

The CAP Superfamily: Cysteine-Rich Secretory Proteins, Antigen 5, and Pathogenesis-Related 1 Proteins—Roles in Reproduction, Cancer, and Immune Defense

Gerard M. Gibbs, Kim Roelants, and Moira K. O'Bryan

Monash Institute of Medical Research (G.M.G., M.K.O.), Monash University, Melbourne, 3168, Australia; Biology Department (K.R.), Vrije Universiteit Brussel, B-1050 Brussels, Belgium; and The Australian Research Council Centre of Excellence in Biotechnology and Development (M.K.O.), Monash University, Melbourne, 3168, Australia

The cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins (CAP) superfamily members are found in a remarkable range of organisms spanning each of the animal kingdoms. Within humans and mice, there are 31 and 33 individual family members, respectively, and although many are poorly characterized, the majority show a notable expression bias to the reproductive tract and immune tissues or are deregulated in cancers. CAP superfamily proteins are most often secreted and have an extracellular endocrine or paracrine function and are involved in processes including the regulation of extracellular matrix and branching morphogenesis, potentially as either proteases or protease inhibitors; in ion channel regulation in fertility; as tumor suppressor or protooncogenic genes in tissues including the prostate; and in cell-cell adhesion during fertilization. This review describes mammalian CAP superfamily gene expression profiles, phy-

logenetic relationships, protein structural properties, and biological functions, and it draws into focus their potential role in health and disease. The nine subfamilies of the mammalian CAP superfamily include: the human glioma pathogenesis-related 1 (GLIPR1), Golgi associated pathogenesis related-1 (GAPR1) proteins, peptidase inhibitor 15 (PI15), peptidase inhibitor 16 (PI16), cysteine-rich secretory proteins (CRISPs), CRISP LCCL domain containing 1 (CRISPLD1), CRISP LCCL domain containing 2 (CRISPLD2), mannose receptor like and the R3H domain containing like proteins. We conclude that overall protein structural conservation within the CAP superfamily results in fundamentally similar functions for the CAP domain in all members, yet the diversity outside of this core region dramatically alters target specificity and, therefore, the biological consequences. (*Endocrine Reviews* 29: 865–897, 2008)

- I. Introduction
- II. The CAP Superfamily Conservation and Evolutionary Distribution
- III. Mammalian CAP Superfamily Proteins
 - A. Cysteine-rich secretory proteins (CRISPs) subfamily
 - B. Glioma pathogenesis related-1 (GLIPR1) subfamily
 - C. Glioma pathogenesis related-2 (GLIPR2)/Golgi-associated pathogenesis related-1 (GAPR-1) subfamily
 - D. Peptidase inhibitor 15 (PI15) subfamily

First Published Online September 29, 2008

Abbreviations: Ag5, Antigen 5; ASP, activation-associated secreted protein; CAP, cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 protein; CAPCL, CAP and CTL domain containing; CAPLD1, CAP and LCCL domain containing protein 1; CRISP, cysteine-rich secretory protein; CRISPLD1, CRISP LCCL domain containing 1; CTL, C-type lectin; DFNA9, deafness autosomal dominant nonsyndromic sensorineural disorder 9; E, embryonic day; EST, expressed sequence tag; GAPR-1, Golgi-associated pathogenesis related-1; GGN1, gametogenetin 1; GLIPR1, glioma pathogenesis related-1; GLIPR1L1, GLIPR1-like 1; GPI, glycosyl-phosphatidylinositol; HPI, hookworm platelet inhibitor; ICR, ion channel regulator; IVF, *in vitro* fertilization; LCCL, limulus factor C, Coch-5b2, and Lgl1; Lgl1, late gestation lung protein 1; MAP3KII, MAPK kinase kinase II; MRL, mannose receptor-like; β -MSP, β -microseminoprotein; NIF, neutrophil inhibitory factor; NMR, nuclear magnetic resonance; PDB, Protein Data Bank; PI15, peptidase inhibitor 15; Pr-1, pathogenesis-related-1; PsTx, pseudochetoxin; R3HDML, R3H domain-like; RTVP1, related to testis specific, vespid and Pr-1; RyR, ryanodine receptor; SCP, sperm-coating glycoprotein; SNP, single nucleotide polymorphism; WT, Wilms' tumor.

Endocrine Reviews is published by The Endocrine Society (<http://www.endo-society.org>), the foremost professional society serving the endocrine community.

- E. Peptidase inhibitor 16 (PI16) subfamily
- F. CRISPLD1/CAPLD1 subfamily
- G. CRISPLD2/CAPLD2 subfamily
- H. Mannose receptor-like/CAP domain and CTL domain containing (CAPCL) subfamily
 - I. Peptidase inhibitor R3H domain containing-like subfamily
- IV. Structural Characteristics of CAP Superfamily Proteins
 - A. The CAP domain
 - B. The cysteine-rich domain or CRISP domain
 - C. C-type lectin (CTL) domains
 - D. The limulus factor C, Coch-5B2, and Lgl1 (LCCL) domains
- V. CAP Proteins in Reproduction
 - A. Spermatogenesis
 - B. Epididymal maturation and sperm capacitation
 - C. CAP's role in sperm and egg fusion
 - D. Female reproductive tract
 - E. Genetic studies associate SNPs in CRISPs with reduced fertility
- VI. CAP Proteins in Cancer
- VII. CAP Proteins in Immune Regulation
 - A. Anclystoma-secreted protein, or activation-associated secreted protein (ASP)
 - B. Antigen 5 (AG5)
 - C. Pathogenesis related-1 (PR-1)
 - D. An immune function for mammalian CAPs
- VIII. CAP Protein in Nonreproductive Tract Pathologies
 - A. Sjögren's syndrome
 - B. Chronic pancreatitis

ular organ systems. Lastly, we have described what is known about individual CAPs in specific biological processes and diseases. For context, several of these sections are illustrated with data from nonmammalian CAP proteins for which considerably more information has been obtained. Although there is limited detail on the biological function of many of the mammalian CAP superfamily proteins, through this systematic analysis a more comprehensive global view of the CAP superfamily is attainable. This is a view that we hope will prove to be a useful basis and stimulus for developing functional hypotheses for all mammalian CAPs, including the many that remain virtually unstudied.

II. The CAP Superfamily Conservation and Evolutionary Distribution

The CAP superfamily is also referred to as the sperm-coating glycoprotein (SCP) protein family (Pfam PF00188). The CAP nomenclature was derived from the three major groups of proteins originally identified and, in contrast to SCP, more accurately reflects the diversity in the superfamily. The CAP superfamily is characterized by several signature sequences, two of which are defined in the PROSITE database (<http://www.expasy.ch/prosite/>) as: CAP1, [GDER][HR][FYWH][TVS][QA][LIVM][LIVMA]Wxx[STN], and CAP2, [LIVMFYH][LIVMFY]xC[NQRHS]Yx[PARH]x[GL]N[LIVMFYWDN]. Two additional motifs are apparent and, while noted previously (1, 11, 12), have not been adopted as part of the consensus for the superfamily within PROSITE. Herein, we define these two additional conserved sequence motifs as CAP3 (HNxxR) and CAP4 (G[EQ]N[ILV]) (Fig. 1).

The tertiary crystal structure of several CAP superfamily proteins has been determined (Section V). They show that the CAP motifs are present in a small, disulfide bond stabilized and structurally conserved 17- to 21-kDa domain referred to as the CAP (or SCP or Pr-1) domain. The presence of these sequence motifs and the resultant CAP domain is characteristic and definitive of the entire superfamily. Using a 175-amino acid region covering most of the CAP domain and incorporating these signature sequences, 1281 CAP proteins in 427 species can be identified using Pfam (13). CAP superfamily proteins typically, but not always, contain a signal sequence directing proteins to specific intracellular compartments or more typically to the extracellular environment, where their disulfide-stabilized structure is presumably important for overall stability. Within this extracellular environment, functions are proposed to be either autocrine or paracrine. In the most broad sense, these involve regulating signal transduction events, *e.g.*, via ion channel gating or receptor activation, or modifying other proteins, for example in association with extracellular matrix remodeling or cellular adhesion.

Although a large number of proteins within the superfamily contain a CAP domain in isolation (*e.g.*, the Ag5 and Pr-1 proteins found in insects and plants), many others contain additional C-terminal extensions. With such a large superfamily, identification and classification of sub-

families has been difficult and often confusing; however, it is becoming apparent that many groupings can be defined based upon similarity in the CAP domain and also the nature of any C-terminal extensions. A schematic overview of the mammalian CAP superfamily architecture is provided in Fig. 1. Within published literature, nomenclature has frequently been loosely applied when naming new proteins in this superfamily, resulting in confusion. Herein we have endeavored to clarify, and simplify, CAP nomenclature. In some cases, we have recommended name changes.

Those proteins that contain the CAP domain in isolation typically terminate shortly after the CAP2 motif. The plant Pr-1, terminal amino acids are a PY dipeptide; and in the Ag5, a similar motif, [ILVP]Y, is very near to the terminus. The PY dipeptide is conserved among all mammalian CAPs and is identified shortly after the CAP2 motif. Given the largely conserved nature of this dipeptide throughout the superfamily, and given that it terminates the CAP domain in a number of cases, this may be considered the termination of the domain throughout the entire superfamily. This general rule is used throughout this review.

As summarized in Fig. 1, each of the CAPs identified in mammals (with the exception of GLIPR2) includes a C-terminal extension. Most have at least a short C-terminal extension containing four cysteines, which has been referred to as a Hinge region (14). Additional domains are sometimes entirely unique and are observed only in association with the CAP domain, *e.g.*, the ion channel regulator (ICR) at the C terminus of the CRISPs (15); some display homology to domains of other protein families, such as the C-type lectin (CTL) domain (Section IV.C) (16) and the LCCL domain (Section IV.D) (17); and others have C-terminal extensions with no homology to other sequences. This multidomain architecture is suggestive of multifunctional proteins. This interpretation is supported by several lines of evidence as detailed throughout this review.

Database analyses that we have undertaken show that at least 43 different protein domain architectures are present for the CAP superfamily across the three major phylogenetic divisions of Archaea, Bacteria and Eukaryota. Within each of these divisions, there are 8, 213, and 216 species, respectively, that are known to contain CAP proteins. Within the Bacteria, the majority of sequences reside within the Proteobacteria (138 sequences) and Firmicutes (132). Within the Eukaryota, the majority of sequences reside within the Viridiplantae (183) and Metazoa (529), primarily distributed between Mollusca, Nematoda (109), Arthropoda (flying winged insects of the Neoptera) (216) and Chordata (170). Within the Chordata, proteins have been primarily identified from squamates (snakes and lizards) (49), primates (36), and rodents (41). The focus of the current review is the CAPs of chordate origin, especially those found within mammals. Some CAPs of nonmammalian origin will, however, be described because they provide a useful context for future investigation.

III. Mammalian CAP Superfamily Proteins

A. Cysteine-rich secretory proteins (CRISPs) subfamily

The mammalian cysteine-rich secretory proteins (CRISPs) were first identified after characterization of major androgen-regulated proteins in rat epididymal fluid. The protein now called CRISP1 was initially identified as two closely migrating proteins, D and E, on a nonreducing and nonreducing protein gel (18). The same protein was subsequently identified as acidic epididymal glycoprotein (19), sialoprotein (20), 32-kDa epididymal protein (21), and protein IV (22). Proteins D and E were later shown to be two variants produced from the same gene (23, 24). After cloning and sequencing of the mouse gene, it was named CRISP1 based upon the high cysteine content, the predicted N-terminal signal sequence, and its extracellular localization (25). Orthologs of CRISP1 have been identified in the human, horse (as AEG1) (26), and mouse (as MEP7) (27).

Within all mammalian species investigated, several CRISP1 paralogs have been identified. CRISP2 was originally identified in the guinea pig as auto-antigen 1 (28, 29), in the mouse as TPX1 (2, 30), and in the rat (31, 32), human (33), and horse (34). CRISP3 was originally identified in the mouse (25), followed by the human (also called specific granule protein 28) (33, 35), and the horse (also called horse seminal plasma protein-3) (34, 36, 37). CRISP4 has been identified only in the mouse and the rat (38, 39).

Expressed sequence tags (ESTs) and genome sequences show that CRISPs are found in many other species including the chimpanzee, cow, pig, dog, chicken, opossum, and platypus (G. M. Gibbs, unpublished data). Owing to their abundance in the male reproductive tract of all of the species studied to date, CRISPs have gained much attention as potential regulators of male fertility.

CRISPs are two-domain proteins with an evolutionary diverse and structurally conserved N-terminal CAP domain of approximately 21 kDa, which contains each of the CAP motifs, and a C-terminal cysteine-rich domain of approximately 6 kDa, which is comprised of a Hinge and an ICR region (Fig. 1). The presence of two structural domains in CRISPs was first indicated by partial disulfide bond mapping of CRISP1 (40), but was formally confirmed after publication of the crystal structure of several snake CRISPs (14, 41, 42) (Section V). The C-terminal extension contains 10 strictly conserved cysteines, which form five disulfide bonds, and is the defining element of the CRISP subfamily. The C-terminal domain of CRISPs is often referred to as the cysteine-rich domain; however, the wealth of other nonrelated but similarly named domains leads to confusion—hence our proposed CRISP nomenclature (15). Throughout this review we refer to the C-terminal extension as the CRISP domain when we consider both the Hinge and the ICR together. The spacing of the cysteines in the Hinge region is different among each of the mammalian CAP subfamilies. The cysteine spacing of the Hinge region in the CRISPs is C_{x2}C_{x3}C_{x4}C.

CRISPs are only found in vertebrates and have been reported in many mammals, reptiles, *Xenopus* frogs, and the parasitic lamprey. All CRISPs contain a predicted signal peptide consistent with their extracellular localization or their

localization to specific intracellular compartments. Although they do not contain transmembrane domains, they are sometimes found associated with membranes potentially either through glycosylation or through interactions with integral membrane proteins. All mammalian CRISPs have a molecular mass of between 25 and 32 kDa, depending upon their glycosylation status or other posttranslational modifications.

Aspects of the biological function of mammalian CRISPs are starting to become apparent. Notably, they are likely to be dual function proteins with an activity associated with both the N-terminal CAP domain and the C-terminal CRISP domain. For example, the N-terminal CAP domain has been implicated in fusion to the oocyte (Section V.C), whereas the C-terminal domain has an ion channel regulatory activity (Section IV.B). In particular, for the former of these functions, analysis of the CAP superfamily and CAP domain containing proteins is useful. CRISPs are expressed in numerous locations throughout the body, and depending upon expression context, their biological functions may differ. The expression context and biochemistry of CRISPs are described throughout this section, and the biological functions are the focus of Sections V–VIII.

1. Mammalian CRISP orthologs and paralogs. Genomic data show that most mammals contain three CRISPs, with the mouse being the only one to date shown to contain four homologous genes. Genes are typically clustered and syntenic between species (reviewed in Ref. 43). Human *CRISP1*, *CRISP2*, and *CRISP3* are localized to chromosome 6p21.3; rat *Crisp1*, *Crisp2*, and *Crisp4* are localized to chromosome 9q12; and mouse *Crisp1*, *Crisp2*, and *Crisp3* are localized to chromosome 17B2, whereas *Crisp4* is located on chromosome 1A3 in a locus rich in other CAP superfamily genes (*Glpr1*, *Glpr111*, *Glpr112*).

A phylogenetic analysis of 99 CAP protein sequences yields a well-resolved consensus tree defining the major mammalian CAP protein clades (Fig. 1B). In this tree, all CRISP proteins cluster together in a single clade composed of two major subclades that receive high node credibility by our analyses. One subclade consists of CRISP1 orthologs plus rodent CRISP4, whereas the other combines the CRISP2 and CRISP3 orthologs. A closer examination of the phylogenetic structure of the CRISP clade (Fig. 2) highlights the current nomenclature inconsistencies between species and between subfamilies (or subclades). The nomenclatural inconsistency for CRISP1, CRISP3, and CRISP4 had its origin in the earliest period of CRISP research and developed through a combination of incomplete characterizations of the *Crisp* genes in the mouse and rat and overlapping expression profiles of CRISPs within the epididymis of the rat and the human. Although phylogenetic differences were recognized between “apparent” CRISP1 orthologs in the rat and human (44), it was not until the identification of mouse and rat CRISP4 (38, 39) that the genetic picture was completed and the true ortholog of human CRISP1 was identified as rodent CRISP4 (39). The long-stem branch preceding the CRISP2/CRISP3 subclade (Fig. 2) indicates that CRISP2 and CRISP3 are closely related paralogs. The clustering of equine and canine CRISP3 with the CRISP2 group (Fig. 2) may be a phylogenetic artifact caused by the overall low divergence between

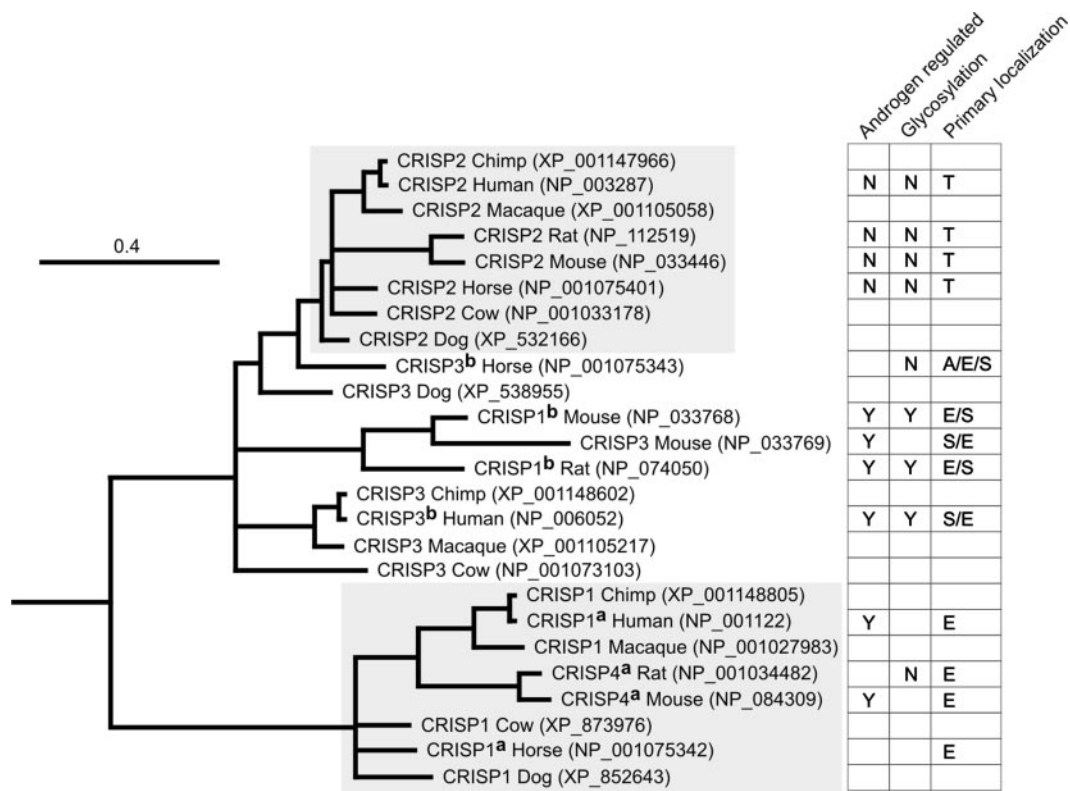


FIG. 2. Phylogenetic relationships of CRISP homologs from human, macaque, chimpanzee, dog, horse, cow, rat, and mouse, combined with a summary of regulated gene expression via androgens, protein modification, and predominant sites of protein localization. The tree is extracted from the complete mammalian CAP superfamily tree (supplementary Fig. 1). The monophyly of two groups of orthologs is highly supported (gray boxes), with less resolution among the CRISP3 orthologs (unboxed). The greater variability in tree structure among CRISP3 compared with earlier analyses (39) may be related to the greater number of sequences included in this analysis. The superscript “a” and superscript “b” on horse, human, mouse, and rat CRISP1, CRISP3, and CRISP4 highlight the phylogenetic similarity and nomenclature inconsistency as discussed in the text. E, Epididymis; S, salivary duct; T, testis; A, ampulla of the vas deferens; N, no; Y, yes.

