RESEARCH COMMUNICATION

A mammalian epididymal protein with remarkable sequence similarity to snake venom haemorrhagic peptides

Anthony C. F. PERRY,* Roy JONES,† Patrick J. BARKER‡ and Len HALL*§

*Department of Biochemistry, University of Bristol, School of Medical Sciences, Bristol BS8 1TD, U.K.,

and †Department of Biochemistry and ‡Microchemical Facility, Institute of Animal Physiology and Genetics Research, Cambridge Research Station, Babraham, Cambridge CB2 4AT, U.K.

Following spermatogenesis in the testis, mammalian spermatozoa pass into the epididymis, where they undergo changes which confer on them forward motility and the ability to recognize and penetrate the egg. Many of these maturation events involve androgen-regulated epididymal proteins which become associated with the sperm membrane, and/or effect changes to integral sperm membrane proteins. Here we report the sequence of an 89 kDa androgen-regulated protein from rat (*Rattus norvegicus*) and monkey (*Macaca fascicularis*) epididymis that is synthesized exclusively in the caput region and is localized on the apical surface of its principal epithelial cells. This protein shows remarkable similarity to a variety of proteases and disintegrins found in snake venoms and is similar to, but distinct from, the guinea-pig sperm surface PH-30 α/β complex recently implicated in sperm-egg recognition and fusion.

INTRODUCTION

Newly synthesized mammalian spermatozoa are immature and incapable of fertilization when they emerge from the testis. However, during subsequent passage down the epididymis they acquire forward motility and the ability to recognize and penetrate the zona pellucida of the egg. The complex biochemical and physiological changes which lead to this acquisition of fertilizing capacity are collectively known as sperm maturation, a co-operative process involving sperm, epididymal fluid, epididymal epithelium and their interstitium. In particular, it is now clear that androgen-related epididymal proteins play an important role in normal sperm maturation.

Recently we have isolated and partially characterized a number of abundant cDNA clones from rat and monkey epididymal libraries. One of these rat clones, prE15, identified a 4 kb androgen-sensitive epididymis-specific transcript which was expressed almost exclusively in the caput region [1]. Epididymal prE15 transcript levels declined very rapidly after castration, but were re-established to around 70 % of normal steady-state levels after subsequent testosterone administration [1].

We now report the isolation and characterization of full-length cDNA clones encoding an 89 kDa protein, corresponding to rat prE15 and its monkey homologue. Western-blot analysis of epididymal extracts, using antisera raised against a synthetic peptide, identify a 90 kDa protein in the caput region, and immunofluorescence studies *in situ* localize it to the apical surface of its principal epithelial cells. This protein, which we have called Epididymal Apical Protein I (EAP I), shows intriguing similarity to a variety of snake venom haemorrhagic peptides and is also similar to, but distinct from, the guinea-pig sperm surface PH-30 α/β complex implicated in sperm-egg recognition and fusion [2].

MATERIALS AND METHODS

Materials

Nitrocellulose filters and poly(vinylidene difluoride) (PVDF) membranes were from Schleicher and Schull and Millipore respectively. Fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit IgG was obtained from Sigma, Poole, Dorset, U.K. All other chemicals were of AnalaR grade or the purest grade available.

Epididymides from Wistar rats (*Rattus norvegicus*) and monkeys (*Macaca fascicularis*) were stored at -80°C until required. Castration and testosterone-replacement procedures were carried out as previously described [1].

Isolation of EAP I cDNA clones and DNA sequence analysis

Approx. 3×10^4 rat epididymal cDNA clones [3] in pAT153 were transferred to nitrocellulose filters and screened with an oligonucleotide probe derived from the rat epididymal cDNA clone prE15 [1]. After clone purification and re-screening, plasmid DNA was isolated from several positive clones. The cDNA insert from one full-length clone, prE-EAPI, was purified and used to isolate several monkey homologues from a monkey epididymal cDNA library [3]. The cDNA sequences of prE-EAPI and one full-length monkey homologue, pmE-EAPI, were completely sequenced on both DNA strands by using a custom primer walking strategy on a Du Pont Genesis 2000 automated sequencer utilizing fluorescently labelled dideoxynucleotides.

Preparation of polyclonal antisera, tissue extracts and Western-blot analysis

Antisera to a rat EAP I C-terminal 25-residue peptide (see Fig.

Abbreviation used: EAP I, epididymal apical protein I.

§ To whom correspondence should be addressed.

The nucleotide sequences of rat (*Rattus norvegicus*) and monkey (*Macaca fascicularis*) EAP I cDNAs will appear in the EMBL Nucleotide Sequence Database under the accession numbers X66140 and X66139 respectively.

