5

The complete monkey cDNA was excised from pBluescript by *Eco*RI/*Rsa*I digestion and cloned into *Eco*RI/*Bam*HI-digested pSG5 (Stratagene) after blunting the *Bam*HI site. The cDNA without the signal peptide was obtained by PCR amplification from the pBluescript clone using the high fidelity Pfu polymerase (Stratagene) and the oligonucleotide primers CCAAGGAATTCCAACCATGGGTGG (forward) and ACAGGGATCCGTGACATTCGAGAAGAAG (reverse). The amplified product was digested with *Eco*RI/*Bam*HI, gel purified, and cloned into the *Eco*RI/*Bam*HI site of pSG5.

Transfection of COS cells for immunostaining and Western blotting

Monkey kidney COS-1 cells (American Type Culture Collection, Rockville, MD) were cultured in DMEM with high glucose. Cells were plated in 10-cm dishes (10^6 cells/dish) for protein extraction for Western blotting. Cells were transfected with 10 μg of pSG5-ESC42 expression vector per 10-cm dish using DEAE-Dextran (Sigma Inc., St. Louis, MO) (16). Cells were washed once in 6 ml of PBS, harvested in 1 ml of PBS, and centrifuged for 1 min at 5,000 $\times$ *g*. Cells were resuspended in 100 μl of RIPA buffer (PBS, pH 7.4, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 0.5 mM phenylmethylsulfonyl fluoride, 10 μM pepstatin, 4 μM aprotinin, 80 μM leupeptin, and 5 mM benzamidine) (17). DNA was sheared through a 27-gauge needle. Lysates were centrifuged at 13,000 $\times$ *g* for 2 min to remove debris. Aliquots of 0.5 μl of COS lysate were denatured in 1% SDS loading buffer loaded on 12% polyacrylamide gels. Surplus swim-up sperm pellets were boiled in 200 μl of 1% SDS loading buffer for 5 min, and 5 μl was applied per lane. Proteins were electroblotted to nitrocellulose membranes and immunodetected using enhanced chemiluminescence (NEN Life Science Products, Boston, MA) or SuperSignal (Pierce Chemical Co., Inc., Rockford, IL).

Attempted amplification of alternative transcripts by PCR

Using a human caput/corpus cDNA library in Lambda Zap (Stratagene) (18) as template, amplification of alternative transcripts was attempted using a single reverse primer (AGTGGGATCCGAGAGGAAGTCATGAGC) and three forward primers, the first to the 5′ region (GCGAATTCCCAAAGGCAAGTCTAAATGTTG), the second to the first alternative exon (GCGAATTCTACCACCTCCTGCTTCCCAAG), and the third to the second alternative exon (CGGAATTCTCTCCTGTTTCCCAGGGAC).

Immunohistochemical staining

Tissues were fixed by immersion in Bouin's fluid and embedded in paraffin according to standard protocols (19). For immunohistochemical staining of rhesus monkey epididymis, the Tyramide Signal Amplification Indirect kit (NEN Life Science Products) (20, 21) was used with diaminobenzidine (Aldrich, Milwaukee, WI) as the chromogen. To visualize ESC42 for confocal microscopy, the fluorophor cyanine-3 conjugated to tyramide was used with the Tyramide Signal Amplification Direct kit (NEN Life Science Products). The DNA in nuclei was stained using blue fluorescent 4′,6-diamidino-2-phenylindole (DAPI) in Vectashield mounting medium (Vector Laboratories, Inc., Burlingame, CA). Photographs of diaminobenzidine-stained sections were taken with a Nikon (Tokyo, Japan) Eclipse E600 microscope using a Spot digital camera and Spot Advanced software (Diagnostic Instruments, Inc., Sterling Heights, MI). Digital photographs of fluorescent sections were taken using a Carl Zeiss (Jena, Germany) 410 laser scanning confocal microscope and Carl Zeiss software.

Surplus swim-up human sperm were provided by Dr. Stan Beyler (Assisted Reproductive Technology Clinic, Department of Obstetrics and Gynecology, University of North Carolina, Chapel Hill). Sperm were immunostained as described previously (18) using affinity-purified antibody 5497. Sperm images were arranged and labeled using Adobe Photoshop 5.0 (Adobe Systems, Mountain View, CA).

Results

New epididymis-specific clones

From a rhesus monkey epididymis-specific cDNA library, 36 nonoverlapping novel cDNAs (Table 1) were isolated and sequenced. In addition, monkey homologs of 7 known epididymis-specific cDNAs encoding AEG (32), EAP-I, GPX-5, HE1, HE3, HE5, and HE6 (33) were found, consistent with expected library composition. Full length cDNA sequences and additional sequence from the human genomic database revealed that a total of 25 independent expressed sequences were isolated. The predicted proteins show motifs characteristic of enzymes, protease inhibitors, ligand transport, extracellular matrix interaction, and nuclear regulatory proteins. The motifs predicted in ESC42, ESC342, and ESC328 suggest extracellular matrix adhesion, a protein group expanded in the human genome compared with other eukaryotes (34).

By contrast, the nuclear regulatory factors, also expanded in the human genome, are not well represented in this subtracted library. They would be eliminated during subtraction procedures if these functions are performed by identical proteins in other organs. Twelve percent or 48% of the 25 cDNAs either contain no satisfactory open reading frame or predict amino acid sequences unrelated to known proteins and containing no recognizable functional motifs. This result is similar to the 41.7% molecular function unknown category in the 26,383 human genes (34).