CRISP2 and CRISP3 proteins and was partially resolved when alternative likelihood methods were applied (K. Rolants, unpublished data). The CRISP3 orthologs appear to have the most sequence variation, and this is represented in the phylogenetic tree. CRISP2 forms a well-resolved clade and is the most clearly defined in terms of interspecies nomenclature. Mouse and rat CRISP1 cluster with human CRISP3, providing supporting evidence that rodent CRISP1 is the ortholog of human CRISP3. For simplicity, within the following sections these subdivisions are referred to as the CRISP1 group, the CRISP2 group, and the CRISP3 group, despite individual instances of nomenclature inconsistency. We considered attempting to rename the CRISPs in this review so that the nomenclature accurately reflected the phylogenetic classification, but felt that this would lead to confusion in the already extensive CRISP literature.

Conservation of expression profiles and protein modification are an additional measure of true protein orthologs. Summarized in Fig. 2 is the androgen-regulated expression, posttranslational modifications, and primary sites of expression of CRISPs. These show that the CRISP2 proteins are not androgen-regulated, are not modified by glycosylation, and have a primary site of expression within the testis and male germ cells. The CRISP1 group and the CRISP3 group display overlapping expression profiles, in particular within the epididymis. Their expression is up-regulated by androgens,

and they have a tendency toward glycosylation. Variation between these two subgroups is evident in their expression profiles, with the CRISP1 group having a narrower expression profile largely restricted to the epididymis in the male reproductive tract. The CRISP3 group, although still enriched in the reproductive tract, is widely expressed, including within the salivary duct, immune tissues, and exocrine glands. Although this restricted expression profile has been the dogma for CRISPs, a recent analysis by our laboratory has shown that all CRISPs in the mouse have a wider expression profile than previously published. This report showed that CRISP1, CRISP2, and CRISP4 were also expressed in a wider range of reproductive and nonreproductive tissues, generally from secretory epithelia and immune tissues (45).

2. *CRISPs expression, localization, and posttranslational modifications.* Rather than reviewing mammalian CRISP expression profile according to nomenclature classification, this summary of CRISP expression is based on tissue distribution. Given the high likelihood of functional redundancy between family members, we believe this organization will assist in functional predictions and overcome some of the difficulties associated with the nomenclature inconsistencies.

3. *Testis.* Germ cells within the testis undergo successive mitotic and meiotic divisions followed by a series of mor-

phogenic processes, collectively called spermiogenesis, to form the highly specialized spermatozoa (46). The differentiated spermatozoa is comprised of several subcellular compartments and can be most simply divided into the head and the tail. The head contains a highly condensed nucleus and an anteriorly positioned lytic vesicle called the acrosome. Lysis of the acrosome (a process called the acrosome reaction) is required to achieve fertilization (47). The sperm tail, which is required for motility, is divided into three regions: the midpiece, principal piece, and end piece. In addition to the axoneme, which all ciliated cells possess, sperm tails have unique ultrastructural components, notably, the outer dense fibers and a fibrous sheath (48, 49).

Based on published data, *Crisp2* is the only CRISP normally expressed in the mammalian testis. Its expression is not regulated by androgens, and CRISP2 protein is not modified by glycosylation (28, 50). CRISP2 has been partially characterized in the guinea pig, mouse, rat, horse, and human. The bulk of the data suggest that CRISP2 is first produced in spermatids; however, its mRNA may be expressed several days earlier, then subjected to translational delay (28, 32, 51–53) via binding to the RNA binding protein DAZL (54). CRISP2 is incorporated into the acrosome contents and is visible from early steps in round spermatid development (28, 29, 51, 53, 55, 56). Based on protein localization studies largely in the rat, CRISP2 is also incorporated into the developing sperm connecting piece (or neck) and accessory structures of the sperm tail (51). Incorporation of CRISP2 into the sperm tail is likely to occur via the granulated bodies (51). Granulated bodies are believed to be protein transport vesicles required for the transfer of proteins from the cytoplasmic lobe of elongating spermatids into the developing sperm tail (57). Within the acrosome of guinea pig sperm, CRISP2 is estimated to be 6.4% of the total protein content (28).

Although murine *Crisp2* expression is clearly abundant within the testis, low levels of cDNA have also been amplified from purified secretory acinar cells from the mouse lacrimal gland (tear duct) (58), and cDNA and protein have been observed in the mouse ovary, lung, and skeletal muscle (45).

4. Epididymal production and sperm localization. Fully differentiated testicular sperm are not capable of fertilizing an oocyte and are required to undergo a period of epididymal maturation to achieve fertilization competence. The epididymis is comprised of several regions, most simply divided into the initial segment, the caput, corpus, and cauda. During passage through the initial segment, the caput and corpus, sperm acquire the ability to fertilize. Sperm are subsequently stored within the cauda epididymis in an inactivated, yet fully functional state. Epididymal maturation is characterized by a dramatic remodeling of the sperm cell, including changes in both the membrane lipid and protein component. Many proteins produced by the epididymis are secreted into the epididymal lumen where they both surround the sperm and are transferred to the sperm membrane.

Mouse, rat, human, and horse epididymis contains several CRISPs, typically from the CRISP1 and CRISP3 clades (Fig. 2). They show partially overlapping expression profiles (25, 33, 37, 39, 59) and, as indicated by the recent knockout of

Crisp1 in the mouse which had normal fertility (60), most likely share some degree of functional redundancy. The epithelium of humans and the horse epididymis secrete abundant CRISP1^a and CRISP3^b into the epididymal lumen (33, 37, 59), whereas rodents secrete abundant CRISP1^b and CRISP4^a (39). The *superscript* following each of these proteins is representative of their phylogenetic group (Fig. 2). Although the nomenclature would indicate different expression profiles between humans and rodents, there is in fact conservation based upon phylogenetic grouping. It is likely that each are functional counterparts whose biological role is determined by posttranslational modifications and the nature of the interaction with sperm. Based on mouse EST analyses, *Crisp3* also appears to be expressed in the epididymis, although, unlike *Crisp1* and *Crisp4*, this is not the primary tissue of production.

Human CRISP1 protein has been detected in all regions of the epididymis, vas deferens, and seminal plasma, and its expression is predominant in the male reproductive tract (33, 39, 44). CRISP1 from the epididymal epithelium becomes enriched within epididymosomes (61, 62), from whence it is transferred, at least in part, to the postacrosomal region of the sperm head (44). The interaction between CRISP1 and sperm is facilitated through N-linked glycosylation, which contributes approximately 13% to the total molecular mass (or 4 kDa) (44). CRISP1 associated with ejaculated sperm can be washed off under mild conditions, suggesting that it interacts weakly with sperm and that a significant proportion remains in the fluid component of the epididymal lumen (33). CRISP1 has been estimated to be present at a concentration of between 8 and 14 $\mu\text{g}/\text{ml}$ (or 0.15–0.25 $\mu\text{g}/\text{mg}$ of total seminal plasma protein) (33).

CRISP3 is also produced by the secretory epithelium of the human epididymis. It is present at highest concentrations in the cauda epididymal lumen and vas deferens (4.6 $\mu\text{g}/\text{mg}$ total protein) (59). CRISP3 exists in two forms (+/– glycosylation), both of which remain associated with human sperm after analysis of the washed highly motile “swim up” fraction (59). CRISP3 is present in seminal plasma at 14.8 $\mu\text{g}/\text{ml}$ with the majority originating from the accessory sex organs, *i.e.*, the ampulla of the vas deferens and the prostate (59).

Within the mouse, *Crisp4* expression and protein localization were observed from 30 d after birth in the principal cells in the epididymal epithelium. Expression levels increased up to d 60 in an androgen-regulated manner (38). Within the rat, *Crisp4* is most highly expressed in proximal regions of the epididymis, *i.e.*, highest levels were observed in the initial segment, segments III–VI of the caput epididymis, and in the corpus epididymis. Although much lower levels of expression were seen in the cauda epididymis (38), the secreted protein accumulates within the corpus and cauda where it attaches to the maturing sperm (39). A similar distribution was observed in the mouse epididymis (39). The majority of the CRISP4 protein, however, is released from sperm under nondetergent extraction conditions, suggesting primarily a nonintegral membrane association (39). Low expression of murine *Crisp4* has also been detected in the oocyte (39), skeletal muscle, spleen, and thymus (45).

Similar to rodent *Crisp4*, rodent *Crisp1* is found in the highest levels in the epididymis (39). In comparison to *Crisp4*, however, *Crisp1* is expressed more distally, and the highest levels of expression were observed in the cauda epididymis (39). *Crisp1* expression is androgen-regulated (50), consistent with the identification of androgen response elements within the promoter (63). The bulk of rodent CRISP1 is produced and secreted from the principal cells of the proximal corpus epididymis through to the cauda, where it can be identified as a 29-kDa protein and a 31-kDa protein with three major and one minor charge variants that are suggestive of differential glycosylation (27, 39, 40, 64).

Within the mouse, CRISP1 comprises 15% of the total cauda epididymal fluid protein content (12.4 $\mu\text{g}/\text{mg}$ of cauda protein) (40). Within the rat, CRISP1 reaches maximal concentrations in the cauda epididymal lumen (19, 65) and is estimated to be present at approximately 15.5 $\mu\text{g}/\text{mg}$ of total protein (20, 66). Such a high concentration is consistent with the proposed role of CRISP1 as a decapacitation factor involved in maintaining sperm in a quiescent state before ejaculation (see Section V.B) (67).

Rat epididymal CRISP1 exists as two isoforms, the smaller D and larger E isoform with an approximate difference of 2 kDa in molecular mass. The smaller D form is synthesized from the principal cells of the caput epididymis and in all subsequent regions of the epididymis. The larger E form is synthesized in the corpus and proximal cauda (64, 68, 69) and is quantitatively the lesser fraction (65, 70). Both forms of rat CRISP1 have been purified to homogeneity from the epididymal fluid, and sequencing of an internal 34-amino acid peptide supports their uniform identity (71). Both isoforms had a blocked N terminus (70), which was likely the result of cleavage of the predicted signal peptide after secretion from the epididymal epithelium. Tryptic peptide maps from each isoform showed a few differences; however, direct sequencing of those peptides failed, suggesting that a distinction between the isoforms resides with the blocked N-terminal tryptic fragment.

Considerable research has gone into determining the type of CRISP1 associated with sperm and the dynamics of its redistribution during capacitation and fertilization. The post-translational differences between the D and E forms has significant implication in this redistribution. Although these studies have been complicated by variations in techniques and analysis tools, a clear picture is beginning to emerge. It is possible to discriminate the D and E isoforms of CRISP1 using the 4E9 monoclonal antibody that binds to a partially characterized epitope on the N terminus of the E form (71). Although both the D and E isoforms are glycosylated (in rat CRISP1 this corresponds to between 6 and 7.5% of its total mass or ~ 2 kDa) (65, 72), this is not the basis for the molecular mass difference (70). Splice variants of rat *Crisp1* have been identified; however, these do not contribute to additional coding sequence in the N terminus (73, 74). Protease and cyanogen bromide cleavage of CRISP1 indicates that the 4E9 epitope involves both additional protein sequence and the presence of an O-linked N-acetyl galactosamine epitope of 204 Da (43, 70, 75).

In the rodent, most sperm bound CRISP1 is loosely associated with the plasma membrane and is easily removed (64,

76). A smaller population, the E form, is more strongly attached to the membrane and remains associated with the sperm head after maturation within the female reproductive tract (27, 64). This more strongly attached population has been proposed to have a role in the fusion to the oocyte and is described in more detail below in Section V.C.

Immunofluorescence staining of sperm from the caput epididymis showed generalized CRISP1 staining over the entire sperm surface, which becomes localized to the mid-piece and the head after epididymal maturation and entry into the cauda epididymis (65). As indicated, within epididymal sperm the D isoform appears to be glycosyl-phosphatidylinositol (GPI) anchored (65) and is seen associated loosely with the dorsal region of the sperm head in fixed sperm (64, 77). The E isoform undergoes C-terminal proteolytic truncation during epididymal transit to create progressively smaller versions, with the smallest of these (26 kDa) associating very strongly with the membrane over the sperm tail (64). The inability of the 11D4 monoclonal antibody, which binds to the 11 most C-terminal amino acids of full length CRISP1 (64), failed to detect the truncated forms on the sperm tail, supporting the notion of proteolysis. This proteolysis interpretation is yet to be confirmed.

Many reports have described the redistribution of CRISP1 D isoform from the dorsal head region to the equatorial segment during capacitation (77, 78). Indirect immunofluorescence on sperm using the 4E9 antibody suggests that a small proportion of the CRISP1 E isoform redistributes to the equatorial segment on the sperm head after *in vitro* capacitation (Section V.C) (compare Fig. 1 in Ref. 64 with Fig. 7 in Ref. 67) and that the total CRISP1 pool (as detected by antibody CAP-A) on the sperm membrane is reduced during capacitation (67). Given, the question of specificity of polyclonal antibodies raised against almost identical D and E isoforms; the relative abundance of the D isoform within the epididymal lumen and on fresh cauda sperm; that the D isoform is not tightly bound to the membrane; and that the E isoform appears to undergo capacitation-associated redistribution to the equatorial segment, the possibility exists that on capacitated sperm the E isoform becomes preferentially detected as the D isoform is depleted (Section V.C)—hence, the appearance of translocation from dorsal to equatorial segment of the D isoform. Although several outstanding questions remain, not least of which is the nature of the truncated E isoforms, this hypothesis is consistent with numerous aspects of the functional models (Section V.C).

5. *CRISP expression in the accessory glands of the male reproductive tract.* As indicated above, CRISP3 is produced by the normal human vas deferens and prostate and contributes to the high levels observed in seminal plasma (14.8 $\mu\text{g}/\text{ml}$) (59). *Crisp3* is also produced in the horse seminal vesicles and is thus likely to contribute to seminal plasma levels in other species (37). Immunoprecipitation and gel filtration experiments have shown that within human seminal plasma, CRISP3 binds to dimeric β -microseminoprotein (β -MSP) with 1:1 stoichiometry and with a K_D of approximately 6.5×10^{-11} M, suggesting a highly specific interaction (79). As the function of β -MSP [also called prostate secretory protein 94

(PSP94)] is not clear, the biological significance of this interaction is not yet apparent. There is, however a significant excess of β -MSP in seminal plasma, meaning that it is likely that all the CRISP3 is complexed. This interaction has been proposed to occur via the CAP domain because protease-inhibitor 16 (a non-CRISP CAP) that is found in human serum plasma also interacted with β -MSP (80) (Section III.E).

Within stallion seminal plasma, CRISP3 is not glycosylated and is present as a monomer in solution and at concentrations up to 1.3 mg/ml. It has been estimated that after removal of seminal plasma, 0.9 to 9 million molecules of CRISP3 remain associated with each stallion sperm on the postacrosomal and equatorial segment and that they may have a function related to oocyte binding (36, 37) in a manner conceptually similar to that proposed for CRISP1 (reviewed in Ref. 76) (Section V.C). To date, this role has not been explored.

6. CRISPs in the immune system. *Crisp3* has a broader expression distribution than other mammalian CRISPs. Within the nondiseased animal, the expression of *Crisp3* cDNA is suggestive of a role in the innate immune system (35, 81). Specifically, CRISP3 has been detected in the secretions from several exocrine glands onto mucosal surfaces, including the lacrimal gland, submandibular (salivary) gland, prostate, and pancreas (25, 33, 58); and within several cells of the immune system including pre-B cells, neutrophils, and eosinophils (81, 82); in the spleen where its expression is regulated by Oct2 (83); and in the thymus (33). The potential role of CRISP3 in the immune system is described in more detail in Section VII. CRISP3 is estimated to be present at 6.3 μ g/ml in human plasma, 21.8 μ g/ml in saliva, 11.2 μ g/ml in seminal plasma, and 0.15 μ g/ml in sweat (82). We showed that, CRISP3 mRNA and protein is also found in skeletal muscle, uterus, mammary gland, and heart (45), but such expression is unlikely to be related to immune function.

Within the mouse submandibular gland, the developmental expression profile of *Crisp3* shows a distinct increase (approximately 6-fold) relative to *Crisp1* from birth up to 20 d of age (84). Consistent with the presence of an androgen-response element in both mouse *Crisp1* and *Crisp3* promoters (85), the absolute level of both mRNAs increased within the granular convoluted tubules of the male submandibular gland from d 30, concordant with increased androgen levels (25, 84). *Crisp3* expression has also been observed in equine submandibular glands (37).

Within neutrophils, CRISP3 is present as two populations including an N-linked glycosylated form (predicted to be at Asn239) with an estimated molecular mass of 29 kDa, and a nonmodified form of 27 kDa (82). Within human plasma, presumably as a product of neutrophil and eosinophil secretory granules, CRISP3 binds to α 1B-glycoprotein (α 1BG) with a 1:1 molar ratio and a K_D in the nanomolar range, suggestive of a highly specific interaction (86). This interaction was not metal ion dependent and has been proposed to inactivate the potentially toxic effect of CRISP3. Such an activity would be analogous to the inactivation of snake venom metalloproteases and myotoxins by α 1BG-like proteins in opossum plasma (87–90). Based on this thesis, CRISP3 may normally be kept in a bound-inactive state until

it reaches, or is delivered to, its endogenous target. Unfortunately, because the activity of CRISP3 is, as yet, unknown, this hypothesis is difficult to test.

It is of note that *CRISP3* expression is dramatically up-regulated in a number of pathologies including in prostate cancer (91–93), squamous carcinoma of the tongue (94), chronic pancreatitis (95), and Sjögren's syndrome (96–98). The potential clinical significance of these changes is described in Sections VII–VIII.

B. Glioma pathogenesis related-1 (GLIPR1) subfamily

The second most well-characterized CAP subfamily in mammals consists of the GLIPR1 proteins. GLIPR1 proteins are also called related to testis specific, vespid and pathogenesis related-1 (RTVP1) proteins. Similar to the CRISPs, the GLIPR1 proteins are a multigene subfamily consisting of three genes in most species, with four identified within the mouse. Each GLIPR1 component group has distinct differences at their C termini (Fig. 1A). Our phylogenetic analysis shows that mammalian GLIPR1 proteins form a single well-supported clade composed of three distinct subclades (Fig. 1B): GLIPR1, GLIPR1-like 1 (GLIPR1L1), and GLIPR1-like 2 (GLIPR1L2). GLIPR1-like 3 (GLIPR1L3) is unique to the mouse and appears to be a duplication of *Glpr1l1* (99). GLIPR1L3 has a high degree of sequence identity with GLIPR1L1 (82%) and groups within the GLIPR1L1 clade after phylogenetic analysis (see supplementary Fig. 1, published as supplemental data on The Endocrine Society's Journals Online web site at <http://edrv.endojournals.org>, and Ref. 325 cited therein). Available expression data suggest distinct differences in expression profiles of each of these groups. Each of the GLIPR1 subfamily members contains the four CAP signature sequence motifs and the Hinge-like sequence with a cysteine spacing of Cx₂Cx₅Cx₄C. Uniquely among the mammalian CAP superfamily, GLIPR1 and GLIPR1L2 proteins are predicted to have C-terminal transmembrane domains (99), which may anchor the extracellular CAP domain to membranes. There are some localization data to support this prediction (99). GLIPR1L2 has a long glutamate-rich domain (99), the function of which remains unknown.

The first *GLIPR1* gene was identified after an examination of the major up-regulated genes within the most aggressive form of human brain cancer, glioblastoma multiforme/astrocytoma, and within glioma cell lines (100, 101). Transcripts were not identified in other neuronal cancer cell lines or within the normal brain, suggesting that *GLIPR1* expression was derived from the cell from which the cancer originated and, as such, may be important for pathogenesis. The same gene (*RTVP-1*) was later identified in prostate cancer cell lines and shown to be p53-regulated and proapoptotic (102). Epigenetic down-regulation of *GLIPR1* has been observed within the cancerous prostate (103). The role of GLIPR1 in cancer is discussed in Section VI.

Human *GLIPR1*, *GLIPR1L1*, and *GLIPR1L2* genes are clustered on chromosome 12q21, and the mouse genes are localized on chromosome 10D1 (99). Each of the human and mouse genes has functional p53 response elements (99, 102). Genomic analyses identified likely orthologs of all GLIPR1, GLIPR1L1, and GLIPR1L2 within many taxa ranging from

teleost fish to the chicken and other mammals (G. M. Gibbs, unpublished data).

Human GLIPR1 expression is observed in the fetal kidney and in multiple adult tissues (99, 104). Quantitative RT-PCR showed that *GLIPR1* expression was highest in the human lung, followed by the testis, bone marrow, prostate, bladder, and kidney (99). Similarly, *GLIPR1L2* has a wide expression profile with highest levels in the testis and lower levels in kidney, prostate, lung, bladder, and bone marrow (99). *GLIPR1L1* is almost exclusively expressed in the testis, with trace amounts in the bladder (99). The cellular localization of *GLIPR1L1* and *GLIPR1L2* in humans is currently unknown. Similarly, a formal analysis of murine expression profiles has not been completed; however, EST expression data are consistent with that reported for the human. Splice variants of *GLIPR1* are observed, and variations in expression profile are reported; however, it is not apparent how these may alter biological function (99, 105).