Fig. 3. Immunofluorescence localisation of EAP I in tissue sections of caput epididymidis

(a) Preimmune serum. (b) EAP I antiserum. Phase contrast is shown on the right-hand side and immunofluorescence on the left. Magnification $\times 337$. Regions shown are from distal caput epididymidis.

and epifluorescence microscopy on a Zeiss Axiophot photomicroscope.

RESULTS AND DISCUSSION

Cloning and characterization of rat and monkey EAP I cDNAs

A probe derived from a rat partial cDNA clone (prE15) known to correspond to 4 kb androgen-sensitive tissue-specific transcript [1] was used to isolate a full-length clone (prE-EAPI) from a rat epididymal cDNA library, and this in turn was used to isolate a monkey homologue (pmE-EAPI). The complete cDNA sequences of both clones (prE-EAPI, 3586 bp; pmE-EAPI, 3360 bp) were determined [see EMBL database accession numbers X66140 (rat) and X66139 (monkey)] and found to possess open reading frames of approx. 780 residues (Fig. 1). The predicted protein products (EAP I; ~ 89 kDa) are highly conserved between the two species (72% overall sequence identity).

Tissue specificity and hormonal regulation of EAP I expression

To investigate the site and hormonal regulation of EAP I expression, a C-terminal peptide was synthesized (Fig. 1), and polyclonal antisera were raised, and used to localize EAP I by Western immunoblotting. A protein of the expected size (~ 90 kDa) was obtained with rat protein extracts from the initial segment, and with proximal and distal caput regions of the epididymis (Fig. 2a), but not with the corpus or cauda regions or a variety of other rat tissues (Fig. 2b). Coincidence between the cloned protein and the antigen detected by the antibody was confirmed by N-terminal microsequencing (see Fig. 1). EAP I could not be detected in the caput epididymidis of 6-day-castrated animals, but was induced in castrated rats subsequently treated

for 6 days with testosterone propionate (Fig. 2c). However, the level of EAP I in testosterone-treated castrated rats was notably lower than that in untreated animals, a result very similar to that found at the RNA level [1], suggesting that testicular factors in addition to testosterone may be necessary for normal levels of expression.

Immunofluorescence localization of EAP I expression

Frozen sections of rat caput epididymidis probed with rat anti-EAP I antiserum showed strong fluorescence around the apical surface in the vicinity of the surface microvilli of the principal epithelial cells that line the duct (Fig. 3b). Only weak background fluorescence was detected in sections of the testis and cauda epididymidis (results not shown), confirming the results obtained by Western blotting. Sections probed with preimmune serum were consistently negative (Fig. 3a).

Sequence similarity between EAP I, snake venom haemorrhagic proteins and the guinea-pig sperm surface PH-30 α/β complex

In an attempt to gain an insight into the function of EAP I, databases were searched and found to reveal an intriguing conservation with several snake venom components (Fig. 4). Initially two regions were identified within EAP I which showed marked similarity to Habu (*Trimeresurus flavoviridis*) venom haemorrhagic protein HR2A [5] (a zinc protease with basement-membrane-hydrolysing activity [6]) and a Habu venom disintegrin [7] (a small soluble integrin ligand which binds to the glycoprotein IIb-IIIa receptor on the platelet surface, and inhibits fibrinogen binding and hence platelet aggregation). These two domains were located near the centre of the deduced EAP I sequence (Fig. 4). More recently a novel Habu mosaic protein,

Fig. 4. Sequence similarity of EAP I to snake venom proteins and to guinea-pig PH-30 α and β subunits

(a) The amino acid sequences of rat EAP I, Habu venom haemorrhagic proteins HR2A (2A) [5] and HR1B (1B) [8], Habu disintegrin (D) [7] and the mature α and β subunits of guinea-pig PH-30 (30 α and 30 β) [2] were aligned by eye with gaps introduced to maximize the alignment: (|) indicates amino acid identity between adjacent sequences. (b) Diagrammatic representation of above sequence alignments to highlight the different domains.

HR1B [8], has been sequenced and found to contain both a zinc protease and a disintegrin domain within the same polypeptide chain (Fig. 4). This protein, which has 10-fold higher haemorrhagic activity than HR2A, making it a major lethal factor in Habu venom, also shows significant sequence alignment with EAP I. Interestingly, HR1B lacks the Arg-Gly-Asp motif [9] characteristic of many integrin ligands, a motif which is also absent from EAP I. The striking similarity of EAP I to these venom components suggests that it may bind to an integrin-like receptor.

Very recently, Blobel and co-workers [2] have reported the deduced sequences of the mature α and β subunits of guinea-pig PH-30 (a sperm surface protein thought to be involved in sperm-egg recognition and subsequent fusion), determined by a combination of PCR and analysis of partial cDNA clones. Of particular interest was the finding that the mature β subunit and α precursor contain domains at their *N*-termini which show a high degree of sequence similarity to a number of integrin ligands found in snake venoms. In addition, the α subunit contains a putative fusion peptide, similar to that found in rubella virus E2. The authors present several lines of indirect evidence to support their hypothesis that PH-30 β binds to an egg plasma membrane integrin, and suggest that the interaction between species-specific pairs of PH-30 homologues and their corresponding integrins could account for the additional specificity of sperm-egg plasma-membrane interactions.