To obtain full-length cDNAs, a nonsubtracted monkey epididymis cDNA library was constructed and screened by plaque hybridization with subtracted library cDNAs. The 11 full length cDNAs currently under investigation are summarized in Table 2. They encode proteins with less than 50% amino acid sequence identity to the most closely related known protein except for the recently published human and dog homologs of ESC342 and a mouse lipocalin related to ESC384. These low levels of homology leave open the possibility that these proteins in the epididymis serve different functions from the related proteins. Expression of each gene in monkey was regionally concentrated and at least partially regulated by androgen, as determined by Northern blot hybridization to RNAs isolated from intact, castrated, and T-replaced rhesus mon-

TABLE 1. Profile of epididymis-specific clones

Epididymal subtracted clone	Relative abundance	Chromosomal location	Homologies
ESC300, ESC78	+	1q31	Zinc protease motif
ESC177	+	3q24	Phosphatase
ESC6	++	5q33.2	Serine protease inhibitor
ESC384	++	9q34.2-34.3	Lipocalin
ESC513, ESC458, ESC63	+++	9q34.2-34.3	Lipocalin
ESC476	+	11p14.2	–[a]
ESC112, ESC115	++	11q24	Zinc finger motif
ESC485, ESC211	+	11q24	–
ESC461, ESC374	+	14p11.2	Ribonuclease A
ESC77	+	14p11.2	–
ESC615	++	16q12, 22q11.2	Carboxylesterase
ESC9, ESC702	+++	19q13.4	Vitamin D-binding motif
ESC342, ESC46	+	19q13.3	Fibronectin II domains
ESC396, ESC88	++	20p12.2-13	–
ESC13, ESC278	++	20p11.21-12.3	Cysteine protease inhibitor
ESC42	+++	20q11.1-11.22	Trefoil motif
ESC54, ESC548	+	20q11.2-12	–
ESC376	+	20q12	–
ESC328	++	–	von Willebrand motif
ESC7	++	–	–
ESC163	+	–	–
ESC363	+	–	–
ESC462	++	–	–
ESC468	+	–	–
ESC507	+	–	–

[a] *Dashes* indicate information is not yet available in GenBank.

TABLE 2. Characterization of full length clones

Clone	ORF (aa)	Epididymis expression[a]			Androgen regulation	Homologous sequences	Identity[b]
		Caput	Corpus	Cauda			
ESC6	94	−	+	++	Yes	Acrosin-trypsin kazal-type inhibitor (22)	34% 63-aa human
ESC9	146	+	+++	+	Testis factor	Vitamin D-binding protein (23)	24% 79-aa rabbit
ESC13	138	+++	++	−	Yes	Cystatin-related CRES (24)	38% 102-aa human
ESC42	123	+++	++	−	Yes	Trefoil-like motif (25), EGF motif	28% 47-aa mouse
ESC112	98	+	+++	++	Yes	SP10 intraacrosomal protein (26)	37% 77-aa human
ESC177	433	+	++	++	Yes	Prostatic acid phosphatase (27)	35% 70-aa rat
ESC342	233	+	+++	+/−	Yes	Fibronectin II domains (28)	98% 223-aa human
ESC384	182	+++	+	−	Testis factor[c]	Retinoic acid-binding protein mEP17 (29)	66% 174-aa mouse
ESC461	204	+/−	+++	++	Yes	Secreted pancreatic ribonuclease A2 (30)	30% 100-aa human
ESC513	168	+++	++	−	Yes	Major urinary protein (31)	38% 156-aa mouse
ESC615	551	+/−	+	++	Yes	Carboxylesterase	47% 501-aa rat

aa, Amino acids; ORF, open reading frame.

[b] +++ indicates highly abundant expression in monkey; − indicates undetected expression.

[b] Identity indicates percent identical amino acids over the indicated length and species found in a BLAST search.

[c] Testis factor indicates mRNA levels lower in castrated than in sham-operated animals but not fully maintained in T-replaced animals. Androgen regulates, but other factor(s) are required for full expression.

keys as described below for ESC42 (data not shown for other clones).

Clone ESC42

ESC42 is the most highly expressed of the 11 full length clones. Of 4×10^4 plaques screened from the nonsubtracted epididymis cDNA library, 59 or 0.15% hybridized to the ESC42 cDNA. The monkey mRNA is 94% identical to the human cDNA, and the open reading frames encode 89% identical proteins (Fig. 1). Signal peptides are predicted for both proteins with cleavage sites between monkey amino acids 22 and 23 and between human amino acids 20 and 21

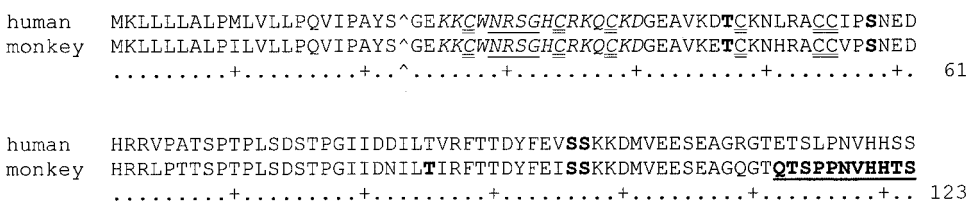

FIG. 1. Alignment of human and rhesus monkey ESC42 amino acid sequences. The predicted functional sites are represented as follows. *Carets* indicate the signal peptide cleavage site, and *stars* indicate the trefoil-like motif. The zinc finger-like motif is *italicized*. The *N*-glycosylation site is shown with a *single underline*. The six cysteines are shown with *double underlines*. Phosphorylation sites are shown in *boldface*. The C-terminal peptide used to raise the antibody for immunostaining is indicated in *boldface* and *underlined*. The GenBank accession numbers for these sequences are AF347073 (human) and AF207834 (monkey).

(35). The predicted proteins have a calculated molecular mass of 11.2 kDa after signal peptide removal. A PROSITE (36) scan identified several sites for posttranslational processing, including an asparagine glycosylation site (**N**RSG, amino acids 29–32), four PKC phosphorylation sites (**T**CK, amino acids 47–49; TIR, amino acids 86–88; **S**SK, amino acids 97–99; and **S**KK, amino acids 98–100), and casein kinase C phosphorylation sites (**S**NED, amino acids 58–61; and **S**KKD, amino acids 98–101). The predicted human and monkey proteins are novel and the functions are unknown, but database searches reveal homologies to functional motifs in the cysteine-rich N-terminal half of the mature protein. A motif similar to the CCHC-type zinc finger found in DNA-binding proteins was identified in amino acids 25–40 by Profilescan (37). Similarities to the trefoil three-loop domain (25) and to epidermal growth factor were recognized in amino acids 35–56 by GeneFIND (38). The cysteine-rich N-terminal region is also 49% similar across 68 amino acids to the extracellular N terminus of low affinity nerve growth factor receptor (p75) (39). This region is the first of four similar cysteine-rich motifs forming the extracellular nerve growth factor binding domain of the receptor. In addition, the N terminus of ESC42 is 48% similar to a 57-amino acid segment near the C terminus of HE2β1 (18). These relationships are consistent with the possibility that disulfide bonds linking the cysteines in the N-terminal half of the ESC42 form functional intrachain loop structures with the potential to mediate macromolecular interactions.