C. Glioma pathogenesis related-2 (GLIPR2)/Golgi-associated pathogenesis related-1 (GAPR-1) subfamily

The *GLIPR2* subfamily possesses several unique characteristics among the mammalian CAP proteins. Although only two *GLIPR2* proteins have been characterized (human and mouse), several unique characteristics suggest that this clade may contain the most primitive mammalian CAPs. This interpretation is supported by our phylogenetic analysis, showing the *GLIPR2* clade to be the earliest diverged CAP subfamily within mammals, which was rooted with the protein PRY1 from the yeast *Saccharomyces cerevisiae* (106) (Fig. 1B). Despite the similarity in nomenclature between *GLIPR1* and *GLIPR2* proteins, they do not display a particularly close phylogenetic relationship within the CAP superfamily (Fig. 1B).

Genomic analysis showed *GLIPR2* orthologs to be present in the widest array of species of all those in the mammalian CAP superfamily (G. M. Gibbs, unpublished data). *GLIPR2* and paralogs can be seen within *S. cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, insects, ascidians, bony fish, amphibians, birds, and mammals. Thirty-six genes have been identified, with 24 in mammals, including the human, chimpanzee, monkey, mouse, rat, dog, and cow. *GLIPR2* is localized on human chromosome 9p12-p13, with orthologous genes localized in syntenic regions in other species.

An analysis of mammalian *GLIPR2* protein sequences shows that each of the CAP sequence motifs is present, but they contain a greater degree of sequence variability than in other subfamilies (G. M. Gibbs, unpublished observation). Uniquely, *GLIPR2* proteins do not contain a predicted signal sequence. This is consistent with its intracellular localization to the Golgi membrane (107). In addition, the conserved PY dipeptide sequence present in most other CAP superfamily proteins is absent, and the Hinge-like sequence present in all other mammalian CAPs is absent from *GLIPR2* (Fig. 1A). In this latter respect, *GLIPR2* is more similar to many nonmammalian CAPs. *GLIPR2* proteins have the greatest degree of identity with the plant Pr-1 proteins (Section VII.C). This, along with their abundance in immune environments, supports the notion that *GLIPR2* proteins may

represent an evolutionary link between the plant and mammalian immune systems, a relationship that has been considered on several occasions but not yet demonstrated (100, 107).

The *GLIPR2* gene (originally called *C9orf19*) was identified in the human and has a wide expression profile. Highest expression levels were seen in the heart, lung, and peripheral blood leukocytes. Lower levels were detected in the skeletal muscle, prostate, and uterus (108). At a similar time, the same protein (originally called GAPR-1) was purified from the lipid-enriched microdomains of CHO cells Golgi complexes, where its myristoylated N terminus was shown to be important for interaction to the cytoplasmic side of the Golgi membrane (107). Weak interaction with caveolin-1 supported enrichment to lipid raft domains.

Experimental analysis showed that *GLIPR2* is predominantly localized in tissues with immunological functions including circulating and developing leukocytes in the spleen. High levels of *GLIPR2* have also been reported in the kidney, pancreas, lung, uterus, embryo, and placenta. Lower levels were apparent in the brain, testis, and skeletal muscle (107).

More recently, *GLIPR2* was identified in the mouse as an up-regulated protein in kidney fibrosis (109). Fibrosis is an accumulation of extracellular matrix that ultimately may replace normal tissue and often results in organ failure (110). Fibrosis has a complex etiology, which is often associated with an accumulation of myofibroblasts producing increased extracellular matrix (111, 112). Myofibroblasts arise from several sources, one of which is the transdifferentiation of local epithelial cells in a process called the epithelial to mesenchymal transition (113). In the normal kidney, *GLIPR2* is secreted (potentially through a non-ER secretory pathway) to the collecting ducts of the medulla. However, within the fibrotic kidney *GLIPR2* was observed within epithelial cells of proximal tubules and within the damaged glomeruli, adjacent to activated myofibroblasts, and directly apposed to zones of peritubular fibrosis. The addition of exogenous *GLIPR2* to cultured renal epithelial cells caused their transformation to a fibroblast-like morphology, supporting the notion that *GLIPR2* is involved in epithelial to mesenchymal transition and in generating a pool of myofibroblasts contributing to fibrosis (109).

D. Peptidase inhibitor 15 (PI15) subfamily

The *PI15* subfamily of CAP proteins was first identified from cultured human glioblastoma cells as a novel 25-kDa trypsin binding protein (p25TI) with weak inhibitory activity (114). Genomic analysis shows *PI15* orthologs in 29 species conserved throughout the Euteleostomi (bony vertebrates), with 23 identified in mammals including the human, chimpanzee, monkey, cow, horse, dog, mouse, and rat (G. M. Gibbs, unpublished data). Human *PI15* is situated on chromosome 8q21.11 and is adjacent to another mammalian CAP superfamily gene, the *CRISPLD1* (Section III.F). Similarly, orthologous *PI15* genes in six mammalian species are localized adjacent to *CRISPLD1* (Section III.F). Our phylogenetic analyses cluster the *PI15* proteins as the sister clade of the *CRISPLD1*, *CRISPLD2*, and *R3HDML* subfamilies (Fig. 1B and supplementary Fig. 1).

An analysis of seven mammalian PI15 sequences shows a minimum of 90% identity over the entire length of the 258 amino acid proteins (G. M. Gibbs, unpublished data). All PI15 subfamily members contain each of the expected CAP motifs. Structurally, they contain a CAP domain terminating with the conserved PY dipeptide, a Hinge-like sequence with cysteine spacing of C_x₂C_x₇C_x₄C, followed by an additional short stretch of 12 amino acids (Fig. 1A and supplementary Fig. 2). N-terminal sequencing of cell culture-derived PI15 showed that the predicted signal peptide was active. Human PI15 appears to be produced as a propeptide and is cleaved at a furin-like protease cleavage site to produce a mature protein of 198 amino acids with one predicted glycosylation site (115).

Within the mouse, PI15 was identified as a marker of podocyte maturation during kidney glomerular development (116). Within the human, *PI15* has been detected within the brain, placenta, and lymphocytes (115). Screening of adult human and mouse tissue blots showed expression of *PI15* in a wide range of tissues including the prostate, mammary gland, salivary gland, thyroid gland, skeletal muscle, smooth muscle, heart, and ovary (117). Within the mouse and human, PI15 expression was observed within glandular structures including the prostate, mammary gland, salivary gland, and thyroid gland (117). This expression profile mirrors that recently reported for the murine CRISPs (45).

Similar to *GLIPR1*, human *PI15* is up-regulated within glioblastoma and neuroblastoma cell lines (115). The increased production of PI15 from cancer cell lines (115) and demonstration of a weak trypsin inhibitory activity (114) provide indirect support for a role in cancer pathogenesis, but this activity is yet to be confirmed, and an association between PI15 and cancer pathogenesis has not been established.

PI15 proteins have been identified in the pancreatic mesoderm of the chicken (originally called sugarcrist) (117). Chicken PI15 is expressed at precise periods during organogenesis suggestive of a role during development (117). Chicken PI15 was observed within the mesoderm of the emerging dorsal pancreatic bud, within the thyroid anlagen between the emerging lung buds, within the lung bud itself, and within the anterior and posterior regions of the limb buds (117). Organogenesis of branched organs requires signaling between mesenchymal and epithelial cells and precise modification of the extracellular matrix through inhibition and activation of numerous intracellular and extracellular factors, including proteases, protease inhibitors, and adhesion molecules (118, 119) (reviewed in Refs. 120 and 121). The localization of chicken PI15 (with 90% identity to human PI15) to regions of bud development is suggestive of a role in the controlled modification of extracellular matrix through the action of proteases and protease inhibitors and in branching morphogenesis. Similarly, rat CRISPLD2 identified during lung development is proposed to have a role in the regulation of extracellular matrix formation (Section III.G).

Mouse PI15 has been identified within the developing podocytes lining Bowman's capsule of the kidney glomerulus (116). *PI15* showed a distinct peak of expression during the S-shaped and capillary-looped stage in the embryonic day (E) 18.5 kidney. Expression of *PI15* mirrored the expres-

sion of the known transcription factor and podocyte marker, Forkhead box c2 (*Foxc2*), but expression of *PI15* was not regulated by *FOXC2* (116).

Although relatively little data are available on PI15 proteins, within multiple species and tissue environments PI15 proteins are secreted and localized to the extracellular matrix. Within this environment, they may have a role in the regulation of the modification of extracellular matrix.

E. Peptidase inhibitor 16 (*PI16*) subfamily

The peptidase inhibitor 16 (PI16) CAP subfamily is not well characterized, and there are only a few reports following their identification in human serum as a PSP94 binding protein (also called β -MSP) (80) and from a murine cardiac cDNA library (122). Although also referred to as cysteine-rich protease inhibitor (CRIPI) and CRISP9 through public databases, the family of orthologous genes was renamed PI16. It remains to be established whether the PI16 proteins have a protease inhibitory activity as reported for the human PI15 protein.

Our *in silico* genomic analysis identified numerous *PI16* orthologs conserved within the Eukaryota, where 26 species containing the *PI16* gene were identified (G. M. Gibbs, unpublished data). Human *PI16* is localized to chromosome 6p21.2 and to syntenic regions within other species including chromosome 17B1 in the mouse, chromosome 12:9.02 Mb in the dog, and chromosome 23:8.86 Mb in the cow.

Analysis of six mammalian protein sequences shows a high degree of identity, with a minimum of 60% (between the dog and the mouse) and is consistent with them forming a distinct subclade after phylogenetic analysis (Fig. 1B and supplementary Figs. 1 and 3). PI16 proteins are normally about 460 amino acids in length. The mouse and rat contain an insertion of about 20 amino acids. Each of the CAP motifs is present, but a distinct feature of this family is the spacing between CAP3 and CAP4, suggesting a major structural difference between this subfamily and all other CAP superfamily proteins (Fig. 1A). Following the conserved PY dipeptide, a Hinge-like sequence with cysteine spacing the same as the *GLIPR1* subfamily (C_x₂C_x₅C_x₄C) is present, followed by an additional approximately 264 amino acids with no homology to any other sequences and no recognized protein domains.

Expression of human *PI16* is observed within the testis, prostate, small intestine, colon, peripheral blood leukocytes, and the ovary with immunohistochemical analyses showing localization to specific cells within the fibromuscular stroma of the normal human prostate, the pituitary gland, parathyroid gland, tonsil, kidney, stomach, liver, and the Leydig cells within the testis (80). This wide expression profile is consistent with the National Center for Biotechnology Information Unigene expression profile. Analysis of serum levels of PI16 in prostate cancer patients showed a significant decrease (80), and it has been used as a prognostic marker of prostate cancer recurrence (123) (Section VII). Within human serum, PI16 is present as two molecular mass forms of 94 and 71 kDa, which reduce to 68 and 52 kDa, respectively, after deglycosylation (80). The nature of the 16-kDa difference in molecular mass is currently unclear.

Mouse PI16 (predicted to be 53 kDa) was identified as a secretory protein from a cardiac cDNA library, where it was shown to have a role in regulating cardiomyocyte size (122). PI16 was produced from a single transcript as three glycosylated forms of 74, 100, and 108 kDa (122). Deglycosylation did not eliminate the mass differences between the bands, suggesting additional posttranslational modification (122), paralleling observations for human PI16 and other CAP superfamily proteins. Within the mouse, PI16 is abundant in skin and aorta, with lower levels in adipose tissue (122). Immunohistochemical analysis of hearts showed barely detectable levels of PI16 within the cytoplasm, but protein had accumulated in the intercellular spaces, consistent with the secretory nature of the CAP superfamily proteins. PI16 is the second most highly up-regulated protein in cardiac myocytes in a mouse model of heart failure, second only to atrial natriuretic peptide (122) (*Section VIII.D*).

F. CRISPLD1/CAPLD1 subfamily

There are presently no primary publications describing the CRISPLD1 proteins. During our analysis of the mammalian CAP superfamily, we identified a large group of proteins displaying clear homology to the CAP superfamily with each of the normal conserved sequence motifs. Available EST data support their expression, and our phylogenetic sequence analysis supports their recognition as a distinct CAP subfamily (Fig. 1B).

Among the mammalian species used in the phylogenetic analysis, there is a minimum 92% sequence identity over the length of the CRISPLD1 proteins (supplementary Fig. 4). Our analysis of selected CRISPLD1 proteins showed that they are approximately 500 amino acids in length and contain each of the four CAP consensus motifs, the conserved PY dipeptide, and a Hinge-like sequence (Cx₂Cx₇Cx₄C), followed by an additional 260 amino acids containing tandem LCCL domains (*Section IV.D* and Fig. 1A). Analysis of expression profiles through Unigene shows a wide expression profile within the human and mouse with highest levels within the ear. In this regard, the presence of the LCCL domain may be significant because a mutation in LCCL domain has been associated with the deafness autosomal dominant nonsyndromic sensorineural disorder 9 (DFNA9) (17) (*Section IV.D*).

CRISPLD1 is also annotated as Cocacrisp, CRISP10, and LCRISP1 on public databases. Despite the prevalence of CRISP nomenclature within this CAP subfamily, members of the subfamily are clearly distinct from the CRISPs because they do not contain the 10 absolutely conserved cysteines or the ICR domain of the CRISPs. As such, we suggest a new nomenclature for this group: CAP and LCCL domain containing protein 1 (CAPLD1).

The CAPLD1 subfamily is conserved within the Euteleostomi (bony vertebrates) where 30 orthologs have been predicted, with 23 of these being in mammals including the human, chimpanzee, mouse, rat, dog, cow, and horse. Human CAPLD1 is localized to chromosome 8q21.11. Orthologous genes in other mammalian species are normally localized to syntenic regions. In each of those species listed above,

CAPLD1 is always within the same locus as PI15. Mouse *Crisp4* is also found within the same gene cluster.

G. CRISPLD2/CAPLD2 subfamily

The first member of this subfamily was identified in the rat as a glucocorticoid-inducible gene with a potential role in branching morphogenesis (124). Within this context, it was named late gestation lung protein 1 (Lgl1). Subsequently, it was named CRISP and LCCL domain containing protein 2 (CRISPLD2) to reflect its homology with the CRISPLD1/CAPLD1 subfamily (Fig. 1A). However, as outlined above, this subfamily of proteins is distinct to the CRISPs, and as such the CRISPLD1 nomenclature is inappropriate (Fig. 1B). As for the CRISPLD1 subfamily, we propose that the CRISPLD2 nomenclature be changed to CAP and LCCL domain containing protein 2 (CAPLD2) to more accurately reflect the phylogenetic structure of the mammalian CAPs.

Our genomic analysis of the orthologous CAPLD2 genes illustrated their conservation throughout the bony vertebrates, where 28 orthologs were identified. At least 22 mammalian species contain CAPLD2 genes, including the human, chimpanzee, mouse, rat, cow, dog, and horse. The human gene is localized to chromosome 16q24.1, and the mouse gene localized to the syntenic region at chromosome 8E1. EST expression profiles indicate that within human and mouse *Capld2* is widely expressed.

In terms of protein structure, CAPLD2 proteins are typically approximately 500 amino acids in length, with the exception of the horse and dog, which, assuming that the theoretical translation is correct, may have extended N termini. Mammalian CAPLD2s show a minimum of 74% identity, and each contains a CAP domain including each of the conserved CAP motifs, followed by the PY dipeptide and the Hinge-like sequence (Cx₂Cx₇Cx₄C) (Fig. 1A and supplementary Fig. 5). Similar to CAPLD1, CAPLD2s have a large C-terminal extension following the Hinge-like sequence. The extensions are about 260 amino acids in length and include tandem LCCL domains (*Section IV.D*). Our phylogenetic analysis shows that the CAPLD2 proteins form a distinct subfamily within the CAP superfamily, which is most closely related to the CAPLD1 and PI15 subfamilies (Fig. 1B and supplementary Fig. 1).

Rat *Capld2* was identified during a screen for glucocorticoid-induced transcripts in fetal d 20 lung fibroblasts (124). It was also shown to be expressed within the fetal heart, kidney, intestine, and within the adult lung, heart, and spleen. Within the fetal lung, *Capld2* was expressed in the mesenchyme from d 12 at the initiation of branching morphogenesis. Increased expression was observed through to d 21, where it became associated with smooth muscle myofibroblasts adjacent to the growing epithelium, which regulates alveolar development (125). Treatment of developing E12–13 rat lungs in explant culture with antisense oligonucleotides to *Capld2* reduced protein levels by 49% and resulted in an apparent stoichiometric decrease of terminal bud development to 47% of control levels, suggesting a crucial role for CAPLD2 in branching morphogenesis (125). Additionally, maximal *Capld2* expression in rat lung mesenchyme at late gestation suggested a role in alveolar development.

Alveolar septation reaches a maximum 7 d after birth (126) in the rat when CAPLD2 is concentrated at the tips of budding secondary alveolar septa (127).

Bronchopulmonary dysplasia is a chronic lung disorder characterized by inflammation and scarring in the lungs that can develop in children born prematurely. A mouse model of bronchopulmonary dysplasia, induced by increased oxygen levels, showed reduced levels of *Capld2* expression following increased oxygen, resulting in arrested alveolar growth. Upon return to normal oxygen levels, alveolar regeneration was observed, and *Capld2* levels returned to normal, suggesting that *Capld2* expression was related to septal growth and alveolar partitioning (127). Similarly, within the E12 mouse kidney corresponding to the initiation of branching, retinoic acid-induced expression of *Capld2* was observed with a localization to the mesenchyme surrounding the developing kidney (128). At E13.5, CAPLD2 was observed within the mesenchymal cells in the nephrogenic zone and in stromal cells surrounding the ureteric bud trunk. At E18.5 and in the adult, CAPLD2 was evident in maturing proximal tubules but not in glomeruli. Heterozygous *Capld1* knockout mice show reduced branching morphogenesis in the kidney suggesting, similar to data on the lung, that secreted CAPLD2 enhanced morphogenesis (128).

Retinoic acid has an indirect, but essential, role in branching morphogenesis. It is believed to stimulate mesenchymal cells to produce a paracrine branching morphogen (129, 130). The stimulation of expression of *Capld2* by retinoic acid in the kidney and lung at early stages of branching may represent a conserved mechanism required for the initiation of branching in branched organs (128). Consistent with this is the localization of PI15 to areas of budding during development in the chicken (117). Branching morphogenesis is, at least in part, regulated via the controlled modification of extracellular matrix through localized concentrations of activating and inhibitory enzymes (118, 131). Supporting this role in extracellular matrix modification is the demonstration that CAP proteins function as proteases (*i.e.*, Tex31, Section IV.A) (132) or as a protease inhibitor (*i.e.*, P151, Section III D) (115).

Mouse *Capld2* is also expressed within the mandible, palate, and nasopharynx regions during craniofacial development at E13.5–17.5 (133). Within the Caucasian and Hispanic populations, the region of the genome containing CAPLD2 has been associated with nonsyndromic cleft lip with or without cleft palate (133). This association was not observed in the Columbian population, and as such the role of CAPLD2 in facial development is inconclusive. However, it is of note that this apparent inconsistency between populations may reflect the complex genetic etiology of the syndrome.

H. Mannose receptor-like/CAP domain and CTL domain containing (CAPCL) subfamily

The MRL gene subfamily is virtually uncharacterized. There have been no primary publications on its potential function or indeed its distribution within mammals. It is however, a clearly annotated group of orthologs within public databases. During our analysis of the CAP superfamily, the MRL were recognized as a distinct CAP subfamily (Fig.

1 and supplementary Fig. 1). This group of orthologous genes was, presumably, named as such because of the presence of the CTL domain. CTL domains are abundant in the mannose receptor (134). The mannose receptor has been implicated in immune regulation through recognition and clearance of microorganisms and serum glycoproteins through interaction with the CTL domain (135, 136). Apart from the presence of a CTL domain, however, the MRL group of proteins within the CAP superfamily has no similarity to the mannose receptor, including not containing a transmembrane domain. As such, we suggest the adoption of the following more informative nomenclature: CAP and CTL domain containing (CAPCL).