Despite an obvious similarity between the subunits of guinea-pig PH-30 and rat and monkey EAP I (Fig. 4), this similarity is confined to the disintegrin-like domain and the position of the putative transmembrane domain relative to it. However, both mature subunits of PH-30 are derived from larger precursors [10] of similar size to EAP I. It is therefore possible that these precursor forms may contain a similar domain to that found in EAP I, and indeed that such proteases may be involved in EAP I and PH-30 processing.

Several lines of evidence indicate that PH-30 and EAP I are not species homologues of the same protein, and indeed possess important differences. (i) PH-30 is synthesized in the testis [2]. PCR amplification of transcripts corresponding to part of the monkey homologue of PH-30, using primers based on the guinea-pig sequence, identifies amplified products with a different sequence from that of monkey EAP I (but similar to PH-30) in monkey testicular cDNA preparations (results not shown). EAP I transcripts, on the other hand, are abundant in the caput region of the epididymis and are undetectable in the testis [1]. PCR amplification of guinea-pig EAP I transcripts using primers based on the rat sequence identify an abundant product of the expected size in epididymal, but not testicular, cDNA preparations and possess a sequence similar to that of rat EAP I but different from that of PH-30 (result not shown). Hence PH-30 is of testicular origin in all three species tested, whereas EAP I is of

epididymal origin. (ii) Rat and monkey EAP I exhibit a very high degree of sequence similarity (72% identity overall, 77% in the disintegrin domain). Guinea-pig PH-30 shows less than 23% sequence identity to EAP I overall (49% in the disintegrin domain). (iii) EAP I lacks the insertion which contains the putative fusion peptide found in PH-30 α and the epidermal-growth-factor-like repeat found in both subunits of PH-30.

Nevertheless, the extensive work on the properties of PH-30, its proposed importance in sperm-egg fusion and its similarity to EAP I, must be borne in mind when attempting to establish a biological role for the latter. Clearly the presence of an additional protein, EAP I, with putative integrin-binding properties, may add further complexities to the role of PH-30 proposed by Blobel and co-workers [2] and suggest that integrin-disintegrin interactions may be involved in epididymal maturation of spermatozoa as well as in the sperm-egg recognition process. The multidomain nature of EAP I, its tissue specificity and its striking similarity to Habu HR1B raise interesting questions about the evolution of these mosaic proteins and may provide a suitable framework for the design of novel, targeted, therapeutic anticoagulants. In addition, further studies on the role of EAP I in a sperm maturation may help in the design of contraceptive vaccines, since the epididymis is an obvious target for fertility control.

This work was jointly supported by the Medical Research Council, Agricultural and Food Research Council and University of Bristol Molecular Recognition Centre. We are grateful to Huntingdon Research Centre for providing monkey tissues.

REFERENCES

1. Walker, J. E., Jones, R., Moore, A., Hamilton, D. W. & Hall, L. (1990) *Mol. Cell. Endocrinol.* **74**, 61–68
2. Blobel, C. P., Wolfsberg, T. G., Turck, C. W., Myles, D. G., Primakoff, P. & White, J. M. (1992) *Nature (London)* **356**, 248–252
3. Perry, A. C. F., Jones, R., Niang, L. S. P., Jackson, R. M. & Hall, L. (1992) *Biochem. J.* **285**, 863–870
4. Jones, R., Brown, C. R., von Glos, K. I. & Gaunt, S. J. (1985) *Exp. Cell Res.* **156**, 31–44
5. Miyata, T., Takeya, H., Ozeki, Y., Arakawa, M., Tokunaga, F., Iwanaga, S. & Omori-Satoh, T. (1989) *J. Biochem. (Tokyo)* **105**, 847–853
6. Ohsaka, A., Just, M. & Habermann, E. (1973) *Biochim. Biophys. Acta* **323**, 415–428
7. Musial, J., Niewiarowski, S., Rucinski, B., Stewart, G. J., Cook, J. J., Williams, J. A. & Edmunds, L. H. (1990) *Circulation* **82**, 261–273
8. Takeya, H., Oda, K., Miyata, T., Omori-Satoh, T. & Iwanaga, S. (1990) *J. Biol. Chem.* **265**, 16068–16073
9. D'Souza, S. E., Ginsberg, M. H. & Plow, E. F. (1991) *Trends Biochem. Sci.* **16**, 246–250
10. Blobel, C. P., Myles, D. G., Primakoff, P. & White, J. M. (1990) *J. Cell Biol.* **111**, 69–78

Received 8 July 1992; accepted 15 July 1992