The human ESC42 gene is located on chromosome 20q11 and includes two predicted alternative first exons of 84 bp each containing a 5′ untranslated region and encoding 19 of the 20 amino acids of the predicted signal peptide (Fig. 2). The two first exons are nearly identical except for the splice junctions and five base substitutions, two of which are silent; one leads to amino acid L13 conversion to F, one leads to V17 conversion to G, and one leads to P19 conversion to L. These changes might alter the structure of the signal peptide but would not change the cleavage site (35). Alternative splicing of the mRNA could result in a coding sequence for either of the two exon 1 signal peptides or none. However, attempts to amplify alternative transcripts by PCR yielded only products containing the first of the two signal peptide exons (data not shown). A putative TATA box near the cDNA 5′ end and the stop codon, polyadenylation signal, and poly(A) tail (tail removed in Fig. 2) in the cDNA suggest that the complete sequence was obtained. The gene is not associated with the cluster of classic trefoil factors (TFFs) on chromosome 21p22.3 (40).

ESC42 protein extracted from ejaculated human sperm migrated with an apparent mol wt of about 20,000, somewhat larger than the 11.2 kDa predicted for the mature peptide and consistent with posttranslational processing such as phosphorylation or glycosylation (Fig. 3). COS-1 monkey kidney cells were transiently transfected with plasmids to overexpress monkey ESC42 with or without the signal peptide. The recombinant proteins appeared smaller than the protein derived from sperm but migrated close to the 18.5-kDa molecular mass marker on this 12% SDS polyacrylamide gel as well as on 10–20% gradient polyacrylamide gels (data not shown). Immunoreactive recombinant species of relative mol wt 32,000, 35,000, and 50,000 may represent dimers and trimers because trefoil domains are known to form intermolecular multimers through disulfide linkage. Attempts to eliminate these species by additional reducing agents succeeded in decreasing the quantity of larger species and increasing the quantity of smaller species (data not shown).

Immunostaining using rabbit antibody raised against the 14-amino acid C-terminal peptide demonstrated expression of ESC42 in epithelial cells throughout the monkey epididymis, but most abundantly in caput. The ESC42 protein in caput (Fig. 4A) was concentrated in specific cells of the columnar epithelium and absent from adjacent, morphologically identical cells, forming a banded pattern characteristic of epididymal gene expression (18). Immunostaining was eliminated by preincubation of antibody with peptide antigen (Fig. 4B). The ESC42 protein in the efferent duct epithelial cells was located in some cells in a rounded pattern suggesting location in nuclei (Fig. 4C). However, confocal microscopy using red fluorescent ESC42 immunostaining and DAPI nuclear stain (Fig. 4, E and F) demonstrated that ESC42 protein was not present in nuclei. In caput, ESC42 was located in the perinuclear regions (Fig. 4, A, D, and F). ESC42 was concentrated at the luminal surface of the epithelium and associated with the microvilli (visible in Fig. 4D and especially clear in Fig. 4F). Staining was also associated with sperm in the lumen. Human ejaculated sperm were immunostained using Texas Red fluorescent dye-conjugated secondary antibody (Fig. 5A). ESC42 was concentrated on the sperm in the postacrosomal head and neck regions. The protein was less abundant on the midpiece and tail.

ESC42 mRNA was highly restricted to the epididymis in the male (Fig. 6). A weakly hybridizing species of the same

```
tttcctgcttttctgtagcaagagcgaaaaaaggtgtttctcaagttcttcagtttagaa  60
 \\                                                          //  //
ctctgcccccactgcagagaggccaataaaactcttgttcagactcacactgcacacagt  1800 TATA box

attctgaactcctggatCTACCACCTCCTGCTTCCCAAGGACCATGAAACTCCTGCTGCT  1860
                                          MetLysLeuLeuLeuLe
                                          M  K  L  L  L  L
                                                              EXON 1
GGCTCTTCCTATGCTTGTGCTCCTACCCCAAGTGATCCCAGgtaatcagaggtcagggaa  1920
uAlaLeuProMetLeuValLeuLeuProGlnValIleProA
 A  L  P  M  L  V  L  L  P  Q  V  I  P  A

gatacaaaaatagaggtatagtgggttctggctcatgaaaattccaatgtgtcacggtac  1980
 \\                                                          //  //
tagggaaacagGGAAACCTCTCCTGTTTCCCAGGGACCATGAAACTCCTGCTGCTGGCTC  3360
                                       MetLysLeuLeuLeuLeuAlaL
                                       M  K  L  L  L  L  A  L
                                                          alternative EXON 1
TTCCTATGCTTGTGTTCCTACCCCAAGGGATTCTGGgtaaccagaagtcagtaatcagaa  3420
euProMetLeuValPheLeuProGlnGlyIleLeuA
  P  M  L  V  F  L  P  Q  G  I  L  A

gttgtgtatcttaaaggacactgccagcagaaattgccaactctgggggtttttagtaat  3480
 \\                                                          //  //
tgaaacagagtctcgctctgttgcccaggctggtgtgcaatgagagcatcagtaacccaa  4020

aatcaatggtctttccttttattattcagCCTATAGTGGTGAAAAAAAATGCTGGAACAG  4080
                              laTyrSerGlyGluLysLysCysTrpAsnAr
                               ^Y  S  G  E  K  K  C  W  N  R      EXON 2

ATCAGGGCACTGCAGGAAACAATGCAAAGATGGAGAAGCAGTGAAAGATACATGCAAAAA  4140
gSerGlyHisCysArgLysGlnCysLysAspGlyGluAlaValLysAspThrCysLysAs
 S  G  H  C  R  K  Q  C  K  D  G  E  A  V  K  D  T  C  K  N

TCTTCGAGCTTGCTGCATTCCATCCAATGAAGACCACAGGCGAGTTCCTGCGACATCTCC  4200
nLeuArgAlaCysCysIleProSerAsnGluAspHisArgArgValProAlaThrSerPr
 L  R  A  C  C  I  P  S  N  E  D  H  R  R  V  P  A  T  S  P

CACACCCTTGAGTGACTCAACACCAGGAATTATTGATGATATTTTAACAGTAAGGTTCAC  4260
oThrProLeuSerAspSerThrProGlyIleIleAspAspIleLeuThrValArgPheTh
 T  P  L  S  D  S  T  P  G  I  I  D  D  I  L  T  V  R  F  T

GACAGACTACTTTGAAGTAAGCAGCAAGAAAGATATGGTTGAAGAGTCTGAGGCGGGAAG  4320
rThrAspTyrPheGluValSerSerLysLysAspMetValGluGluSerGluAlaGlyAr
 T  D  Y  F  E  V  S  S  K  K  D  M  V  E  E  S  E  A  G  R

GGGAACTGAGACCTCTCTTCCAAATGTTCACCATAGCTCATGACTTCCTCTCGGCTATCA  4380
gGlyThrGluThrSerLeuProAsnValHisHisSerSerEnd
 G  T  E  T  S  L  P  N  V  H  H  S  S  *

CTCACCCCTGTCCTCAGAGTGATAAACTAAGTCACATACAGATAAAGCACTGAAAACACC
ACAGTGACCCTCCCACCCCCCACCAATATGTAATTCTATTAATAGAAACAGCTGTGTAAA
GAAGTCTAAAATTTTCACTATTTCCAATGATAAACTCTTCAGTGCTCTTCTTGAaatgtc  4560 polyA site

acattatttccacaacaagttataacctatttttagtatttcttgtttgctagtgaccta  4620
 \\                                                          //  //
caggtgaaagcattcattctcctactaactatggccttggagccaggttttatctctcac  4860
```