The theoretical translations of CAPCL sequences available through public databases display some variability. This is particularly true for the N-terminal region, where start codons may be incorrectly applied. Across the major homologous region of the CAPCL proteins, a minimum of 78% identity was observed (G. M. Gibbs, unpublished data). The CAPCL proteins have each of the four CAP consensus sequence motifs, the PY dipeptide, the Hinge-like sequence with a cysteine spacing of C_{x2}C_{x6}C_{x10}C, and an approximately 215-amino acid C-terminal extension that contains the CTL domain (Fig. 1A and supplementary Fig. 6). Following phylogenetic analysis of nine sequences from six mammalian species, the CAPCL formed a discrete highly conserved clade within the mammalian CAP superfamily (Fig. 1B and supplementary Fig. 1). The CAPCL proteins, like most CAPs, contain a predicted signal sequence and are likely to be secreted into the extracellular matrix.

There are three predicted human CAPCL paralogous genes that are localized to 16q22.1–16q22.3 and a single predicted mouse gene to chromosome 8E1. Our genomic analysis indicated 27 orthologous genes in the bony vertebrates, with 24 in mammals including human, monkey, mouse, cow, horse, and dog. ESTs show a restricted expression profile to the kidney and testis in the human and the mouse.

I. Peptidase inhibitor R3H domain containing-like subfamily

As with the CAPCL family of proteins, there are no primary publications for the peptidase inhibitor R3H domain (R3HDML), yet they are evident as a large group of annotated proteins in public databases. The R3H motif has a consensus sequence RxxxH and has a proposed single stranded (ss) DNA binding function (137). The motif R3H is present in many mammalian CAP proteins, as well as Na-Asp-2 from the nematode parasite (138), which has a known tertiary structure. It is not known whether CAPs containing this motif are ssDNA binding proteins; however, given their typically extracellular localization, this may be unlikely.

The predicted human R3HDML gene is located on chromosome 20q13.12, and its mouse ortholog is located on chromosome 2H3. Our genomic analysis predicts 26 orthologous genes within the bony vertebrates, with 23 orthologs in mammals, including the human, chimpanzee, mouse, dog, and cow. Limited EST data show a restricted expression profile to eye, intestine, and lung in the mouse.

Protein sequence alignments (supplementary Fig. 7) show a minimum of 78% identity between the mouse and human. This subfamily of the CAPs shares each of the CAP motifs, the PY dipeptide, the Hinge-like sequence (Cx₂Cx₇Cx₄C), and a short C-terminal extension of 12 amino acids. Our phylogenetic analyses provide strong support for a close relationship of R3HDMML proteins with PI15 proteins and CAPLD proteins (Fig. 1B).

IV. Structural Characteristics of CAP Superfamily Proteins

The crystal structure of several CAPs has been determined. Conserved structural features among the family provide insight into potential functions, and these are reviewed briefly. As described earlier, numerous subfamilies contain C-terminal extensions including the CRISP domain, the CTL domain, and the LCCL domain, and these likely impart additional functionality. Each of these additional domains will be briefly reviewed.

A. The CAP domain

The tertiary structure of CAP superfamily proteins shows a remarkable conservation, despite often low overall identity and significant evolutionary distance between organisms. Core structural elements are retained (the CAP motifs), resulting in a conserved structure with a conserved putative active site. Variation outside of these sequences may account for the different activities (binding partner and/or enzymatic function) that have been reported in different environments.

1. An enzymatic function? The only enzymatic function described to date for a CAP superfamily protein is as a substrate-specific protease Tex31 from the cone snail *Conus textile* (132). Tex31 showed a specific proteolytic activity for the cleavage of the propeptide of the conotoxin TxVII₁, which was enhanced 5-fold by calcium and inhibited by serine protease inhibitors, suggesting a metalloprotease- or serine protease-like activity. As such, Tex31 may be a novel class of substrate-specific protease (132).

A CAP protein with 67% identity to Tex31 was recently identified from *Conus marmoreus*. This protein, originally identified as GlaCrisp (139) and later as Mr30 (140), has been investigated for proteolytic function. Hansson *et al.* (139) showed no proteolytic activity when tested against the substrate for Tex31. Qian *et al.* (140) showed a very low serine protease-like activity when tested against a substrate not related to the Tex31 substrate. These latter authors could not discount the possibility of trace levels of contaminant proteolytic function and concluded that Mr30 did not have protease activity. Stecrisp from *Trimeresurus stejnegeri* snake venom was tested in protease assay against the Tex31 substrate with no detectable proteolytic activity (14). It remains unclear whether Tex31 and related proteins have a very high degree of protease substrate specificity that precludes observation of proteolytic function in related proteins, or whether variation in the putative active site changes the function of the protein.

2. Structural conservation. The nuclear magnetic resonance (NMR) structure of P14a from the tomato (141) and the crystal structure of several CAPs has been determined (Fig. 3A). These include vespid Ves v5 (11), human GPR-1 (142), nematode Na-ASP-2 (138), and numerous CRISPs from snake venom, *e.g.*, stecrisp, triflin, and natrin (14, 41, 42, 143). In addition, structural coordinates for the venom CRISPs pseudochetoxin (PsTx) (144) and pseudocin (145) are in the Brookhaven Protein Data Bank (PDB). These structural determinations showed that CAP domains possessed a similar overall structure consisting of a unique α - β - α fold that is stabilized by a buried hydrogen bonding network and, with the exception of GPR-1, several disulfide bonds (Fig. 3A). In all except GPR-1, the conserved cysteines within β -strand 3 and β -strand 4, forming the central β -sheet (Fig. 3), form the only conserved disulfide bond present within the superfamily. Other disulfides are typically conserved within a given subfamily. These features provide the thermal, pH, and proteolytic stability reported for CAP proteins, which is consistent with the structural requirements of an extracellular functioning protein (141). The absence of disulfide bonding in GPR-1 proteins is consistent with their intracellular localization (Section III.C). The GPR-1 proteins do contain two conserved cysteines, but structural assessment showed they are not close enough to form a disulfide bond (142).

Protein tertiary structure alignments of all CAPs for which structural data is available show that all the major secondary structure elements overlap (Fig. 3B). This alignment shows 30 amino acids that either are conserved (17) or have conserved properties (13). A ribbon diagram of triflin (41) shows that many of the conserved amino acids are present in the CAP motifs in the core of the domain and many are stabilized by hydrogen bonding interactions (Fig. 3C, i and ii). Of these 30 amino acids, only seven have side chains exposed to the solvent (at least 10% exposure, determined using Swiss PDB Viewer). Solvent exposed amino acids are important for interaction with other proteins in solution, and conserved solvent exposed amino acids may be important for specific interactions or enzymatic active sites. Mapping of these seven conserved solvent exposed amino acids to the ribbon structure and to the surface map of triflin shows that four are located in a solvent exposed cleft, two are located at the top of the cleft, and one is outside of the cleft (Fig. 3C, iii and iv). This solvent exposed cleft is a putative enzyme active site. Several of these conserved solvent exposed amino acids are stabilized via direct and indirect hydrogen bonds to conserved amino acids in the core of the domain underlying the significance of the structural core to the putative active site (11). Of particular note are the conserved histidine residues (H₆₀ and H₁₁₅ in triflin) and the acidic amino acids (normally glutamic acid) that directly support a number of hydrogen bonds (Fig. 3D). A comparison of all CAP structures shows that these two histidines are oriented to the center of the cleft, alluding to their significance in any potential enzymatic function.

The histidines, with an ability to complex divalent cations, form a structure with some similarity to protease active sites (11, 41), and this is consistent with the calcium-activated serine protease-like activity reported for Tex31 (132). In con-

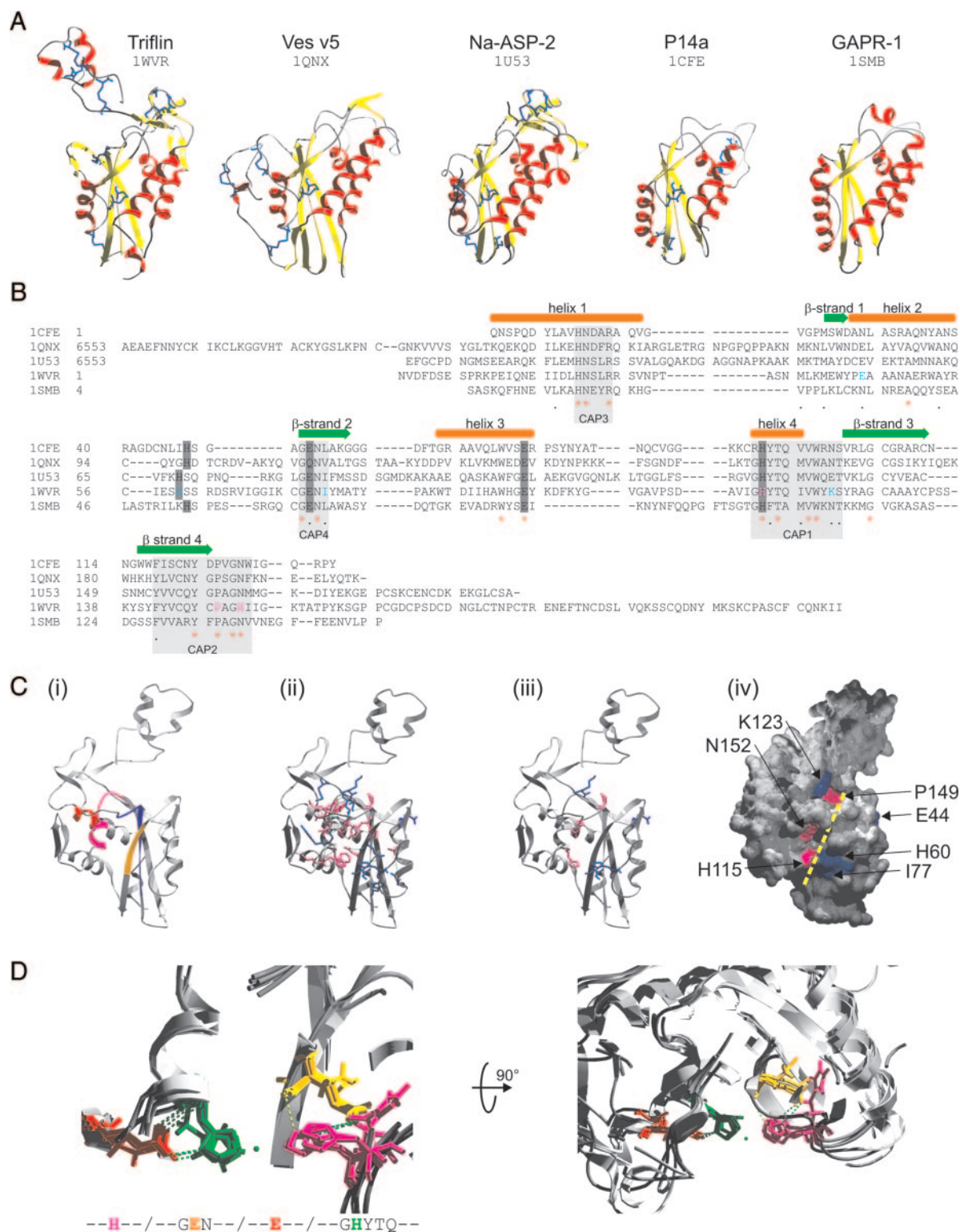


FIG. 3. A, The tertiary structures of CAP superfamily proteins triflin (snake), Ves v5 (wasp), Na-ASP-2 (hookworm), P14a (tomato), and GAPR1 (human) showing the characteristic α - β - α structure forming the core of the protein (α - β - α are red, yellow, and red, respectively). Disulfide bonds forming cysteines are indicated in blue when present. The PDB structure codes are shown below the protein name. B, Alignment of CAP proteins based upon secondary structure elements. The major conserved secondary structure elements are indicated above the relevant amino acid sequences (orange bar, α -helix; green arrow, β -strand). Conserved amino acids from this alignment are indicated by an asterisk, and amino acids with conserved properties are indicated by the dots. Each of the CAP signature motifs are marked (light gray boxes), and the histidines and acidic amino acids involved the putative active site are also indicated (darker gray boxes). Of the conserved or similar amino acids, based upon the structural alignment, those that are solvent exposed (at least 10% exposure) are marked (blue text or pink text, according to conservation

trast, metalloproteases require a catalytic triad of histidines to complex metal ions with a catalytic glutamate (reviewed in Ref. 146), and serine proteases require a conserved serine within a catalytic triad (reviewed in Refs. 147 and 148). From a protease perspective, the putative active site therefore appears incomplete (14, 41, 138). To circumvent this apparent limitation, Serrano *et al.* (142) proposed that the dimerization of GAPR-1 completed the formation of the active site.

From a structural perspective, CAP domains have some of the attributes of a protease, yet they lack the complete functionally active site seen in more “classical” proteases. It is unknown whether this “incomplete” putative active site is the structural basis of the protease inhibitory activity reported for human PI15 (114) or whether this site in the CAP domain represents a new class of protease active sites as proposed for Tex31 (132). The definitive demonstration of protease activity for a mammalian CAP has not been demonstrated; nor has it been exhaustively tested.

Each of the four CAP motifs has been mapped to the ribbon structure of triflin to show their tertiary structure context (Fig. 3). The CAP1 sequence forms the core of the domain and includes two solvent exposed amino acids, notably a conserved histidine oriented to the center of the soluble cleft and a lysine in the turn at the apex of the cleft. The CAP2 sequence forms one of the three major β strands through the core of the protein, followed by a turn that is initiated by a conserved proline. The CAP2 motif contains two conserved solvent exposed amino acids, including an asparagine directed toward the conserved histidines and the proline at the apex of the cleft. Indeed, a synthetic peptide corresponding to the CAP2 motif has been shown to interfere with fertilization in a competitive inhibition assay, which is suggestive of a functionally important interaction at these sites (149) (Section V.C). The CAP3 motif is part of an α -helix on the outer face of the CAP domain; however, the conserved amino acids in this motif often point to the interior of the structure and are stabilized by a network of hydrogen bonds. These interactions in turn stabilize the surface exposed helix. The CAP4 sequence contains invariant amino acids that provide direct stabilizing interactions to a solvent exposed histidine forming part of the proposed active site.

As a point of reference, P14a, from the tomato, and human GAPR-1 represent what may be considered the most basic functional CAP unit. They do not have additional loops, insertions, or C-terminal extensions (Fig. 3A).

B. The cysteine-rich domain or CRISP domain

In addition to the CAP domain, the CRISP proteins are characterized by their approximately 60-amino acid cysteine-rich C-terminal extension. This extension contains 10 conserved cysteines that form five disulfide bonds. Numerous reptile venom CRISP crystal structures (14, 41, 42) and the NMR structure of the mouse CRISP2 CRISP domain (15) show that CRISP domains exist as two distinct structural regions (1), the Hinge region that contains four cysteines forming two disulfide bonds and (2) the ICR region, which contains six cysteines forming three disulfide bonds. The NMR solution structure of mouse CRISP2 CRISP domain showed a very high degree of flexibility between the Hinge and the ICR (15). Combined with crystal structure data obtained from snake CRISPs showing that the Hinge is tightly associated with the CAP domain, with no interactions between the ICR and the CAP domain (14, 42), this suggests that the ICR has a relatively high degree of rotational freedom in comparison to the CAP domain and Hinge region. Although unpublished, the NMR structure of human CRISP2 ICR has been determined with the structure deposited into the Brookhaven PDB (PDB code: 2cq7). This shows the same structure in the ICR as other CRISPs.

In describing the structural domains of the CRISPs, the Hinge and ICR have typically been grouped as one: as the cysteine-rich domain or the CRISP domain. However, in view of the broader mammalian CAP family (Fig. 1), it is apparent that the ICR is the region that is uniquely present in the CRISPs whereas the Hinge is present broadly across the mammalian CAPs, with the strong likelihood that the Hinge is disulfide-bonded in the same manner across the superfamily. As such, the CRISP subfamily of the CAP superfamily may be defined, in addition to their CAP signature sequences, by the unique ICR containing six cysteines and three disulfide bonds. This interpretation also suggests that the Hinge region may have functional significance in its own right. This interpretation is supported by the identification of CRISP2 binding partners that rely on Hinge sequences for their interaction, specifically MAPK kinase kinase II (MAP3KII) and gametogenetin 1 (GGN1) (56, 150).

Uniquely among the CAP superfamily, the CRISPs have ion channel regulatory activity. This was first characterized in helothermine, a CRISP from the venom of the poisonous lizard *Heloderma horridum horridum* (Mexican beaded lizard) (151). Helothermine inhibited I_A type and delayed rectifier voltage gated K^+ channels, the ryanodine receptor (RyR) and

status and consistent with the coloring in part C). C, (i) Mapping the CAP1 (pink), CAP2 (blue), CAP3 (red), and CAP4 (orange) signature motifs on to the triflin structure. (ii) The 30 conserved amino acids identified from the structure alignment are indicated in the ribbon structure of triflin showing a clustering in the hydrophobic core of the protein. Many of these amino acids are from the CAP1 and CAP2 motifs and are primarily involved in forming the structural core of the protein (Section V). (iii) Seven of the 30 conserved amino acids are solvent exposed (at least 10% exposure) and are indicated to contrast with the core amino acids. (iv) The solvent exposed structurally conserved amino acids are mapped to the surface of triflin showing a solvent exposed cleft (yellow dashed line) in the structure. Amino acids are colored according to their conservation shown in the alignment above (pink are conserved amino acids, and blue are similar amino acids). D, The solvent exposed cleft contains a conserved dihistidine motif with conserved acidic amino acids (Asp, Glu) providing a supporting hydrogen bond network (dashed lines). Shown (left) is a close-up of the aligned putative active for each of the proteins in (A) with the same orientation to that shown in (C). The color-coded amino acids in the close-up view are matched with the consensus primary sequence context shown. The small green sphere is a divalent cation complexed with the histidines, which was determined from the triflin crystal structure. Rotation of the structures (right) shows the solvent exposed cleft with histidines penetrating to the center. This view corresponds to the axis shown by the yellow line in (C) from top to bottom. Tertiary structure alignments and structural assessments were completed in Swiss-PdbViewer version 3.7 (324).

voltage gated Ca^{2+} channels (152–154). CRISPs from numerous poisonous reptiles have subsequently been investigated for their ion channel regulatory activity (reviewed in Refs. 6 and 7).

Numerous detailed characterizations of the ion channel regulatory activity from venom CRISPs have shown that they have an activity that appears either broad acting, as in the case of helothermine, or very specific, as evidenced by a comparison of PsTx and pseudocin (reviewed in Refs. 6 and 7). PsTx purified from the venom of the Australian King Brown snake was the first characterized peptide regulator of cyclic nucleotide-gated type A channels (144). Pseudocin, which differs from PsTx by three amino acids in the CRISP domain, has a 30-fold lower activity against cyclic nucleotide-gated type A 2, indicating that certain amino acids determine specific activity and regulate function (145). This has yet to be systematically investigated through site-directed mutagenesis. The inhibitory activities of CRISPs are typically voltage-independent and may directly block ion current by binding to the ion channel pore (153, 155).

The characterization of venom CRISPs has used native proteins containing both the CAP and CRISP domains. As such, it was unclear which part of the protein was involved in ion channel regulatory function. This was recently resolved through the determination of CRISP crystal structures and through the demonstration that recombinant CRISP domain of murine CRISP2 was sufficient for an ion channel regulatory function (14, 15).

Although there is no primary sequence homology, the ICR of CRISPs has a conserved disulfide architecture and an overall molecular shape with a high degree of similarity to the sea anemone toxins ShK and BgK, which regulate potassium channels (156–158). This structural similarity suggests that the ICR alone will be sufficient for ion channel regulatory function.

To date, the ability of mammalian CRISPs to regulate ion channels has only been examined for CRISP2 (15). Specifically, when applied to the cytoplasmic side of RyR prepared from skeletal muscle (RyR1) and heart (RyR2) sarcoplasmic reticulum, the CRISP domain of CRISP2 resulted in the activation of RyR1 and the inhibition through RyR2. When applied to the luminal domain of RyRs, CRISP2 inhibited RyR1 and RyR2 in a voltage-dependent manner. As outlined in Section V.B, several biological properties of other mammalian CRISPs are entirely consistent with roles in regulation of ion channel activity, *e.g.*, CRISP1 as a sperm decapacitation factor and CRISP3 in the pathogenesis for Sjögren's syndrome (Section VIII.A).