FIG. 2. Human ESC42 gene aligned with the amino acid sequence. The gene sequence was extracted from GenBank (accession number AL031650). The cDNA nucleotide sequence is indicated in *uppercase letters*. Introns are shown in *lowercase letters*. The cDNA start site is near the TATA box (*double underline*) and the poly(A)$^+$ tail is attached shortly after the poly(A)$^+$ addition site (*double underline*). The predicted amino acid sequences are indicated in one- and three-letter abbreviations. ^, Indicates predicted signal peptide cleavage site.

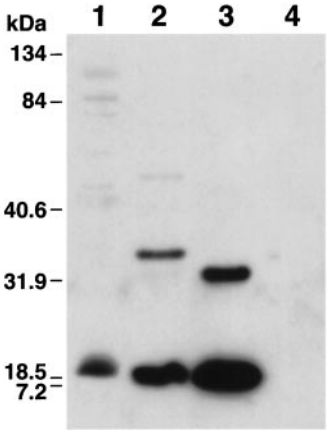

FIG. 3. ESC42 protein extracted from sperm appears larger than predicted from the cDNA sequence. ESC42 was resolved by SDS-PAGE and immunodetected using antibody to the ESC42 C-terminal peptide. Lane 1, Human sperm extract; lanes 2–4, extracts of COS-1 cells transfected with pSG5 encoding full length ESC42 protein (lane 2), pSG5 encoding the mature ESC42 protein without signal peptide (lane 3), and pSG5 empty vector (lane 4).

size was detected in oviduct, suggesting either low levels of expression of ESC42 or the presence of a related sequence. To determine if ESC42 expression is regulated by T, a key hormone in sperm maturation, RNA was obtained from rhesus monkeys that were sham operated, castrated at 6 d, and castrated at 6 d but given a single injection of 400 mg of T enanthate immediately after testis removal (Fig. 7). By 6 d after castration, ESC42 mRNA levels in caput and corpus declined to 10% of normal levels. Exogenous T maintained expression in caput but not in corpus. Serum T levels in monkeys that were sham operated, castrated, and castrated/androgen replaced before testis removal were 3.0, 3.8, and 1.4 ng/ml, respectively. Six days after injection, at the time the epididymides were removed, serum T levels were 2.0, 0.15, and 64.9 ng/ml in monkeys that were sham operated, castrated, and castrated/androgen replaced, respectively.

Discussion

The epididymis-specific cDNA library containing 25 new and 7 known unique cDNAs described in this report demonstrates the power of the subtracted library to open new avenues of investigation into epididymal function. Our results show that the epididymis produces mRNAs encoding a substantial number of novel proteins and novel variants of known proteins. They show that in the epididymis, motifs that have defined functions in other organs appear modified in sequence and in context, perhaps to serve the unique requirements of the epididymis. The specific expression of these mRNAs, androgen regulated and regionally localized in epididymis, is consistent with a role for their proteins in sperm maturation. Our results lay a foundation for further understanding of the proteomic profile of epididymal sperm maturation.

BLAST searches using the 25 cDNAs identified 18 of the human gene sequences, confirming the similarity of these monkey and human gene sequences and the usefulness of the monkey model system in understanding human epididymal function. The remaining 7 genes should become available as the human genome sequencing project progresses. Alternatively, these 7 may not be well conserved and may represent sequences substantially different in the rhesus monkey.

The 25 cDNAs encode proteins related to modifying enzymes (ESC177, ESC300, ESC615) and inhibitors (ESC6, ESC13) that could activate or protect sperm surface proteins. Two new lipocalins (ESC384, ESC513) from the cluster on human chromosome 9 bind undetermined ligands but are similar to known proteins that bind and transport retinoic acid and pheromones, both of which are important in fertility (41). The vitamin D-binding motif of ESC9 may also serve a hydrophobic ligand-binding/transport function. Recent evidence supports a host defense role for secreted ribonucleases similar to ESC461 (30). ESC112 contains a motif similar to the C3HC4-class zinc finger (38) in addition to its low homology with the secreted acrosomal protein SP10. A predicted signal peptide (35) in the ESC112 amino acid sequence suggests a secreted protein more than a nuclear protein The fibronectin II domains of ESC342 may bind collagen or phospholipids, as proposed for HE12 (28). The von Willebrand (ESC328) cysteine-rich motifs may mediate interactions with other proteins. ESC342 and ESC328 may interact with extracellular matrix proteins to stimulate intracellular signal transduction.

Characterization of these epididymis-specific clones has begun with the most abundant. The ESC42 mRNA is highly epididymis specific and androgen regulated. It is most abundant in caput and corpus in the sham-operated animal, reduced in caput and corpus in the castrated animal, and maintained only in caput in the T-replaced animal. It is puzzling that the mRNA is not detectable in the corpus of the T-replaced epididymis. Similar loss of mRNA in the corpus in the T-replaced animal was seen for 3 of the 10 other full length clones (data not shown). Region-specific regulation of gene expression by androgens in the epididymis is complex and was discussed in a recent review (42). The delivery of exogenous T through the general circulation may affect regionalized gene expression differently from T reaching the caput from the testis at high concentrations in luminal fluid. The T administered to the T-replaced animals raised circulating serum total T to levels similar to those reported for caput fluid in the rat (43). Testis factors may have a role in regulating gene expression in different regions of the epididymis, and the absence of factors could have contributed to an alteration in ESC42 mRNA stability or synthesis in response to exogenous T. Further experiments would be required with this castrated model to explain the regional differences in T response.

The ESC42 protein is expressed mainly in the caput and efferent ducts, but it is also detected in areas of the corpus and cauda. Binding to sperm was indicated by immunofluorescent detection on ejaculated human sperm and on Western blots of sperm extract. The protein contains a cysteine-rich region similar to the trefoil motif characteristic of a family that includes breast cancer estrogen-inducible gene (TFF1), stomach spasmolytic polypeptide (TFF2), and intestinal trefoil factor (TFF3) (for review, see Wong et al. [25]). Trefoil proteins are induced in ulceration and carcinoma and are involved in the cell migration and differentiation in damage repair and in cancer (44). TFFs are abundant in the gastro-

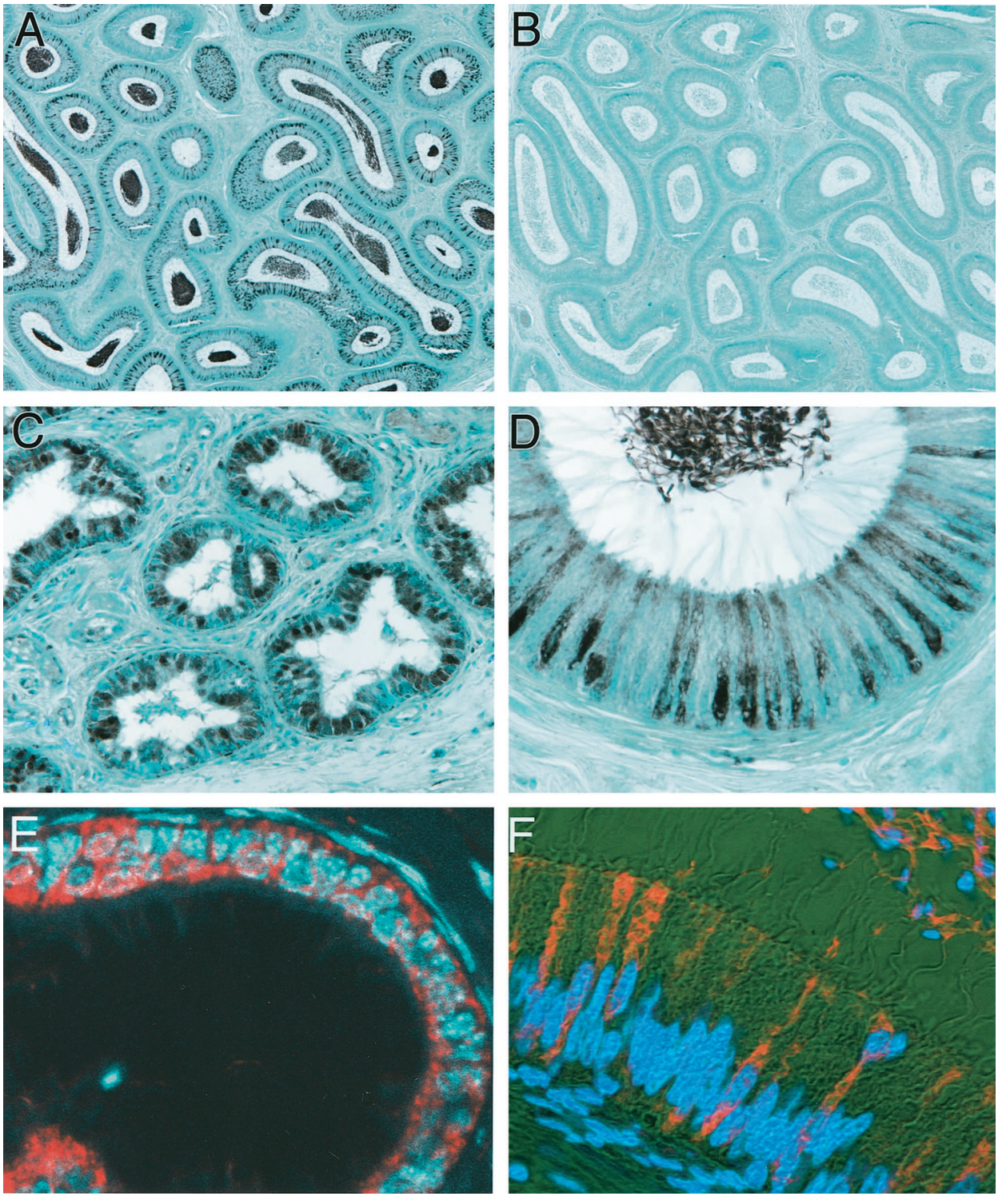

FIG. 4. Immunolocalization of ESC42 protein in rhesus monkey epididymis. A and B, Caput (4× objective). A, ESC42 appears *brown* against toluidine *blue* counterstain. B, Antibody was preabsorbed with peptide antigen. C, Efferent ducts (10× objective). D, Caput (40× objective). E, *Red* fluorescent stain indicates ESC42 cyanine-3-conjugated immunostaining; DAPI *blue* indicates nuclei. F, *Orange* indicates ESC42 protein; DAPI *blue* indicates nuclei. Artificial color was used to enhance contrast.

intestinal tract in association with mucins, where they modulate cell adhesion and migration in mucosal defense and repair. The types of injury to the epididymis in which ESC42 might function include damage to the epithelium by invading pathogens. Interaction of the TFF1 trefoil motif with a von Willebrand motif in mucins was recently reported (45), suggesting that the TFF may cross-link mucins, potentially affecting viscosity and sperm mobility.