Despite conclusive data on the function of CRISP domains, the function of the CAP domain in CRISP proteins largely remains unknown. Alluding to a function, however, was the recent identification of a CRISP protein from the parasitic Japanese river lamprey, *Lethenteron japonicum* (159), and data suggesting its immune suppressive function. The lamprey is one of the most primitive vertebrates and one of the few vertebrate parasites (160). The parasitic lamprey attaches to the host and feeds on blood for extended periods of time. To facilitate this, it secretes factors to assist in feeding, and in this regard, it is similar to the invertebrate parasites such as the human hookworm (Section VII.A). The lamprey secretes two major proteins from its buccal gland. One of these, a CRISP

protein, was shown to have an effect on smooth muscle contraction possibly through high voltage-activated (L)-type Ca^{2+} channels (159), similar to several snake CRISPs. This activity was proposed to assist in feeding through vasodilatation and is likely to be attributed to the CRISP domain. Given the conservation of the CAP domain in parasitic blood meal feeders (Section VII.A), the CAP domain in this context may be involved in immune suppression or in inhibition of platelet aggregation by analogy to the nematode CAPs (Section VII.A).

C. C-type lectin (CTL) domains

CTL domains are C-terminal extensions found in association with the CAP domain in the CAPCL subfamily of the mammalian CAPs (Fig. 1; Section III.H). CTL domains are 100 to 130 amino acids in length and contain four conserved cysteines, which form two disulfide bonds, and a conserved tryptophan (PROSITE PDOC00537). They function in calcium-dependent carbohydrate recognition (reviewed in Refs. 16 and 161).

In addition to their association with the CAP domain, CTL domains are seen as a C-terminal anchor in the macrophage mannose receptor (134); as a C-terminal domain in the collectins, which are implicated as modulators of the innate immune system (reviewed in Ref. 162); as an extracellular domain in the selectins, which are membrane-anchored cellular adhesion molecules with roles in inflammation and cancer (*e.g.*, reviews in Refs. 163 and 164); and are also associated with the LCCL domain. The function of the CTL domain is unknown.

D. The *limulus* factor C, *Coch-5b2*, and *Lgl1* (LCCL) domains

The LCCL domain (also called the factor C homology domain) was named after recognition of homology between the *Limulus* (horseshoe crab) factor C, COCH, and Lgl1 (Section III.G) (17). LCCL domains are normally found in association with other protein domains, including, but not limited to, CTL domains in factor C (165), von Willebrand type-A domains in cochlin and akhirin (166, 167), or CAP domains in CAPLD2 (17, 124). LCCL domains are about 100 amino acids in length, and their C-terminal part contains a highly conserved histidine in a conserved motif YxxSxxCxAAVHxGVI (PROSITE PS50820).

Little is known about the function of the LCCL domain, although it is present in multidomain proteins whose roles span diverse functions. LCCL domains are found in potential complement factors in the case of *limulus* factor C and the Plasmodium Scavenger Receptor Cysteine-Rich, LCCL, Adhesive-like Protein (PSLAP) (165, 168), in the human deafness associated gene (169–173), in akhirin (and cochlin), which have been implicated in eye development (167, 174, 175), and also within CAPLD2. It has been proposed that proteins containing LCCL domains may be involved in cellular adhesion (167, 168), although it is not known whether this is associated with the LCCL domain specifically or with sequences.

V. CAP Proteins in Reproduction

Proteins of the CAP superfamily are intimately associated with the reproductive tract and fertilization. Both published and unpublished data suggest that several CAPs are expressed during spermatogenesis and that they remain associated with sperm through to the final stages of fertilization. Similarly, several CAPs are produced within the epididymis and associate with sperm during epididymal maturation. Yet others are secreted into components of seminal plasma. Bearing in mind the potential for redundancy between closely related protein family members and poorly defined specificity of some analytical reagents, outlined below are the data supportive of a role for CAP proteins in male fertility. The primary mRNA and protein distribution for each member has been described in *Section III*.

A. Spermatogenesis

Spermatogenesis occurs in the testis and is the process whereby a germ cell develops into a sperm. Based on published data, CRISP2 is the only mammalian CAP protein characterized as being produced during spermatogenesis, but an analysis of EST databases and unpublished data suggests that several other superfamily members are also produced in the testis.

In all mammalian species studied to date, CRISP2 is produced initially in either late pachytene spermatocytes or round spermatids, and CRISP2 becomes localized to the acrosome covering the sperm head, the accessory structures of the sperm tail, and on the developing germ cell membrane (28, 29, 31, 32, 51). CRISP2 appears to be a component of the soluble and diffusible fraction of the electron dense acrosomal matrix (176).

A role for CRISP2 in cell-cell adhesion has been proposed. Experiments whereby rat spermatogenic cell-derived cDNAs were transfected into Jurkat Tag cells identified CRISP2 (at that time called TPX1) as a protein putatively involved in mediating adhesion between developing germ cells and Sertoli cells (31). This result was further supported through the ability of CRISP2 antiserum to interfere with the adhesion of purified germ cells to cultured Sertoli cells. Cell binding ability was localized to the N-terminal-most 101 amino acids, and was completely independent of the CRISP domain (53). This aspect of CRISP2 function has not been explored further.

Within the sperm tail, the function of CRISP2 is currently unknown; however, it may be related to motility and its dependence on ion channel regulation. Immunoelectron microscopy and Western blotting data generated from the rat indicate that CRISP2 is a component of the connecting piece, the outer dense fibers, and the longitudinal columns of the fibrous sheath (51).

Within the male reproductive tract, although sperm may have the capacity for motility, they are kept in a quiescent state by decapacitation factors. Upon entering the female reproductive tract, or culture media, sperm cells begin to move with a modal type of motility. At this time, sperm cells also begin to undergo an additional set of maturation events collectively called “capacitation,” during which extensive protein rearrangements and posttranslational modifications

occur (reviewed in Ref. 177), and they gain the ability to undergo the acrosome reaction, bind an oocyte, and exhibit hyperactivated motility. All of these processes are critically regulated by Ca^{2+} flux (178–182). In relation to Ca^{2+} regulation, we have recently shown that CRISP2 can regulate Ca^{2+} flux through RyRs (15). Although the RyR subtypes have not been defined, RyRs have been localized to the connecting piece of human sperm and have been implicated in the generation of a stable oscillating pattern of intracellular Ca^{2+} correlating with sperm tail movement (183). Although the RyR data are suggestive of a role for CRISP2 in motility regulation, it is not conclusive. Given the range of ion channels regulated by venom CRISPs (*Section IV.B*) and the high number of ion channel types found on sperm (reviewed in Ref. 184), it is plausible that CRISP2 may also regulate other channel types.

Recent experiments from our laboratory have shown that CRISP2 exists in a complex within sperm. Specifically, CRISP2 binds to MAP3KII (also called MLK3), and they co-localize in the acrosome (150) and to GGN1 in the sperm tail (56). Within somatic cells, MAP3KII is an upstream activator of Jun kinase (185, 186); however, in the absence of sperm transcription and the localization to the acrosome, it is more likely that MAP3KII is involved in posttranslational phosphorylation of target proteins (187). MAP3KII may therefore regulate CRISP2 function during capacitation. GGN1 on the other hand, is a protein of unknown function that appears to bind to CRISP2 in the sperm tail (56).

In addition to CRISP2, an analysis of Unigene (188) EST expression profiles suggests that all members of the GLIPR subfamily are expressed in the testis (*i.e.*, GLIPR1, GLIPR1L1, GLIPR1L2 and GLIPR1L3), as are PI16 (human), CAPLD1 (human), CAPLD2 (human and mouse), and the CAPCL proteins (human and mouse). The cellular localization and function of these proteins in testis function and fertility remain unknown.

Recently, the involvement of CAP superfamily proteins within nonvertebrate sperm was highlighted (189). The ascidian (sea squirt) *Halocynthia roretzi* produces two CAP superfamily proteins called HrUrafin and HrUrafin-long (L) within their testes and on sperm. These proteins have highest homology to GLIPR1L1 and GLIPR1L2, respectively, and are GPI-anchored to the sperm membrane and localized to lipid rafts (189). It is of note that MAK248 (another GLIPR1L1 ortholog) is also GPI anchored to sperm within the epididymal lumen of the macaque (190). Ascidians are hermaphroditic and have developed mechanisms to select against self-fertilization, thus ensuring the maintenance of genetic diversity. Although HrUrafin appears to be an adhesion protein in a manner conceptually similar to that proposed for CRISP1 (*Section V.C*), it has a role in the negative selection against “self” oocytes (189). HrUrafin is the first CAP superfamily protein reported within nonvertebrate sperm.

B. Epididymal maturation and sperm capacitation

Because epididymal maturation is essential for fertilization (reviewed in Refs. 191 and 192), proteins of epididymal origin may also be essential for sperm function (61, 193–195). This has been a central tenet of approaches to sperm-specific

contraceptive development over numerous years (196, 197). Proposed functions for CRISPs in this context can be broadly categorized into decapacitation factors and sperm-oocyte receptor complex formation.

During epididymal maturation, sperm cells come into contact with at least CRISP1, CRISP3, and CRISP4. CRISPs within the epididymal fluid surround sperm in high concentrations and interact with the sperm membrane with varying affinities. The mechanisms of attachment and their relative strength dictate whether CRISPs disassociate from sperm early in the female reproductive tract, or whether they stay associated with sperm until the final stages of fertilization, and fundamentally effect the ultimate biological function.

Capacitation is an incompletely defined continuum of molecular events that occurs in sperm in the female reproductive tract and is required to achieve fertilization (198). The molecular mechanisms involved during capacitation are complex, and many excellent reviews exist; however, a few pertinent points will be raised here. Capacitation is initiated and regulated by pathways including cholesterol removal and plasma membrane remodeling with the redistribution/dissolution of lipid rafts (199–202), intracellular alkalinization from HCO_3^- , and an increase in intracellular cAMP (203, 204) (reviewed in Ref. 205), resulting in the activation of numerous signal transduction pathways (reviewed in Ref. 206). Correlates of capacitation include a massive up-regulation of tyrosine phosphorylation associated with the sperm tail and head (207), the manifestation of hyperactivated motility (208, 209), the ability to bind the zona pellucida of the oocyte, and the subsequent induction of the acrosome reaction (210, 211).

Decapacitation factors are proteins that inhibit sperm capacitation, and their removal from sperm is required to achieve capacitation (212–214). Decapacitation factors, of which there appear to be many, are often of epididymal or accessory organ origin (67, 215–219) and are removed from sperm after entry into the reproductive tract or by incubation *in vitro* in capacitating conditions.

The larger glycosylated form of CRISP1 has been shown to function as a sperm decapacitation factor in rats and to reversibly inhibit capacitation-associated protein tyrosine phosphorylation and the acrosome reaction in a dose-dependent manner (67). CRISP1 was also identified in a fraction of low molecular mass diffusible proteins from mouse sperm, which upon readdition to sperm caused the inhibition of the progesterone-induced acrosome reaction and the inhibition of tyrosine phosphorylation in membrane proteins, without grossly affecting tyrosine phosphorylation within the head or tail (215). The differences observed in the effect of CRISP1 on rat and mouse sperm protein phosphorylation may be due to species variation or the heterogeneous protein pool used in the latter investigation. Roberts *et al.* (67) proposed that CRISP1 interacts with a specific lipid or protein on the sperm surface to inhibit capacitation. This is entirely consistent with the inhibition of ionic signaling of CRISP proteins as demonstrated for CRISP2 (15), the high concentrations of CRISP1 and CRISP4 in the epididymal lumen, and the reversible association of the majority pro-

portion of the CRISPs to the plasma membrane. A similar role for CRISP4 as a decapacitation factor is possible.

Given first, the demonstrated ion channel regulatory activity (primarily inhibitory) of CRISP domains; second, the high concentrations of CRISPs in the cauda epididymal lumen; and third, the decapacitation factor activity of CRISP1, it is most likely that at least one function of the epididymal CRISPs is to store sperm in a quiescent state through inhibition of ion signal transduction. After entry into the female reproductive tract and the dissociation of much of the CRISP content from the surface of sperm, or through their potential inactivation by binding to other proteins, as in the case of CRISP3 binding to β -MSP (79), specific classes of ion channels may become activated, thus allowing ion signal transduction and sperm capacitation.

Surprisingly, given the above data, an analysis of male *Crisp1* knockout mice showed normal fertility and produced litters of normal size. Even more surprisingly, capacitation as indicated by increased global tyrosine phosphorylation, was reduced (60). Membrane protein phosphorylation was not specifically assessed. From the data presented, it was unclear whether sperm capacitated more slowly or whether the full suite of tyrosine phosphorylation is not required for fertility. These data do, however, suggest that the removal of CRISP1 results in deregulated signal transduction. To dissect the true role of CRISPs in capacitation on an equivalent footing to that occurring in human sperm, it may be necessary to produce *Crisp1-Crisp4* double knockout mice.

Calcium plays a central role in sperm capacitation and is essential both for achieving hyperactivated motility (220, 221) and for the initiation of the acrosome reaction (*e.g.*, reviews in Refs. 222 and 223). An increase in intracellular Ca^{2+} within the sperm tail is required for the initiation of hyperactivated motility, and subsequent cyclic changes in intracellular Ca^{2+} correlate with the beating of the sperm tail. The *Catsper* ion channels are specific to the principal piece of the sperm tail, and the *Catsper1–4* null mice are all sterile through an inability to undergo hyperactivated motility (224–227). Infertility is attributed to a lack of Ca^{2+} ion entry into the sperm tail from extracellular sources (180), but it has recently been shown that stimulation of release of Ca^{2+} from intracellular stores at the base of the sperm head in the *Catsper1* and *Catsper2* null mouse is sufficient for sperm to hyperactivate (181). The most likely candidate for the intracellular Ca^{2+} store at the base of the sperm head is the redundant nuclear envelop, and the central role for this compartment in regulating hyperactivated motility is becoming apparent (181, 228). The signal transduction pathways regulating the release of intracellular Ca^{2+} from the redundant nuclear envelop are unknown, and the identity of the ion channels in this compartment remain unclear; however, both the IP3R and the RyR have been identified in different species (183, 229). Thus, as outlined above, given the localization of CRISP2 to the connecting piece and its ability to regulate RyR activity, it is plausible that CRISP2 is involved in this process. Given that CRISP1 (and likely CRISP4 and CRISP3) diffuse off sperm during the early stages of capacitation, it is also possible that their removal alleviates repression of specific ion channels resulting in the initiation of signal transduction cascades that contribute to hyperactivation.

The binding of CRISPs to β -MSP, or similar proteins, may be functionally important and relevant to all CRISPs in the reproductive tract. This is supported by the interaction of PI16 with PSP94 (β -MSP) in human serum (80) and the binding of triffin to SSP-2 (β -MSP) in snake serum (230, 231). Conservation of binding partners identified from diverse sources supports this interaction as being functionally significant. In each case, the high affinity interaction was proposed to have an inactivating/self-protective function. β -MSP is one of the most abundant proteins secreted from the prostate (232). β -MSP binds to the sperm midpiece and head and regulates motility through inhibition of Na^+ , K^+ -AT-Pases, it comprises a large proportion of the protein released from sperm after ejaculation, and it inhibits the spontaneous acrosome reaction (233–236). Increased expression of β -MSP may be associated with reduced fertility in men (236). As yet, the significance of the high-affinity interaction and the importance of these proteins in regulation of sperm capacitation remain unclear.

To date, only one non-CRISP CAP superfamily protein has been identified in the mammalian epididymis, MAK248 from the macaque (190). MAK248 is most closely related to GLIPR1L1 in the mouse (supplementary Fig. 1). However, differences in the expression of these two proteins (exclusively epididymal *vs.* exclusively testicular) suggest that they may not be strict orthologs. MAK248 is a GPI-anchored protein that appears to undergo epididymal-associated proteolytic cleavage into two disulfide-linked peptides and is localized to the equatorial segment and the posterior head of capacitated macaque sperm (190). The function of MAK248 remains unknown, yet the presence of a CAP domain containing protein distinct to the CRISPs suggests that the CAP domain specifically has a conserved function within the reproductive tract.

C. CAP's role in sperm and egg fusion

As indicated in part A of this section, transfection data have suggested that a sequence in the N-terminal domain of CRISP2 may be involved in germ cell binding to Sertoli cells (31). Although this specific function has not been explored further, these data do fit with recent data suggesting that peptides corresponding to the CAP2 signature motif of either rat CRISP2 or CRISP1 can interfere with *in vitro* fertilization (IVF) in the rat, specifically sperm-oocyte fusion (149). Although the essential nature of CRISP1 in oocyte fusion is refuted by a recent knockout of the mouse *Crisp1* gene (60), these data suggest that there are sequences within the CAP proteins that are fusogenic. Conflicting data exist concerning the ability of CRISP2 antibodies to interfere with IVF; specifically, data from the Cuasnicu laboratory showed that the addition of anti-CRISP2 serum to human sperm in the zona-free hamster oocyte assay inhibited sperm binding (55). However, antibodies to CRISP2 in a guinea pig assay did not inhibit IVF (28), nor did CRISP2 antibodies inhibit the binding of human sperm to excess human oocytes (M. K. O'Bryan, unpublished data). The concrete interpretation of these data is currently further limited by the high degree of sequence homology between CAP members in the CAP signature motifs, uncertainties as to the presentation of CRISP2 on the

surface of either developing germ cells or sperm, and the finding that based on available crystal structures, seven of the 12 amino acids in the fusogenic inhibitory peptide sequence are buried within the internal structure of the CAP domain (Section IV.A). An analysis of the tertiary structure context of the CAP2 motif from the venom CRISP triffin (Fig. 3) suggests that two of the five amino acids not buried within the core of the structure have side chains that are likely to be solvent exposed. These are the conserved proline and asparagine, which are localized to the top of the solvent exposed cleft that contains the putative active site (Fig. 3). Whether this pentamer containing two solvent exposed amino acids is sufficient for binding has not been determined. Regardless, because we currently know very little about the posttranslation processing for sperm surface proteins, we believe that it remains plausible that a protein bearing a CAP2 signature motif, and thus a CAP, is involved in binding male germ cells to either Sertoli cells or oocytes.

Although *Crisp1* knockout males have normal fertility, several lines of evidence suggest a subtle role in the fertilization process. Certainly after capacitation in a number of species, a proportion of CRISP1 remains associated with the fusogenic portion of the sperm, namely the equatorial segment (Section III.A). CRISP1 antibodies have been shown to inhibit rat sperm fertilization (237–239), and CRISP1 protein has been shown to bind to the oocyte, suggesting complementary binding sites (240–242). The immunization of male rats with CRISP1 significantly and reversibly reduced male fertility through inhibition in egg penetration by the sperm, with accumulation of sperm in the perivitelline space (243, 244). Anti-CRISP1 antibodies in the epididymis of immunized rats bound to CRISP1 without affecting expression levels or the localization of CRISP1 to the sperm plasma membrane. Sperm capacitation and the binding of sperm to the plasma membrane of zona pellucida free eggs were also not affected. The same results have been obtained for mouse CRISP1 when using purified CRISP1 in competitive binding assays (241), and for mouse CRISP2 and human CRISP1 when using immunological inhibition (55, 242). A similar activity was demonstrated for Izumo in the mouse (245). Izumo is generally regarded as a *bona fide* sperm oocyte receptor protein.

The reason for the disconnection of *in vitro* and *in vivo* CRISP1 activity in particular may well be the functional redundancy between CAP family members or cross-reactivity of analytical reagents with other CAPs. As such, it may only be after the generation of combination null mice that the true story will become clear.

D. Female reproductive tract

The extent of expression of CAP superfamily proteins within the mammalian female reproductive tract is largely unexplored. We have recently shown expression of CRISP1 and CRISP3 within the murine uterus produced in the secretory epithelial cells (45), and numerous members of the CAP superfamily may be expressed within the ovary and uterus based upon EST analysis. CAP superfamily proteins have, however, been characterized in the female reproductive tract of nonmammalian species, in particular

from *Xenopus*, where they have a demonstrated biological function.

Organisms that undergo external fertilization require chemoattraction to ensure that sperm can find the oocyte and achieve fertilization (reviewed in Ref. 246). The three layers of jelly surrounding the *Xenopus* oocyte are comprised of a matrix of high molecular mass proteins and low molecular mass diffusible proteins (247). An analysis of the diffusible components from *Xenopus laevis* egg jelly identified a protein, called allurin, as having a specific sperm chemoattractant activity (248). Allurin is secreted from the jelly-secreting ducts of the *Xenopus* upper oviductal epithelium, where it becomes localized to the surface of the ciliated cells (249). During passage of the oocyte through the oviduct, the ciliated cells deposit allurin into the developing egg jelly layers (249). After transport through the oviduct, allurin makes up 3% of the total egg jelly protein (249). After spawning, most of the allurin within the outer jelly layer is released within 5 min, whereby it functions in sperm chemoattraction (250).

Although the mechanism of chemoattraction remains unknown, the biological activity appears conserved among *Xenopus* species. Allurin from *X. tropicalis* has a chemotactic function for both *X. tropicalis* and *X. laevis* sperm (251). Similarly, *X. laevis* allurin has a chemotactic activity for *X. tropicalis* sperm (251). Allurin was the first vertebrate protein isolated to have a sperm chemotactic activity.