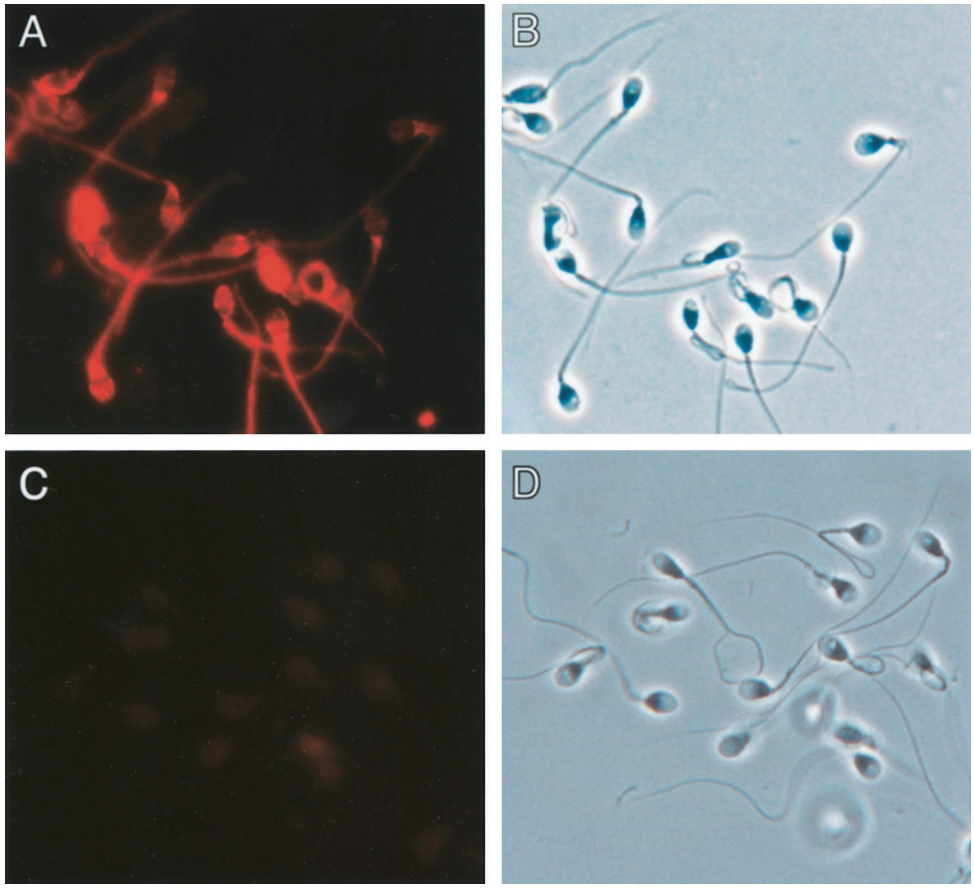

FIG. 5. Immunolocalization of ESC42 on human sperm. A, Texas Red immunofluorescent staining using affinity-purified anti-ESC42 is most intense in the postacrosomal head and neck regions. B, Phase contrast view of the sperm in A. C, Texas Red immunofluorescent staining using affinity-purified anti-ESC42 preabsorbed with peptide antigen. D, Phase contrast view of the sperm in C. Magnification ×500.

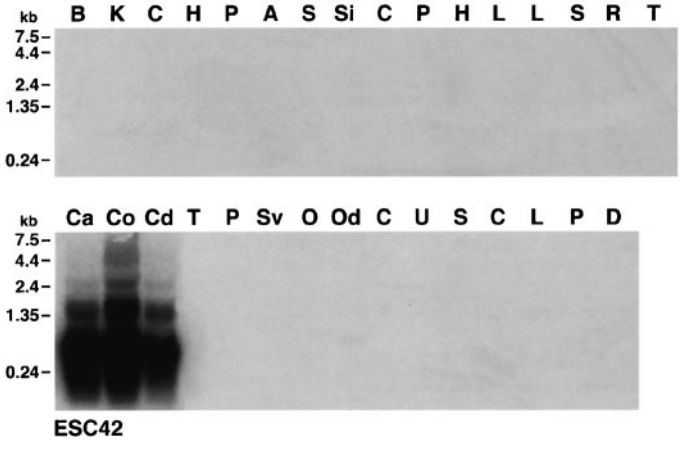

FIG. 6. Expression of ESC42 mRNA in different rhesus monkey tissues. Total RNA (12 μg/lane) was isolated from (left to right) bladder, kidney, cerebrum, hypothalamus, pituitary, adrenal, stomach, small intestine, colon, pancreas, heart, liver, lung, spleen, retina, tonsil, caput, corpus, cauda, testis, prostate, seminal vesicle, ovary, oviduct, cervix, uterus, salivary gland, CV-1 cells, LNCaP cells, PC3 cells, and DU145 cells. The Northern blot was hybridized to ESC42 cDNA encoding the mature protein. Film was overexposed to reveal hybridization in nonepididymal RNAs.

A related trefoil-like motif is found in the carbohydrate-degrading enzymes sucrase-isomaltase and lysosomal α-glucosidase, in which the motif is suggested to bind carbohydrate, raising the possibility of a role in the degradation of extracellular glycoprotein matrix (25).

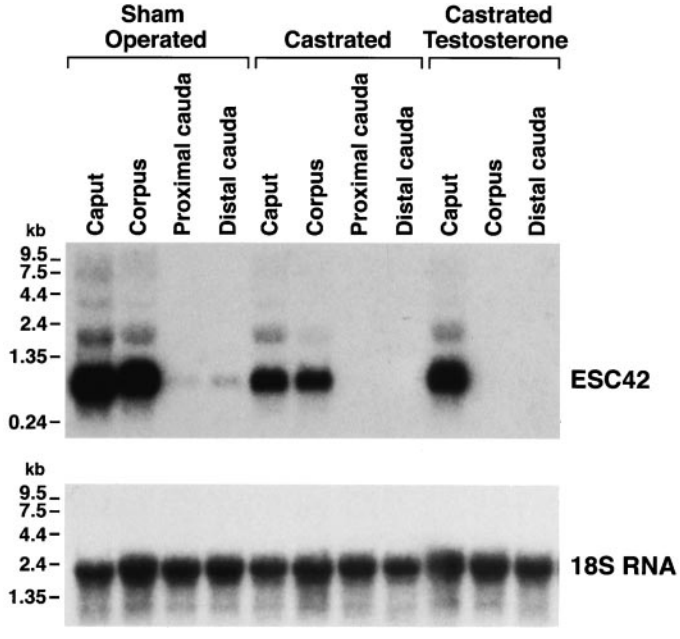

FIG. 7. Androgen regulation of ESC42 mRNA. Rhesus monkeys were sham operated, castrated, or castrated and androgen replaced with immediate injection of T enanthate. Epididymides were removed 6 d after castration. *Top*, Total RNAs (10 μg/lane) from the indicated epididymal regions were analyzed by Northern hybridization to ^{32}P-labeled ESC42 full-length cDNA. *Bottom*, The same blot was hybridized to ^{32}P-labeled 18S rRNA.

Large trefoil-like motifs appear in pig and rabbit zona pellucida protein 3 (46), heavily glycosylated extracellular matrix proteins surrounding oocytes, as well as in uromodulin, a protein thought to protect against urinary tract infections by certain microorganisms (47). Interaction of the trefoil motifs in frog mucin FIM-A.1 with surface carbohydrates of oral microorganisms was suggested to protect against pathological invasion (48).

The motif most closely matching the cysteine-rich region of ESC42 is the 50% identical N terminus of another epididymis-specific protein, ESP13.2 (49) (GenBank accession number AJ236910). The function of this protein is not known, but the authors compare the cysteine-rich region to similar domains in defensins, molecules of innate immunity. Recent reports describe the antimicrobial functions in the epididymis of human cationic antimicrobial protein (50) and the rat Bin1b (51). Bin1b prevents the growth of *Escherichia coli* in epididymal cultures and exhibits structural similarity to HE2β1 (18), a human sperm-binding protein. The cysteine-rich domain in Bin1b is also similar to the six-cysteine array in ESC42. Antimicrobial peptides bound to the sperm surface may promote fertility by protecting sperm against attack by bacteria or viruses in the male and female reproductive tracts.

Acknowledgments

We thank Dr. James L. Mohler for performing surgical castrations of monkeys in the androgen regulation experiment, Zhang Jining, Zang De-Ying, and Raymond Johnson for expert technical assistance, and Alain Burrette and Ting Qiang for confocal microscopy teaching expertise, reagents, and protocols. Special thanks to Richard L. Pippin and Betty F. Horton for expert preparation of the figures.

Received March 21, 2001. Accepted June 5, 2001.

Address all correspondence and requests for reprints to: Susan H. Hall, Laboratories for Reproductive Biology, CB 7500, University of North Carolina, Chapel Hill, North Carolina 27599. E-mail: shh@med.unc.edu.

This work was supported by the Consortium for Industrial Collaboration in Contraceptive Research Program of the Contraceptive Research and Development Program, Eastern Virginia Medical School (Norfolk, VA). The views expressed by the authors do not necessarily reflect the views of the Contraceptive Research and Development Program or the Consortium for Industrial Collaboration in Contraceptive Research. This work was also supported by NIH Grant R37-HD-04466, by the National Institute of Child Health and Human Development/NIH through cooperative agreement U54-HD-35041 as part of the Specialized Cooperative Centers Program in Reproduction Research, and by Fogarty International Center Training and Research in Population and Health Grant D43TW/HD-00627.

* Present address: State Key Laboratory of Molecular Biology, Shanghai Institute for Biological Science, Institute of Biochemistry and Cell Biology, Sinica Academia, Shanghai 200031, China.

† Present address: Department of Biochemistry, Indian Institute of Science, Bangalore, India 560012.