Allurin has highest sequence identity (about 40%) to the CRISPs (252), and similar to the mammalian CRISPs, it has a Hinge-like sequence with the same spacing of cysteine residues (Cx₂Cx₃Cx₄C). Beyond this however, the sequence is terminated, and there is no ICR domain. The expression profile similarity and structural similarity raise the possibility of a chemotactic activity of the mammalian CAPs, although this remains undetermined.

E. Genetic studies associate SNPs in CRISPs with reduced fertility

To date, the only CAPs to be assessed for their potential as a genetic cause of infertility are the CRISPs. An analysis of nonsynonymous single nucleotide polymorphisms (SNPs) in equine *Crisp3* was suggestive of a role in male fertility (253). The authors identified three polymorphisms within *Crisp3*, and one heterozygous SNP encoding E208K was significantly associated with reduced stallion fertility (253). E208K is localized to the CRISP domain of CRISP3 in a region of increased amino acid sequence variability among the CRISP protein family. A biochemical basis for reduced CRISP3 function has not, as yet, been investigated. The study also identified several polymorphisms in the equine *Crisp1* and *Crisp2* genes; however, no association with impaired fertility was detected (253).

Several studies on human male fertility have proposed a role for CRISP2; however, concrete proof of this is lacking. The analysis of an Albanian family identified a t(6;21)(p21.1;p13) translocation as the causal factor in male-specific infertility characterized by severe oligoasthenoteratospermia (254). CRISP2 was contained within the deleted region. Similarly, a larger analysis of break points in infertile male patients carrying balanced translocations identified

CRISP2 as a candidate causal gene in three cases of male infertility (255).

More recently, our group has undertaken a detailed analysis of polymorphisms in the CRISP2 gene in a large number of infertile patients with a range of phenotypes (52). This study identified 21 CRISP2 polymorphisms, indicating an extremely high level of polymorphism between individuals. Three polymorphisms resulted in amino acid substitutions, including one that resulted in the loss of a strictly conserved cysteine involved in intramolecular disulfide bonding in the Hinge region (C196R). This mutation was only observed in a heterozygous state and occurred in similar frequencies in the infertile and control populations; however, a yeast two-hybrid analysis showed that in a homozygous state this polymorphism resulted in the loss of the ability of CRISP2 to bind to its binding partner GGN1 (52, 56).

VI. CAP Proteins in Cancer

Increasingly, expression data are suggesting that the deregulation of particular CAP proteins, notably GLIPR1 and CRISP3, is associated with, and perhaps involved in, the development and progression of cancers. These observations have raised the possibility that CAPs may be used as diagnostic agents and perhaps ultimately as therapeutic targets.

Several lines of evidence suggest a role for CRISP3 in the progression of prostate cancer. CRISP3 is expressed at low levels within the normal human prostate (33), yet it is highly up-regulated in the cancerous prostate. By comparing microdissected prostate cancer tissue to adjacent “normal” tissue in a microarray analysis, Ernst *et al.* (256) showed that several genes within 6p21 were significantly up- or down-regulated. Of these, CRISP3 was the most highly up-regulated with a 21-fold overexpression in cancerous tissue compared with matched control tissue. This conclusion was supported by other studies where a 20- to 2000-fold up-regulation was observed (91, 92). The majority of CRISP3 production came from the prostatic epithelium, rather than the surrounding stromal tissue (256). Bjartell *et al.* (258) showed, using immunohistochemistry on prostate biopsies, that CRISP3 was significantly up-regulated in regions of high-grade cancer (Gleason scores 4/5). Furthermore, the mRNA and protein for the CRISP3 binding protein β -MSP was dramatically down-regulated in prostate cancer (257), suggesting that prostate cancer progression may be critically affected by the ratio of bound to free CRISP3.

Although the mechanism of CRISP3 activity in prostate cancer is yet to be revealed, CRISP3 has been explored as a marker of long-term patient outcomes and as a marker of regression. Patients with increased expression of CRISP3 in cancerous regions of the prostate had a lower recurrence-free probability (258). Serum CRISP3 concentration was found to correlate with progression in both a univariate and multivariate model with β -MSP. However, the inclusion of CRISP3 in existing predictive models (*i.e.*, prostate serum antigen and pathological staging) did not improve their power, potentially as a consequence of numerous confounding endogenous sources of CRISP3 (93,

258). As expected for a gene containing an androgen-response element in its promoter, the levels of serum CRISP3 fell after orchiectomy (93).

CRISP3 was also shown to be significantly down-regulated in tongue tissue from patients with squamous carcinomas, compared with healthy tissue (94). The significance of this change in expression is yet to be explored.

The cancerous human prostate also showed aberrant expression of a member of the GLIPR subfamily. In contrast to CRISP3, however, GLIPR1 was suppressed in prostate cancers. Within the mouse, GLIPR1 (also known as RTVP1) was originally identified as a p53-stimulated prostate transcript (102). The transfection of GLIPR1 into a number of mouse and human prostate cell lines resulted in increased apoptosis by both p53-dependent and p53-independent pathways (102, 103). Similarly, viral delivery of *Glipr1* in a preclinical model of prostate cancer resulted in decreased tumor growth and metastasis (103, 259). In addition to increased rates of apoptosis, decreased tumor progression after virally mediated delivery *in vivo* appeared to depend, at least in part, on antiangiogenic and immunostimulatory activities (259). The preimmunization of mice with a *Glipr1*-modified tumor cell vaccine was also partially protective against the development of prostate cancer in an orthotopic cancer model (260), and, perhaps most compellingly, mice lacking *Glipr1* had a significantly increased predisposition to spontaneous tumor growth when compared with wild-type littermates (261). It is also of note that *Glipr* heterozygous animals were significantly more disposed to developing tumors than wild-type animals, suggesting that the proapoptotic effect of GLIPR1 is stochastic in nature (261). Tumors were observed in several tissue types. Concurrent p53 inactivation greatly increased the rates of tumor development; however, rates were still significantly above that of p53 inactivation alone, suggesting that for GLIPR1, the p53-independent apoptosis is functionally significant in carcinogenesis.

In terms of the mechanism of apoptosis, transfection experiments suggest that GLIPR1 production results in elevated reactive oxygen species production, leading to the sustained activation of c-Jun N-terminal kinase (JNK); a shift in the balance between pro- and antiapoptotic BCL2 family members toward the apoptotic pathway; and ultimately apoptosis by both caspase-dependent and caspase-independent pathways (261).

With one notable exception, there is also a compelling case for GLIPR1 being a human tumor suppressor gene. A case for using gene therapy for delivery of GLIPR1 to prostate cancers is most well developed; however, expression analyses show that GLIPR1 is also suppressed in human cancer cell lines derived from multiple sources, including lung, colon, and bladder cancers and lymphomas, and that the forced up-regulation of GLIPR1 resulted in apoptosis (261). GLIPR1 gene therapy may therefore be of value for a range of cancers. Indeed, at the time of writing this review, a phase I/II clinical trial on the potential of adenoviral-mediated GLIPR1 cancer therapy was under way (IND13033). mRNA and protein analyses using human prostate tissue have shown that GLIPR1 production is suppressed, largely via methylation of the promoter, in cancerous regions compared with normal prostate tissue (103).

Perhaps the loudest note of caution against widespread GLIPR1 delivery at the moment is the identification of GLIPR1 as one of the most highly up-regulated transcripts in human gliomas (astrocyte-derived brain tumors) (100, 101). In contrast to the situation for GLIPR1 in prostate cancer, GLIPR1 is up-regulated in astrocyte-derived tumors and cell lines, and the degree of expression correlates with the degree of invasiveness *i.e.*, glioblastoma > anaplastic astrocytomas > low-grade astrocytomas > normal brain (100, 104). GLIPR1 is expressed in very low levels in normal brain tissue, and the delivery of interfering RNAs into glioma cells results in elevated apoptosis. Conversely, the transfection of GLIPR1 into glioma cell lines increases both their proliferation rate and their ability to form colonies in agar (104).

In an almost mirror image of the effect of GLIPR1 on prostate cells, the transfection of GLIPR1 into a glioma cell line resulted in decreased phosphorylation (suppression) of JNK and a shift in the BCL2 family members toward a prosurvival phenotype (104). Clearly these data indicate that the GLIPR1 function is context dependent, and at this time the pivotal point between pro- and antitumorigenic effects is not known, but it may involve the relative levels of each of the four GLIPR1 transcripts that have been observed (101, 104). Regardless, the potential to use GLIPR1 gene delivery or suppression as a targeted therapy is an exciting prospect, and the outcome of clinical trials is eagerly awaited.

Elevated GLIPR1 expression has also been observed in a number of other tumor types, although its mechanism has not been determined. GLIPR1 is, for example, highly up-regulated in Wilms' tumor (WT). WT is a renal cancer that arises during embryonic development and typically occurs in children. WT has a complex etiology and is associated with both genetic and epigenetic changes (262–266). Using a comparative analysis of differentially methylated regions in the genomes of normal fetal and WT kidney tissue, Chilukamarri *et al.* (267) identified the GLIPR1 promoter as being a frequently hypomethylated locus in WT. Consistent with this, GLIPR1 mRNA and proteins were significantly up-regulated in the majority of WT tissues examined (267).

VII. CAP Proteins in Immune Regulation

Several mammalian CAP proteins show a notable expression bias to cells and tissues involved in immune surveillance. These include expression within cells involved in both the innate and adaptive immune systems and in the body's response to diseases with an immune component, including cancers and autoimmune disorders. Although the mechanisms of action of CAPs within these settings is currently largely unknown, it is tempting to draw parallels with the roles of nonvertebrate CAPs such as activation-associated secreted proteins (ASPs), Ag5, and the Pr-1 family of CAPs having an immune regulatory function. In the hope of shedding light on the role of mammalian CAP function in the immune system, several of these proteins are briefly described.

A. *Anclystoma*-secreted protein, or activation-associated secreted protein (ASP)

CAP superfamily proteins have been identified in several species of both free living (*C. elegans*) and parasitic nematode. Parasitic worms gain residence in their host through suppression and evasion of the host immune system (268). *Anclystoma caninium* (dog hookworm) produces numerous CAP proteins that are secreted in high abundance immediately after the transition of the free living larvae to the parasitic form. These include the neutrophil inhibitory factor (NIF), the hookworm platelet inhibitor (HPI), ASP1, and ASP2 (269–272). These proteins are produced at the crucial early stage of parasitic invasion, whereby the parasite attempts to evade the host immune system through secretion of immune evasion molecules, ultimately to facilitate and enhance feeding.

NIF was identified after a screen for inhibitors of leukocyte function (269). The 41-kDa glycoprotein blocked adhesion of activated neutrophils to vascular endothelial cells via direct high-affinity binding to the extracellular domain of CD11b/CD18 integrin (269, 273). Integrins are involved in the transfer of signals from the extracellular matrix to the cytoplasm and are key to the control of these signals (*e.g.*, review in Ref. 274). The CD11b/CD18 integrin (otherwise known as macrophage-1 antigen, complement-receptor 3, or $\alpha_M\beta_2$ integrin) is involved in the mammalian innate immune system and mediates inflammation through the regulation of leukocyte adhesion (275). HPI has 36% identity to NIF, interacts with the integrin glycoprotein IIb/IIIa and GPIa/GPIIa, and inhibits platelet aggregation (270).

ASP1 and ASP2 each have about 30% identity to NIF. ASP-related proteins are expressed throughout many, if not all, nematode species, are produced in great abundance shortly after transition to the parasitic form (268), are secreted, and appear to have a conserved function related to evasion of host immunity or to enhanced feeding (270, 272, 276–282). One of the major human parasitic nematodes, *Necator americanus*, contains a family of nine ASP genes that are the most highly represented within EST libraries (282), alluding to their expression abundance. Based upon similarity in sequence and expression profile to the ASP from the dog hookworm, these may have a similar function in immune evasion or inhibition of platelet aggregation.

As an aside, EST analysis of *N. americanus* identified a group of small molecules with similarity to potassium channel blockers from the sea anemone (282). The hookworm secretes, among others, two classes of proteins in high abundance, one with a CAP domain and one with an ICR function. The ICR in the hookworm has similarity to the BgK toxin from the sea anemone (158). The ICR domain of CRISPs (Sections III.A and IV.B) has structural similarity to the BgK toxin (14, 15). Thus, it appears that mammals convergently evolved CRISPs that incorporate both domains into a single protein. In the very same vein, CTL domain-containing proteins are secreted from *N. americanus* (282). The CTL domain from the human hookworm has greater similarity to the mammalian P-selectins than they do to the lectins from *C. elegans* (282). Mammals produce CAP and CTL domain-con-

taining proteins as a single gene product, notably the CAPCL genes (Section III.H).

B. Antigen 5 (Ag5)

The Ag5 proteins form a major and distinct clade of the CAP superfamily. Ag5 proteins were first identified within the fire ants and numerous species of wasps reviewed in Ref. 8). They were later also identified within the midgut of *Drosophila* (1, 283) and are secreted from the saliva of blood-feeding ticks (284), sand flies (285), stable flies (286), and mosquitoes (287, 288).

The Ag5 proteins are the most abundant proteins within wasp venom and are frequently associated with eliciting a strong allergenic response in people (289–292). Although detailed characterizations of the allergenic response have shown both an IgE and an IgG component (293–295), the function of Ag5 proteins within wasp and fire ant venom remains largely unexplored.

Within the blood-feeding ticks, flies, and mosquitoes (284–288), the Ag5 proteins are part of a cocktail of salivary proteins that are believed to function either in suppression of the host immune system or in prevention of clotting to prolong feeding (reviewed in Ref. 296). In this regard, the biological function is similar to that proposed for the ASP and for a CRISP from the parasitic river lamprey (159). Because the Ag5, ASP, and lamprey CRISP have the CAP domain in common, it is most likely that function is encoded within the CAP domain.

C. Pathogenesis related-1 (Pr-1)

Plants display numerous responses after pathogen infection that are aimed at the isolation and elimination of the pathogen (reviewed in Refs. 9 and 297). Among these are the hypersensitive response resulting in the isolation and death of infected cells. The hypersensitive response sees a dramatic up-regulation of the Pr proteins (reviewed in Refs. 9 and 297–299). Of the 17 distinct groups of Pr proteins (9), the Pr-1 proteins are the most highly up-regulated and contribute to 10% of the total protein within the infected leaf (4). The Pr-1 proteins are a major group of the CAP superfamily (298) and were first identified within the tobacco plant after infection by the tobacco mosaic virus (300, 301). Although it has been demonstrated that Pr-1 proteins are not antiviral (302), they may possess an antioomycete (water mold) activity (303, 304) through negative regulation of β -(1–3)-glucanase (305).

The increased expression of Pr-1 proteins throughout non-infected parts of the plants suggests a role in the plant systemic acquired resistance response (9, 298, 306), which is analogous to acquired immunity in mammals. Pr-1 expression also occurs during distinct developmental stages of the normal healthy plants (307). Therefore, similar to the situation in mammals, various roles of Pr-1 proteins may be dependent upon expression context. For a more expansive review of the Pr-1 protein within plants, readers are referred to numerous excellent reviews (9, 297, 298, 306).

D. An immune function for mammalian CAPs

The CAPs from nonmammalian sources with immune regulatory function operate through two general mechanisms:

1) introduction by foreign or invasive organisms for evasion of host defense/enhanced feeding; and 2) up-regulation of CAPs by the host for defensive purposes. In all of the characterized instances relating to immune invasion, the targets of CAP protein binding are different. The changed binding partner interactions may be a consequence of molecular evolution, and these changes may determine their biological function. Although this premise makes it impossible to predict immune function of mammalian CAPs based upon similarity to distant homologs, it does provide a basis for the examination of their function within immune environments.

As noted already, our recent analysis of the CRISPs in normal mouse tissues, CRISP1, CRISP3, and CRISP4 shows a marked expression bias to tissues with roles in the immune system (45). These include the thymus and the spleen, which contain *Crisp1*, *Crisp3*, and *Crisp4*. CRISP3 is produced within B cells where its expression is regulated by OCT2 (83) and in neutrophils where it has been localized to secretory granules (35). Neutrophil granules are rich in matrix degradation proteins whose function is to contribute indirectly to microbicidal potential through the degradation of the extracellular matrix (308–310). Localization of CRISP3 to this granule is supportive of a role in matrix degradation and is consistent with the proposed role of rat CAPLD2 (Lg11) in branching morphogenesis, which requires the controlled remodeling of the extracellular matrix (311). Functional characterization of PI15 as a protease inhibitor (114) or Tex31 as a protease (132) is also consistent with these observations. The distribution of CRISP1 and CRISP4 within mouse immune tissues suggests that these may also have a role in immune regulation (45), although these have not been tested.

GLIPR1 has also been identified within cells of the macrophage lineage and up-regulated in response to $1\alpha,25$ -dihydroxyvitamine D3 and TGF β -induced maturation (100, 312). The role of GLIPR1 within this setting and its role in macrophage function have not been extensively explored. No immune defects have been reported in the *Glipr1* knock-out mouse line to date (261).

VIII. CAP Protein in Nonreproductive Tract Pathologies

A. Sjögren's syndrome

CRISP3 in particular has been identified as a differentially regulated gene product in several pathologies; however, it is unknown whether this is a cause or consequence of disease. CRISP3 is expressed in an androgen-dependent manner in both the submandibular salivary glands and lacrimal glands of the healthy mouse and human (45, 58, 85, 98). Indeed, within saliva from healthy women, CRISP3 protein is produced in the demilunes of acini and secreted into saliva at an average rate of 97.6 μg per 15 min in a resting state, and at a rate of 240.1 μg per 15 min after dehydroepiandrosterone stimulation (98). Women with the autoimmune disease Sjögren's syndrome produce significantly less (4- to 5-fold) CRISP3 than age-matched controls (98). Sjögren's syndrome is a disease of exocrine glands (reviewed in Ref. 97) that is most frequently manifested as dry eyes and mouth. The timing of onset is usually in the peri- or postmenopausal

years, concurrent with the drop in both systemic estrogens and androgens (reviewed in Ref. 97). Indeed, the production of CRISP3 in humans is regulated by local androgen levels, including dehydroepiandrosterone, and is likely to explain the notable patient gender bias in Sjögren's syndrome (9:1, females to males) (98). Because Sjögren's syndrome is characterized by disturbed ion channel (principally aquaporin-5) distribution and function, it is tempting to speculate that this involves the ion channel regulatory activity of the ICR region of CRISP3; however, to date this has not been examined.

B. Chronic pancreatitis

Chronic pancreatitis results in a significant up-regulation of CRISP3 (~21-fold) (313). Chronic pancreatitis is an inflammatory disease characterized by progressive and irreversible destruction of the pancreas, and it has both environmental, *e.g.*, alcohol consumption, and genetic components (314). *In situ* hybridization and immunohistochemical analyses showed a distinct up-regulation of CRISP3 in the cytoplasm of degenerating acinar cells in chronic pancreatitis tissues, compared with adjacent normal acinar cells and to acinar cells of the normal pancreas (95). The significance of these observations is currently unknown.

C. Neuropathic pain

Crisp1 up-regulation has been observed in relation to neuronal pathology, specifically in a rat model of partial sciatic nerve injury leading to neuropathic pain. *Crisp1* was identified as a differentially regulated gene within the dorsal root ganglion cells associated with the perception of pain (315). Partial injury of the nervous system through, for example, spinal cord injury, multiple sclerosis, stroke, infection, toxins, or genetic or immune-mediated disorders can result in the development of neuropathic pain (316, 317). Using the rat model, an analysis of expression profiles within the dorsal root ganglion 2 wk after partial sciatic nerve injury identified the expression of *Crisp1* (318). *In situ* hybridization of dorsal root ganglion cross-sections showed an abundance of transcripts skewed toward the smallest cells with a high proportion being nociceptive (those that respond to high intensity, painful stimulus). Although it remains to be confirmed that CRISP1 protein is produced in nociceptive cells, the aberrant expression of a potential ICR in conditions of nerve damage may result in the inappropriate transmission of sensory information. Mechanisms controlling *Crisp1* expression under these conditions are unknown; however, androgen receptor expression in dorsal root ganglia has been demonstrated (319). The pain response of *Crisp1* null animals has not as yet been studied.

D. Heart disease

As outlined earlier, PI16 has been identified as a secreted protein involved in the regulation of cardiomyocyte size (122). In a mouse model of heart failure, PI16 was shown to be highly up-regulated ($470 \pm 55\%$). A similar increase was observed in damaged human heart tissue ($323 \pm 130\%$) (122). PI16 production by rat cardiomyocytes *in vitro* significantly inhibited their growth. RNA interference-mediated inhibi-

tion results in increased cardiomyocyte size, and the *in vivo* production of PI16 in a transgenic mouse model resulted in apparently healthy mice with small hearts made up of hypertrophic cardiomyocytes (122). Interestingly, cardiomyocytes transfected with small interfering RNA against PI16 are larger and showed reorganized sarcomeres (the basic muscle cell unit between Z-bands) (122). As a consequence of its secretion, the presence of a CAP domain with homology to a protease (132), and also homology to the human PI15 protease inhibitor, Frost and Engelhardt (122) hypothesized that PI16 may have a role in the matrix remodeling component of heart failure (320).

E. Bronchopulmonary dysplasia

Consistent with the proposed role of CAPLD2 in lung branching morphogenesis, it is significantly down-regulated in rat models of bronchopulmonary dysplasia (127). Bronchopulmonary dysplasia is a chronic lung disease most often observed in severely underweight premature children. Its pathology appears to be due, at least in part, to abnormal alveolar and airway development or inflammation (321). Within both the normal and treated rat, CAPLD2 was associated with the growing alveolar tips during periods of rapid growth (127). Such a localization is consistent with a role for CAPLD2 as a regulator of the extracellular matrix. The biochemistry underlying CAPLD2 function is currently unknown.

IX. Conclusions and Perspectives

The CAP superfamily of proteins is extremely diverse, spanning the most ancient of organisms to the most advanced mammals. This diversity suggests involvement in fundamental biological processes. The conservation of several sequence motifs in the CAP domain throughout evolutionary history suggests highly conserved functions, however, these are largely buried within the core of the tertiary structure and involved in maintaining overall protein shape. Conserved solvent exposed amino acids are largely restricted to a putative active site that is suggestive of an enzymatic function. Apart from these motifs, the CAP superfamily proteins have undergone a significant diversification of the CAP domain among subfamilies, particularly in surface exposed amino acids important for interactions. In addition, many have acquired C-terminal extensions or domains that can impart functionality in their own rite. This diversity within mammals, whereby nine distinct subfamilies of the CAP superfamily are identified, points to a highly adaptive evolutionary history.

CAP superfamily proteins represent striking cases of mosaic evolution, with both evolutionary stasis (in CAP motifs) and change (throughout the rest of the protein). These two apparently opposing evolutionary patterns have resulted in a superfamily with conserved tertiary structure, yet a diversity of function. This variability may be at the core of the difficulty in defining the activity/function of the CAP superfamily. In numerous cases, specific high-affinity interactions have been described, and these may provide the basis for a competitive negative regulation by the CAPs. If an

enzymatic function for CAP domains is confirmed (*e.g.*, protease activity), the specific targets of that function may be different based on the different protein interactions controlled through variation in surface amino acids and potentially associated domains. An example of this is provided by NIF and HPI, activation-associated proteins from the nematode. Both of these bind integrins, yet their target is different resulting, in differing biological functions. NIF inhibits neutrophil adhesion, and HPI inhibits platelet aggregation (269, 270).

In the most general terms, CAP superfamily proteins are identified in three environments: 1) as a normal component in a cell or tissue; 2) as being up-regulated during disease; and 3) as being introduced by a foreign agent for a toxic/immune evasive effect. Within this review we have been concerned primarily with the two former situations, although examples of the latter have been provided and are useful in consideration of the former. In numerous cases where CAP expression is chronically up-regulated, an association between extracellular matrix remodeling, immune system regulation/dysfunction, or otherwise general disruption of cellular homeostasis is observed. Determination of the function of CAPs in these instances of disease will undoubtedly be the focus of future investigation.

Some aspects of CAP protein expression appear contradictory because they are up-regulated in cases of chronic disease and cancer where positive correlations have been drawn, yet are also present in the normal healthy adult. Although this in itself is not extraordinary, in many cases there appears to be a dichotomy in function. An example of this is borne out through the CRISP subfamily. CRISPs are highly expressed and conserved within the mammalian male reproductive tract, but also within the venom of poisonous reptiles (322, 323). Their retention implies an important function in each of these environments, which are at first glance, opposing; one is important for the generation of life, and the other has a toxic effect potentially leading to death. In reality, both processes are critically dependent on ion channel regulation.

An accumulating body of evidence suggests an important role for CAPs in mammalian biology, specifically relating to reproductive function, the immune system, cancer invasiveness, and numerous chronic diseases, organogenesis, and development. Some very significant questions remain—specifically, and most notably, what is the underlying mode of action of the CAP domain, an enzyme or high-affinity interacting protein functioning via competitive inhibition? Characterization of this activity will be of enormous benefit to understanding their biological significance in the diverse environments in which they are expressed and may offer new treatment possibilities. For the readers of this journal, the function of CAP proteins is particularly notable for the processes of sperm development, maturation, and function, but also for prostate development and disease. The possibilities and potential importance of the superfamily is further highlighted when examining the many uncharacterized members of the GLIPR1, GLIPR2, CAPLD1, CAPLD2, PI15, and PI16 subfamilies. The majority are highly represented in endocrine-regulated tissues and disease.

Acknowledgments

The authors acknowledge Dr. Bryan Fry for the helpful discussions in relation to the phylogenetic relationships of the CAP superfamily proteins, and Anne O'Connor for review of this manuscript.

Received July 14, 2008. Accepted September 4, 2008.

Address all correspondence and requests for reprints to: Dr. Gerard M. Gibbs, Monash Institute of Medical Research, Monash University, 27-31 Wright Street, Clayton 3168, Australia. E-mail: gerard.gibbs@med.monash.edu.au.

This work was supported by research grants from the Australian Research Council and the National Health and Medical Research Council of Australia. M.K.O. is the recipient of a National Health and Medical Research Council Senior Research Fellowship. K.R. received a postdoctoral fellowship from the Fund for Scientific Research-Flanders (FWO-Vlaanderen).

Disclosure Summary: The authors have nothing to disclose.

References

- Schreiber MC, Karlo JC, Kovalick GE 1997 A novel cDNA from *Drosophila* encoding a protein with similarity to mammalian cysteine-rich secretory proteins, wasp venom antigen 5, and plant group 1 pathogenesis-related proteins. *Gene* 191:135–141
- Mizuki N, Sarapata DE, Garcia-Sanz JA, Kasahara M 1992 The mouse male germ cell-specific gene Tpx-1: molecular structure, mode of expression in spermatogenesis, and sequence similarity to two non-mammalian genes. *Mamm Genome* 3:274–280
- Fang KS, Vitale M, Fehlner P, King TP 1988 cDNA cloning and primary structure of a white-face hornet venom allergen, antigen 5. *Proc Natl Acad Sci USA* 85:895–899
- Cornelissen BJ, van Huijsduijnen RA, Van Loon LC, Bol JF 1986 Molecular characterization of messenger RNAs for 'pathogenesis-related' proteins 1a, 1b and 1c, induced by TMV infection of tobacco. *EMBO J* 5:37–40
- Pfaffner UM, Goodman HM 1987 Isolation and characterization of cDNA clones encoding pathogenesis-related proteins from tobacco mosaic virus infected tobacco plants. *Nucleic Acids Res* 15:4449–4465
- Gibbs GM, O'Bryan MK 2007 Cysteine rich secretory proteins in reproduction and venom. *Soc Reprod Fertil Suppl* 65:261–267
- Yamazaki Y, Morita T 2004 Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* 44:227–231
- King TP, Spangfort MD 2000 Structure and biology of stinging insect venom allergens. *Int Arch Allergy Immunol* 123:99–106
- van Loon LC, Rep M, Pieterse CM 2006 Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 44:135–162
- Szyperki T, Fernandez C, Mumenthaler C, Wuthrich K 1998 Structure comparison of human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein P14a indicates a functional link between the human immune system and a plant defense system. *Proc Natl Acad Sci USA* 95:2262–2266
- Henriksen A, King TP, Mirza O, Monsalve RI, Meno K, Ipsen H, Larsen JN, Gajhede M, Spangfort MD 2001 Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily. *Proteins* 45:438–448
- Kovalick GE, Griffin DL 2005 Characterization of the SCP/TAPS gene family in *Drosophila melanogaster*. *Insect Biochem Mol Biol* 35:825–835
- Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL, Bateman A 2006 Pfam: clans, web tools and services. *Nucleic Acids Res* 34:D247–D251
- Guo M, Teng M, Niu L, Liu Q, Huang Q, Hao Q 2005 Crystal structure of cysteine-rich secretory protein stecrisp reveals the cysteine-rich domain has a K⁺-channel inhibitor-like fold. *J Biol Chem* 280:12405–12412
- Gibbs GM, Scanlon MJ, Swarbrick J, Curtis S, Gallant E, Dulhunty AF, O'Bryan MK 2006 The cysteine-rich secretory protein domain of Tpx-1 is related to ion channel toxins and regulates ryanodine receptor Ca²⁺ signaling. *J Biol Chem* 281:4156–4163
- Zelensky AN, Gready JE 2005 The C-type lectin-like domain superfamily. *FEBS J* 272:6179–6217
- Trexler M, Banyai L, Patthy L 2000 The LCCL module. *Eur J Biochem* 267:5751–5757
- Cameo MS, Blaquier JA 1976 Androgen-controlled specific proteins in rat epididymis. *J Endocrinol* 69:47–55
- Lea A, Petrusz P, French FS 1978 Purification and localization of acidic epididymal glycoprotein (AEG): a sperm coating protein secreted by the rat epididymis. *Int J Androl Vol 1 (Suppl 2B)*: 592–607
- Faye JC, Duguet L, Mazzuca M, Bayard F 1980 Purification, radioimmunoassay, and immunohistochemical localization of a glycoprotein produced by the rat epididymis. *Biol Reprod* 23:423–432
- Wong PY, Tsang AY 1982 Studies on the binding of a 32K rat epididymal protein to rat epididymal spermatozoa. *Biol Reprod* 27:1239–1246
- Jones R, Brown CR 1982 Association of epididymal secretory proteins showing α -lactalbumin-like activity with the plasma membrane of rat spermatozoa. *The Biochemical journal* 206:161–164
- Charest NJ, Joseph DR, Wilson EM, French FS 1988 Molecular cloning of complementary deoxyribonucleic acid for an androgen-regulated epididymal protein: sequence homology with metalloproteins. *Mol Endocrinol* 2:999–1004
- Brooks DE, Means AR, Wright EJ, Singh SP, Tiver KK 1986 Molecular cloning of the cDNA for androgen-dependent sperm-coating glycoproteins secreted by the rat epididymis. *Eur J Biochem* 161:13–18
- Haendler B, Kratzschmar J, Theuring F, Schleuning WD 1993 Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/AEG) and the novel related CRISP-3 are expressed under androgen control in the mouse salivary gland. *Endocrinology* 133:192–198
- Giese A, Jude R, Kuiper H, Piumi F, Schambony A, Guerin G, Distl O, Topfer-Petersen E, Leeb T 2002 Molecular characterization of the equine AEG1 locus. *Gene* 292:65–72
- Rankin TL, Tsuruta KJ, Holland MK, Griswold MD, Orgebin-Crist MC 1992 Isolation, immunolocalization, and sperm-association of three proteins of 18, 25, and 29 kilodaltons secreted by the mouse epididymis. *Biol Reprod* 46:747–766
- Hardy DM, Huang Jr TT, Driscoll WJ, Tung KK, Wild GC 1988 Purification and characterization of the primary acrosomal autoantigen of guinea pig epididymal spermatozoa. *Biol Reprod* 38:423–437
- Foster JA, Gerton GL 1996 Autoantigen 1 of the guinea pig sperm acrosome is the homologue of mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich secretory protein (CRISP) family. *Mol Reprod Dev* 44:221–229
- Kasahara M, Gutknecht J, Brew K, Spurr N, Goodfellow PN 1989 Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene. *Genomics* 5:527–534
- Maeda T, Sakashita M, Ohba Y, Nakanishi Y 1998 Molecular cloning of the rat Tpx-1 responsible for the interaction between spermatogenic and Sertoli cells. *Biochem Biophys Res Commun* 248:140–146
- O'Bryan MK, Loveland KL, Herszfeld D, McFarlane JR, Hearn MT, de Kretser DM 1998 Identification of a rat testis-specific gene encoding a potential rat outer dense fibre protein. *Mol Reprod Dev* 50:313–322
- Kratzschmar J, Haendler B, Eberspaecher U, Roosterman D, Donner P, Schleuning WD 1996 The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3. *Eur J Biochem* 236:827–836
- Giese A, Jude R, Kuiper H, Raudsepp T, Piumi F, Schambony A, Guerin G, Chowdhary BP, Distl O, Topfer-Petersen E, Leeb T 2002 Molecular characterization of the equine testis-specific protein 1 (TPX1) and acidic epididymal glycoprotein 2 (AEG2) genes encoding members of the cysteine-rich secretory protein (CRISP) family. *Gene* 299:101–109
- Kjeldsen L, Cowland JB, Johnsen AH, Borregaard N 1996 SGP28, a novel matrix glycoprotein in specific granules of human neutrophils with similarity to a human testis-specific gene product and a rodent sperm-coating glycoprotein. *FEBS Lett* 380:246–250

36. Magdaleno L, Gasset M, Varea J, Schambony AM, Urbanke C, Raida M, Topfer-Petersen E, Calvete JJ 1997 Biochemical and conformational characterisation of HSP-3, a stallion seminal plasma protein of the cysteine-rich secretory protein (CRISP) family. *FEBS Lett* 420:179–185
37. Schambony A, Gentzel M, Wolfes H, Raida M, Neumann U, Topfer-Petersen E 1998 Equine CRISP-3: primary structure and expression in the male genital tract. *Biochim Biophys Acta* 1387: 206–216
38. Jalkanen J, Huhtaniemi I, Poutanen M 2005 Mouse cysteine-rich secretory protein 4 (CRISP4): a member of the crisp family exclusively expressed in the epididymis in an androgen-dependent manner. *Biol Reprod* 72:1268–1274
39. Nolan MA, Wu L, Bang HJ, Jelinsky SA, Roberts KP, Turner TT, Kopf GS, Johnston DS 2006 Identification of rat cysteine-rich secretory protein 4 (Crisp4) as the ortholog to human CRISP1 and mouse Crisp4. *Biol Reprod* 74:984–991
40. Eberspaecher U, Roosterman D, Kratzschmar J, Haendler B, Habenicht UF, Becker A, Quensel C, Petri T, Schleuning WD, Donner P 1995 Mouse androgen-dependent epididymal glycoprotein CRISP-1 (DE/AEG): isolation, biochemical characterization, and expression in recombinant form. *Mol Reprod Dev* 42:157–172
41. Shikamoto Y, Suto K, Yamazaki Y, Morita T, Mizuno H 2005 Crystal structure of a CRISP family Ca^{2+} -channel blocker derived from snake venom. *J Mol Biol* 350:735–743
42. Wang J, Shen B, Guo M, Lou X, Duan Y, Cheng XP, Teng M, Niu L, Liu Q, Huang Q, Hao Q 2005 Blocking effect and crystal structure of natriin toxin, a cysteine-rich secretory protein from *Naja atra* venom that targets the BKCa channel. *Biochemistry* 44:10145–10152
43. Roberts KP, Johnston DS, Nolan MA, Wooters JL, Waxmonsky NC, Piehl LB, Ensrud-Bowlin KM, Hamilton DW 2007 Structure and function of epididymal protein cysteine-rich secretory protein-1. *Asian J Androl* 9:508–514
44. Hayashi M, Fujimoto S, Takano H, Ushiki T, Abe K, Ishikura H, Yoshida MC, Kirchhoff C, Ishibashi T, Kasahara M 1996 Characterization of a human glycoprotein with a potential role in sperm-egg fusion: cDNA cloning, immunohistochemical localization, and chromosomal assignment of the gene (AEG1). *Genomics* 32:367–374
45. Reddy T, Gibbs GM, Kerr J, O'Bryan MK 2008 Cysteine-rich secretory proteins are not exclusively expressed in the male reproductive tract. *Dev Dyn* 237:3313–3323
46. de Kretser DM, Kerr JB 1994 The cytology of the testis. In: Knobil E, Neill D, eds. *The physiology of reproduction*. New York: Raven Press; 1177–1290
47. Fraser LR 1992 Requirements for successful mammalian sperm capacitation and fertilization. *Arch Pathol Lab Med* 116:345–350
48. Fawcett DW 1975 The mammalian spermatozoon. *Dev Biol* 44: 394–436
49. Eddy EM, Toshimori K, O'Brien DA 2003 Fibrous sheath of mammalian spermatozoa. *Microsc Res Tech* 61:103–115
50. Haendler B, Habenicht UF, Schwidetzky U, Schuttke I, Schleuning WD 1997 Differential androgen regulation of the murine genes for cysteine-rich secretory proteins (CRISP). *Eur J Biochem* 250:440–446
51. O'Bryan MK, Sebire K, Meinhardt A, Edgar K, Keah HH, Hearn MT, de Kretser DM 2001 Tpx-1 is a component of the outer dense fibers and acrosome of rat spermatozoa. *Mol Reprod Dev* 58: 116–125
52. Jamsai D, Reilly A, Smith SJ, Gibbs GM, Baker HW, McLachlan RI, de Kretser DM, O'Bryan MK 2008 Polymorphisms in the human cysteine-rich secretory protein 2 (CRISP2) gene in Australian men. *Hum Reprod* 23:2151–2159
53. Maeda T, Nishida J, Nakanishi Y 1999 Expression pattern, subcellular localization and structure-function relationship of rat Tpx-1, a spermatogenic cell adhesion molecule responsible for association with Sertoli cells. *Dev Growth Differ* 41:715–722
54. Jiao X, Trifillis P, Kiledjian M 2002 Identification of target messenger RNA substrates for the murine deleted in azoospermia-like RNA-binding protein. *Biol Reprod* 66:475–485
55. Busso D, Cohen DJ, Hayashi M, Kasahara M, Cuasnicu PS 2005 Human testicular protein TPX1/CRISP-2: localization in spermatozoa, fate after capacitation and relevance for gamete interaction. *Mol Hum Reprod* 11:299–305
56. Jamsai D, Bianco DM, Smith SJ, Merriner DJ, Ly-Huynh JD, Herlihy A, Niranjan B, Gibbs GM, O'Bryan MK 2008 Characterization of gametogenetin 1 (GGN1) and its potential role in male fertility through the interaction with the ion channel regulator, cysteine-rich secretory protein 2 (CRISP2) in the sperm tail. *Reproduction* 135:751–759
57. Schalles U, Shao X, van der Hoorn FA, Oko R 1998 Developmental expression of the 84-kDa ODF sperm protein: localization to both the cortex and medulla of outer dense fibers and to the connecting piece. *Dev Biol* 199:250–260
58. Haendler B, Toda I, Sullivan DA, Schleuning WD 1999 Expression of transcripts for cysteine-rich secretory proteins (CRISPs) in the murine lacrimal gland. *J Cell Physiol* 178:371–378
59. Udby L, Bjartell A, Malm J, Egesten A, Lundwall A, Cowland JB, Borregaard N, Kjeldsen L 2005 Characterization and localization of cysteine-rich secretory protein 3 (CRISP-3) in the human male reproductive tract. *J Androl* 26:333–342
60. Da Ros V, Maldera J, Willis W, Cohen D, Goulding E, Gelman D, Rubinstein M, Eddy E, Cuasnicu P 2008 Impaired sperm fertilizing ability in mice lacking Cysteine-Rich Secretory Protein 1 (CRISP1). *Dev Biol* 320:12–18
61. Sullivan R, Frenette G, Girouard J 2007 Epididymosomes are involved in the acquisition of new sperm proteins during epididymal transit. *Asian J Androl* 9:483–491
62. Thimon V, Frenette G, Saez F, Thabet M, Sullivan R 2008 Protein composition of human epididymosomes collected during surgical vasectomy reversal: a proteomic and genomic approach. *Hum Reprod* 23:1698–1707
63. Schwidetzky U, Schleuning WD, Haendler B 1997 Isolation and characterization of the androgen-dependent mouse cysteine-rich secretory protein-1 (CRISP-1) gene. *Biochem J* 321:325–332
64. Roberts KP, Ensrud KM, Hamilton DW 2002 A comparative analysis of expression and processing of the rat epididymal fluid and sperm-bound forms of proteins D and E. *Biol Reprod* 67:525–533
65. Tubbs CE, Hall JC, Scott RO, Clark VP, Hermon TL, Bazemore-Walker C 2002 Binding of protein D/E to the surface of rat epididymal sperm before ejaculation and after deposition in the female reproductive tract. *J Androl* 23:512–521
66. Kohane AC, Pineiro L, Blaquier JA 1983 Androgen-controlled synthesis of specific proteins in the rat epididymis. *Endocrinology* 112:1590–1596
67. Roberts KP, Wamstad JA, Ensrud KM, Hamilton DW 2003 Inhibition of capacitation-associated tyrosine phosphorylation signaling in rat sperm by epididymal protein Crisp-1. *Biol Reprod* 69: 572–581
68. Brooks DE 1982 Purification of rat epididymal proteins "D" and "E", demonstration of shared immunological determinants, and identification of regional synthesis and secretion. *Int J Androl* 5:513–524
69. Garrett SH, Garrett JE, Douglass J 1991 In situ histochemical analysis of region-specific gene expression in the adult rat epididymis. *Mol Reprod Dev* 30:1–17
70. Xu W, Hamilton DW 1996 Identification of the rat epididymis-secreted 4E9 antigen as protein E: further biochemical characterization of the highly homologous epididymal secretory proteins D and E. *Mol Reprod Dev* 43:347–357
71. Xu W, Ensrud KM, Hamilton DW 1997 The 26 kD protein recognized on rat cauda epididymal sperm by monoclonal antibody 4E9 has internal peptide sequence that is identical to the secreted form of epididymal protein E. *Mol Reprod Dev* 46:377–382
72. Lea OA, French FS 1981 Characterization of an acidic glycoprotein secreted by principal cells of the rat epididymis. *Biochim Biophys Acta* 668:370–376
73. Klemme LM, Roberts KP, Hoffman LB, Ensrud KM, Siiteri JE, Hamilton DW 1999 Cloning and characterization of the rat Crisp-1 gene. *Gene* 240:279–288
74. Roberts KP, Hoffman LB, Ensrud KM, Hamilton DW 2001 Expression of crisp-1 mRNA splice variants in the rat epididymis, and comparative analysis of the rat and mouse crisp-1 gene regulatory regions. *J Androl* 22:157–163
75. Roberts KP, Ensrud KM, Wooters JL, Nolan MA, Johnston DS,