References

1. **Yanagimachi R** 1994 Mammalian fertilization. In: Knobil E, Neill JD, eds. The physiology of reproduction. 2nd ed. New York: Raven Press Ltd.; 189–317
2. **Cooper TG** 1992 Epididymal proteins and sperm maturation. In: Nieschlag E, Habenicht UF, eds. Spermatogenesis-fertilization-contraception. Berlin: Springer-Verlag; 285–318
3. **Eddy EM, O'Brien DA** 1994 The spermatozoon. In: Knobil E, Neill JD, eds. The physiology of reproduction. 2nd ed. New York: Raven Press Ltd.; 29–77
4. **Brooks DE** 1987 Androgen-regulated epididymal secretory proteins associated with post-testicular sperm development. In: Orgebin-Crist MC, Danzo BJ, eds. Cell biology of the testis and epididymis. New York: New York Academy of Sciences; 179–194
5. **Okamura N, Tamba M, Liao HJ, et al.** 1995 Cloning of complementary DNA encoding a 135-kilodalton protein secreted from porcine corpus epididymis and its identification as an epididymis-specific α-mannosidase. Mol Reprod Dev 42:141–148
6. **Syntin P, Dacheux F, Druart X, Gatti JL, Okamura N, Dacheux JL** 1996 Characterization and identification of proteins secreted in the various regions of the adult boar epididymis. Biol Reprod 55:956–974
7. **Fouchécourt S, Métayer S, Locatelli A, Dacheux F, Dacheux JL** 2000 Stallion epididymal fluid proteome: qualitative and quantitative characterization, secretion and dynamic changes of major proteins. Biol Reprod 62:1790–1803
8. **Zini A, Schlegel PN** 1997 Identification and characterization of antioxidant enzyme mRNAs in the rat epididymis. Int J Androl 20:86–91
9. **Cornwall GA, Hsia N** 1997 ADAM7, a member of ADAM (a disintegrin and metalloprotease) gene family is specifically expressed in the mouse anterior pituitary and epididymis. Endocrinology 138:4262–4272
10. **Kirchhoff C, Osterhoff C, Pera I, Schroter S** 1998 Function of human epididymal proteins in sperm maturation. Andrologia 30:225–232
11. **Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ** 1979 Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299
12. **Ausubel FM, Brent R, Kingston RE, et al., eds.** 1994 Current protocols in molecular biology. New York: John Wiley & Sons
13. **Hamil KG, Hall SH** 1994 Cloning of rat Sertoli cell follicle-stimulating hormone primary response complementary deoxyribonucleic acid: regulation of TSC-22 gene expression. Endocrinology 134:1205–1212
14. **Sambrook J, Fritsch EF, Maniatis T** 1989 Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 7.51–7.52
15. **Zhang X-Z, Donovan MP, Williams BT, Mohler JL** 1996 Comparison of subcapsular and total orchiectomy for treatment of metastatic prostate cancer. Urology 47:402–404
16. **Lopata MA, Cleveland DW, Sollner-Webb B** 1984 High level transient expression of a chloramphenicol acetyl transferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. Nucleic Acids Res 12:5707–5717
17. **Gregory CW, Kim D, Ye P, et al.** 1999 Androgen receptor up-regulates insulin-like growth factor binding protein-5 (IGFBP-5) expression in a human prostate cancer xenograft. Endocrinology 140:2372–2381
18. **Hamil KG, Sivashanmugam P, Richardson RT, et al.** 2000 HE2β and HE2γ, new members of an epididymis-specific family of androgen-regulated proteins in the human. Endocrinology 141:1245–1253
19. **Preece A** 1972 A manual for histologic technicians. 3rd ed. Boston: Little, Brown & Co.
20. **Bobrow MN, Harris TD, Shaughnessy KJ, Litt GJ** 1989 Catalyzed reporter deposition, a novel method of signal amplification: application to immunoassays. J Immunol Methods 125:279–285
21. **Adams JC** 1992 Biotin amplification of biotin and horseradish peroxidase signals in histochemical stains. J Histochem Cytochem 40:1457–1463
22. **Horii A, Tomita N, Yokouchi H, et al.** 1989 On the cDNAs for two types of rat pancreatic secretory trypsin inhibitor. Biochem Biophys Res Commun 162:151–159
23. **Osawa M, Tsuji T, Yukawa N, Saito T, Takeichi S** 1994 Cloning and sequence analysis of cDNA encoding rabbit vitamin D-binding protein (Gc globulin). Biochem Mol Biol Int 34:1003–1009
24. **Brown WM, Dziegielewska KM** 1997 Friends and relations of the cystatin superfamily: new members and their evolution. Protein Sci 6:5–12
25. **Wong WM, Poulsom R, Wright NA** 1999 Trefoil peptides. Gut 44:890–895
26. **Freemerman AJ, Flickinger CJ, Herr JC** 1995 Characterization of alternatively spliced human SP-10 mRNAs. Mol Reprod Dev 41:100–108
27. **Roiko K, Jänne OA, Vihko P** 1990 Primary structure of rat secretory acid phosphatase and comparison to other acid phosphatases. Gene 89:223–229
28. **Saalmann A, Munz S, Ellerbrock K, Ivell R, Kirchhoff C** 2001 Novel sperm-binding proteins of epididymal origin contain four fibronectin type II modules. Mol Reprod Dev 58:88–100
29. **Lareyre J-J, Winfrey VP, Kasper S, et al.** 2001 Gene duplication gives rise to a new 17-kilodalton lipocalin that shows epididymal region-specific expression and testicular factor(s) regulation. Endocrinology 142:1296–1308
30. **Rosenberg HF, Domachowske JB** 1999 Eosinophils, ribonucleases and host defense: solving the puzzle. Immunol Res 20:261–274
31. **Bocskei Z, Groom CR, Flower DR, et al.** 1992 Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. Nature 360:186–188
32. **Sivashanmugam P, Richardson RT, Hall S, Hamil KG, French FS, O'Rand MG** 1999 Cloning and characterization of an androgen-dependent acidic epididymal glycoprotein/CRISP-1 like protein from the monkey. J Androl 20:384–393
33. **Kirchhoff C** 1999 Gene expression in the epididymis. Int Rev Cytol 188:133–202
34. **Venter JC, Adams MD, Myers EW, et al.** 2001 The sequence of the human genome. Science 291:1304–1351
35. **Nielsen H, Engelbrecht J, Brunak S, von Heijne G** 1997 Identification of

prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6
36. **Hofmann K, Bucher P, Falquet L, Bairoch A** 1999 The PROSITE database, its status in 1999. Nucleic Acids Res 27:215–219
37. **Lüthy R, Zenarios I, Bucher P** 1994 Improving the sensitivity of the sequence profile method. Protein Sci 3:139–146
38. **Wu C, Shivakumar S, Shivakumar CV, Chen S** 1998 GeneFIND webserver for protein family identification and information retrieval. Bioinformatics 14:223–224
39. **Barker PA, Murphy RA** 1992 The nerve growth factor receptor: a multicomponent system that mediates the actions of the neurotrophin family of proteins. Mol Cell Biochem 110:1–15
40. **Beck S, Schmitt H, Shizuya H, Blin N, Gött P** 1996 Cloning of contiguous genomic fragments from human chromosome 21 harbouring three trefoil peptide genes. Hum Genet 98:233–235
41. **Flower DR** 1996 The lipocalin protein family: structure and function. Biochem J 318:1–14
42. **Orgebin-Crist M-C** 1996 Androgens and epididymal function. In: Bhasin S, Gabelnick HL, Spieler JM, Swerdloff RS, Wan C, Kelly C, eds. Pharmacology, biology, and clinical applications of androgens. New York: Wiley-Liss, Inc.; 27–38
43. **Turner TT, Jones CE, Howards SS, Ewing LL, Zegeye B, Gunsalus GL** 1984 On the androgen microenvironment of maturing spermatozoa in the adult rat. Endocrinology 115:1925–1932
44. **Efstathiou JA, Pignatelli M** 1998 Modulation of epithelial cell adhesion in gastrointestinal homeostasis. Am J Pathol 153:341–347
45. **Tomasetto C, Masson R, Linares J-L, et al.** 2000 pS2/TFF1 interacts directly with the VWFC cysteine-rich domains of mucins. Gastroenterology 118:70–80
46. **Bork P** 1993 A trefoil domain in the major rabbit zona pellucida protein. Protein Sci 2:669–670
47. **Reinhart HH, Obedeanu N, Robinson R, Korzeniowski O, Kaye D, Sobel JD** 1991 Urinary excretion of Tamm-Horsfall protein in elderly women. J Urol 146:806–808
48. **Hauser F, Gertzen EM, Hoffmann W** 1990 Expression of spasmolysin (FIM-A.1): an integumentary mucin from *Xenopus laevis*. Exp Cell Res 189:157–162
49. **Perry ACF, Jones R, Moisyadi S, Coadwell J, Hall L** 1999 The novel epididymal secretory protein ESP13.2 in *Macaca fascicularis*. Biol Reprod 61:965–972
50. **Nilsson MF, Sandstedt B, Sørensen O, Weber G, Borregaard N, Ståhle-Bäckdahl M** 1999 The human cationic antimicrobial protein (hCAP18), a peptide antibiotic, is widely expressed in human squamous epithelia and colocalizes with interleukin-6. Infect Immun 67:2561–2566
51. **Li P, Chan HC, He B, et al.** 2001 An antimicrobial peptide gene found in the male reproductive system of rats. Science 291:1783–1785