- Hamilton DW 2006 Epididymal secreted protein Crisp-1 and sperm function. *Mol Cell Endocrinol* 250:122–127
76. Cohen DJ, Da Ros VG, Busso D, Ellerman DA, Maldera JA, Goldweic N, Cuasnicu PS 2007 Participation of epididymal cysteine-rich secretory proteins in sperm-egg fusion and their potential use for male fertility regulation. *Asian J Androl* 9:528–532
 77. Rochwerger L, Cuasnicu PS 1992 Redistribution of a rat sperm epididymal glycoprotein after in vitro and in vivo capacitation. *Mol Reprod Dev* 31:34–41
 78. Cohen DJ, Rochwerger L, Ellerman DA, Morgenfeld MM, Busso D, Cuasnicu PS 2000 Relationship between the association of rat epididymal protein “DE” with spermatozoa and the behavior and function of the protein. *Mol Reprod Dev* 56:180–188
 79. Udby L, Lundwall A, Johnsen AH, Fernlund P, Valtonen-Andre C, Blom AM, Lilja H, Borregaard N, Kjeldsen L, Bjartell A 2005 β -Microseminoprotein binds CRISP-3 in human seminal plasma. *Biochem Biophys Res Commun* 333:555–561
 80. Reeves JR, Xuan JW, Arfanis K, Morin C, Garde SV, Ruiz MT, Wisniewski J, Panchal C, Tanner JE 2005 Identification, purification and characterization of a novel human blood protein with binding affinity for prostate secretory protein of 94 amino acids. *Biochem J* 385:105–114
 81. Udby L, Calafat J, Sorensen OE, Borregaard N, Kjeldsen L 2002 Identification of human cysteine-rich secretory protein 3 (CRISP-3) as a matrix protein in a subset of peroxidase-negative granules of neutrophils and in the granules of eosinophils. *J Leukoc Biol* 72:462–469
 82. Udby L, Cowland JB, Johnsen AH, Sorensen OE, Borregaard N, Kjeldsen L 2002 An ELISA for SGP28/CRISP-3, a cysteine-rich secretory protein in human neutrophils, plasma, and exocrine secretions. *J Immunol Methods* 263:43–55
 83. Pfisterer P, Konig H, Hess J, Lipowsky G, Haendler B, Schleuning WD, Wirth T 1996 CRISP-3, a protein with homology to plant defense proteins, is expressed in mouse B cells under the control of Oct2. *Mol Cell Biol* 16:6160–6168
 84. Mizuki N, Kasahara M 1992 Mouse submandibular glands express an androgen-regulated transcript encoding an acidic epididymal glycoprotein-like molecule. *Mol Cell Endocrinol* 89:25–32
 85. Schwidetzky U, Haendler B, Schleuning WD 1995 Isolation and characterization of the androgen-dependent mouse cysteine-rich secretory protein-3 (CRISP-3) gene. *Biochem J* 309:831–836
 86. Udby L, Sorensen OE, Pass J, Johnsen AH, Behrendt N, Borregaard N, Kjeldsen L 2004 Cysteine-rich secretory protein 3 is a ligand of α 1B-glycoprotein in human plasma. *Biochemistry* 43:12877–12886
 87. Jurgilas PB, Neves-Ferreira AG, Domont GB, Perales J 2003 PO41, a snake venom metalloproteinase inhibitor isolated from *Philander opossum* serum. *Toxicon* 42:621–628
 88. Rocha SL, Lomonte B, Neves-Ferreira AG, Trugilho MR, Junqueira-de-Azevedo Ide L, Ho PL, Domont GB, Gutierrez JM, Perales J 2002 Functional analysis of DM64, an antimyotoxic protein with immunoglobulin-like structure from *Didelphis marsupialis* serum. *Eur J Biochem* 269:6052–6062
 89. Neves-Ferreira AG, Perales J, Fox JW, Shannon JD, Makino DL, Garratt RC, Domont GB 2002 Structural and functional analyses of DM43, a snake venom metalloproteinase inhibitor from *Didelphis marsupialis* serum. *J Biol Chem* 277:13129–13137
 90. Catanese JJ, Kress LF 1992 Isolation from opossum serum of a metalloproteinase inhibitor homologous to human α 1B-glycoprotein. *Biochemistry* 31:410–418
 91. Asmann YW, Kosari F, Wang K, Cheville JC, Vasmataz G 2002 Identification of differentially expressed genes in normal and malignant prostate by electronic profiling of expressed sequence tags. *Cancer Res* 62:3308–3314
 92. Kosari F, Asmann YW, Cheville JC, Vasmataz G 2002 Cysteine-rich secretory protein-3: a potential biomarker for prostate cancer. *Cancer Epidemiol Biomarkers Prev* 11:1419–1426
 93. Bjartell A, Johansson R, Bjork T, Gadaleanu V, Lundwall A, Lilja H, Kjeldsen L, Udby L 2006 Immunohistochemical detection of cysteine-rich secretory protein 3 in tissue and in serum from men with cancer or benign enlargement of the prostate gland. *Prostate* 66:591–603
 94. Ye H, Yu T, Temam S, Ziober BL, Wang J, Schwartz JL, Mao L, Wong DT, Zhou X 2008 Transcriptomic dissection of tongue squamous cell carcinoma. *BMC Genomics* 9:69
 95. Liao Q, Kleeff J, Xiao Y, Guweidhi A, Schambony A, Topfer-Petersen E, Zimmermann A, Buchler MW, Friess H 2003 Preferential expression of cysteine-rich secretory protein-3 (CRISP-3) in chronic pancreatitis. *Histol Histopathol* 18:425–433
 96. Tapinos NI, Polihronis M, Thyphronitis G, Moutsopoulos HM 2002 Characterization of the cysteine-rich secretory protein 3 gene as an early-transcribed gene with a putative role in the pathophysiology of Sjogren’s syndrome. *Arthritis Rheum* 46:215–222
 97. Kontinen YT, Porola P, Kontinen L, Laine M, Poduval P 2006 Immunohistopathology of Sjogren’s syndrome. *Autoimmunity Rev* 6:16–20
 98. Laine M, Porola P, Udby L, Kjeldsen L, Cowland JB, Borregaard N, Hietanen J, Stahle M, Pihakari A, Kontinen YT 2007 Low salivary dehydroepiandrosterone and androgen-regulated cysteine-rich secretory protein 3 levels in Sjogren’s syndrome. *Arthritis Rheum* 56:2575–2584
 99. Ren C, Ren CH, Li L, Goltsov AA, Thompson TC 2006 Identification and characterization of RTVP1/GLIPR1-like genes, a novel p53 target gene cluster. *Genomics* 88:163–172
 100. Murphy EV, Zhang Y, Zhu W, Biggs J 1995 The human glioma pathogenesis-related protein is structurally related to plant pathogenesis-related proteins and its gene is expressed specifically in brain tumors. *Gene* 159:131–135
 101. Rich T, Chen P, Furman F, Huynh N, Israel MA 1996 RTVP-1, a novel human gene with sequence similarity to genes of diverse species, is expressed in tumor cell lines of glial but not neuronal origin. *Gene* 180:125–130
 102. Ren C, Li L, Goltsov AA, Timme TL, Tahir SA, Wang J, Garza L, Chinault AC, Thompson TC 2002 mRTVP-1, a novel p53 target gene with proapoptotic activities. *Mol Cell Biol* 22:3345–3357
 103. Ren C, Li L, Yang G, Timme TL, Goltsov A, Ren C, Ji X, Addai J, Luo H, Ittmann MM, Thompson TC 2004 RTVP-1, a tumor suppressor inactivated by methylation in prostate cancer. *Cancer Res* 64:969–976
 104. Rosenzweig T, Ziv-Av A, Xiang C, Lu W, Cazacu S, Taler D, Miller CG, Reich R, Shoshan Y, Anikster Y, Kazimirsky G, Sarid R, Brodie C 2006 Related to testes-specific, vespid, and pathogenesis protein-1 (RTVP-1) is overexpressed in gliomas and regulates the growth, survival, and invasion of glioma cells. *Cancer Res* 66:4139–4148
 105. Xiang C, Sarid R, Cazacu S, Finniss S, Lee HK, Ziv-Av A, Mikkelsen T, Brodie C 2007 Cloning and characterization of human RTVP-1b, a novel splice variant of RTVP-1 in glioma cells. *Biochem Biophys Res Commun* 362:612–618
 106. Miosga T, Schaaff-Gerstenschlager I, Chalwatzis N, Baur A, Boles E, Fournier C, Schmitt S, Velten C, Wilhelm N, Zimmermann FK 1995 Sequence analysis of a 33.1 kb fragment from the left arm of *Saccharomyces cerevisiae* chromosome X, including putative proteins with leucine zippers, a fungal Zn(II)-Cys6 binuclear cluster domain and a putative α 2-SCB- α 2 binding site. *Yeast* 11:681–689
 107. Eberle HB, Serrano RL, Fullekrug J, Schlosser A, Lehmann WD, Lottspeich F, Kaloyanova D, Wieland FT, Helms JB 2002 Identification and characterization of a novel human plant pathogenesis-related protein that localizes to lipid-enriched microdomains in the Golgi complex. *J Cell Sci* 115:827–838
 108. Eisenberg I, Barash M, Kahan T, Mitrani-Rosenbaum S 2002 Cloning and characterization of a human novel gene C9orf19 encoding a conserved putative protein with an SCP-like extracellular protein domain. *Gene* 293:141–148
 109. Baxter RM, Crowell TP, George JA, Getman ME, Gardner H 2007 The plant pathogenesis related protein GLIPR-2 is highly expressed in fibrotic kidney and promotes epithelial to mesenchymal transition in vitro. *Matrix Biol* 26:20–29
 110. Kisseleva T, Brenner DA 2008 Mechanisms of fibrogenesis. *Exp Biol Med* (Maywood) 233:109–122
 111. Powell DW, Mifflin RC, Valentich JD, Crowe SE, Saada JI, West AB 1999 Myofibroblasts. I. Paracrine cells important in health and disease. *Am J Physiol* 277:C1–C9
 112. Desmouliere A, Chaponnier C, Gabbiani G 2005 Tissue repair, contraction, and the myofibroblast. *Wound Repair Regen* 13:7–12

113. Kalluri R, Neilson EG 2003 Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest* 112:1776–1784
114. Koshikawa N, Nakamura T, Tsuchiya N, Isaji M, Yasumitsu H, Umeda M, Miyazaki K 1996 Purification and identification of a novel and four known serine proteinase inhibitors secreted by human glioblastoma cells. *J Biochem* 119:334–339
115. Yamakawa T, Miyata S, Ogawa N, Koshikawa N, Yasumitsu H, Kanamori T, Miyazaki K 1998 cDNA cloning of a novel trypsin inhibitor with similarity to pathogenesis-related proteins, and its frequent expression in human brain cancer cells. *Biochim Biophys Acta* 1395:202–208
116. Takemoto M, He L, Norlin J, Patrakka J, Xiao Z, Petrova T, Bondjers C, Asp J, Wallgard E, Sun Y, Samuelsson T, Mostad P, Lundin S, Miura N, Sado Y, Alitalo K, Quaggin SE, Tryggvason K, Betsholtz C 2006 Large-scale identification of genes implicated in kidney glomerulus development and function. *EMBO J* 25:1160–1174
117. Smith DM, Collins-Racie LA, Marigo VA, Roberts DJ, Davis NM, Hartmann C, Schweitzer R, LaVallie ER, Gamer L, McCoy J, Tabin CJ 2001 Cloning and expression of a novel cysteine-rich secreted protein family member expressed in thyroid and pancreatic mesoderm within the chicken embryo. *Mech Dev* 102:223–226
118. Gill SE, Pape MC, Leco KJ 2006 Tissue inhibitor of metalloproteinases 3 regulates extracellular matrix–cell signaling during bronchiole branching morphogenesis. *Dev Biol* 298:540–554
119. Gill SE, Pape MC, Khokha R, Watson AJ, Leco KJ 2003 A null mutation for tissue inhibitor of metalloproteinases-3 (Timp-3) impairs murine bronchiole branching morphogenesis. *Dev Biol* 261:313–323
120. Pohl M, Stuart RO, Sakurai H, Nigam SK 2000 Branching morphogenesis during kidney development. *Annu Rev Physiol* 62:595–620
121. Karihaloo A, Nickel C, Cantley LG 2005 Signals which build a tubule. *Nephron Exp Nephrol* 100:e40–e45
122. Frost RJ, Engelhardt S 2007 A secretion trap screen in yeast identifies protease inhibitor 16 as a novel antihypertrophic protein secreted from the heart. *Circulation* 116:1768–1775
123. Reeves JR, Dulude H, Panchal C, Daigneault L, Ramnani DM 2006 Prognostic value of prostate secretory protein of 94 amino acids and its binding protein after radical prostatectomy. *Clin Cancer Res* 12:6018–6022
124. Kaplan F, Ledoux P, Kassamali FQ, Gagnon S, Post M, Koehler D, Deimling J, Sweezey NB 1999 A novel developmentally regulated gene in lung mesenchyme: homology to a tumor-derived trypsin inhibitor. *Am J Physiol* 276:L1027–L1036
125. Oyewumi L, Kaplan F, Gagnon S, Sweezey NB 2003 Antisense oligodeoxynucleotides decrease LGL1 mRNA and protein levels and inhibit branching morphogenesis in fetal rat lung. *Am J Respir Cell Mol Biol* 28:232–240
126. Massaro D, Massaro GD 2002 Invited review: pulmonary alveoli: formation, the “call for oxygen,” and other regulators. *Am J Physiol Lung Cell Mol Physiol* 282:L345–L358
127. Nadeau K, Jankov RP, Tanswell AK, Sweezey NB, Kaplan F 2006 Lgl1 is suppressed in oxygen toxicity animal models of bronchopulmonary dysplasia and normalizes during recovery in air. *Pediatr Res* 59:389–395
128. Quinlan J, Kaplan F, Sweezey N, Goodyer P 2007 LGL1, a novel branching morphogen in developing kidney, is induced by retinoic acid. *Am J Physiol Renal Physiol* 293:F987–F993
129. Batourina E, Choi C, Paragas N, Bello N, Hensle T, Costantini FD, Schuchardt A, Bacallao RL, Mendelsohn CL 2002 Distal ureter morphogenesis depends on epithelial cell remodeling mediated by vitamin A and Ret. *Nat Genet* 32:109–115
130. Batourina E, Gim S, Bello N, Shy M, Clagett-Dame M, Srinivas S, Costantini F, Mendelsohn C 2001 Vitamin A controls epithelial/mesenchymal interactions through Ret expression. *Nat Genet* 27:74–78
131. Pohl M, Sakurai H, Bush KT, Nigam SK 2000 Matrix metalloproteinases and their inhibitors regulate in vitro ureteric bud branching morphogenesis. *Am J Physiol Renal Physiol* 279:F891–F900
132. Milne TJ, Abbenante G, Tyndall JD, Halliday J, Lewis RJ 2003 Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J Biol Chem* 278:31105–31110
133. Chiquet BT, Lidral AC, Stal S, Mulliken JB, Moreno LM, Arcos-Burgos M, Valencia-Ramirez C, Blanton SH, Hecht JT 2007 CRISPLD2: a novel NSCLP candidate gene. *Hum Mol Genet* 16:2241–2248
134. Ezekowitz RA, Sastry K, Bailly P, Warner A 1990 Molecular characterization of the human macrophage mannose receptor: demonstration of multiple carbohydrate recognition-like domains and phagocytosis of yeasts in Cos-1 cells. *J Exp Med* 172:1785–1794
135. McGreal EP, Miller JL, Gordon S 2005 Ligand recognition by antigen-presenting cell C-type lectin receptors. *Curr Opin Immunol* 17:18–24
136. Lee SJ, Evers S, Roeder D, Parlow AF, Risteli J, Risteli L, Lee YC, Feizi T, Langen H, Nussenzweig MC 2002 Mannose receptor-mediated regulation of serum glycoprotein homeostasis. *Science* 295:1898–1901
137. Grishin NV 1998 The R3H motif: a domain that binds single-stranded nucleic acids. *Trends Biochem Sci* 23:329–330
138. Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, Liu S, Borgstahl GE, Hotez PJ 2005 X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. *J Mol Biol* 346:801–814
139. Hansson K, Thamlitz AM, Furie B, Furie BC, Stenflo J 2006 A single γ -carboxyglutamic acid residue in a novel cysteine-rich secretory protein without propeptide. *Biochemistry* 45:12828–12839
140. Qian J, Guo ZY, Chi CW 2008 Cloning and isolation of a Conus cysteine-rich protein homologous to Tex31 but without proteolytic activity. *Acta Biochim Biophys Sin (Shanghai)* 40:174–181
141. Fernandez C, Szyperski T, Bruyere T, Ramage P, Mosinger E, Wuthrich K 1997 NMR solution structure of the pathogenesis-related protein P14a. *J Mol Biol* 266:576–593
142. Serrano RL, Kuhn A, Hendricks A, Helms JB, Sinning I, Groves MR 2004 Structural analysis of the human Golgi-associated plant pathogenesis related protein GAPR-1 implicates dimerization as a regulatory mechanism. *J Mol Biol* 339:173–183
143. Wang F, Li H, Liu MN, Song H, Han HM, Wang QL, Yin CC, Zhou YC, Qi Z, Shu YY, Lin ZJ, Jiang T 2006 Structural and functional analysis of natrin, a venom protein that targets various ion channels. *Biochem Biophys Res Commun* 351:443–448
144. Brown RL, Haley TL, West KA, Crabb JW 1999 Pseudechotoxin: a peptide blocker of cyclic nucleotide-gated ion channels. *Proc Natl Acad Sci USA* 96:754–759
145. Yamazaki Y, Brown RL, Morita T 2002 Purification and cloning of toxins from elapid venoms that target cyclic nucleotide-gated ion channels. *Biochemistry* 41:11331–11337
146. Visse R, Nagase H 2003 Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 92:827–839
147. Perona JJ, Craik CS 1997 Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J Biol Chem* 272:29987–29990
148. Perona JJ, Craik CS 1995 Structural basis of substrate specificity in the serine proteases. *Protein Sci* 4:337–360
149. Ellerman DA, Cohen DJ, Da Ros VG, Morgenfeld MM, Busso D, Cuasnicu PS 2006 Sperm protein “DE” mediates gamete fusion through an evolutionarily conserved site of the CRISP family. *Dev Biol* 297:228–237
150. Gibbs GM, Bianco DM, Jamsai D, Herlihy A, Ristevski S, Aitken RJ, Kretser DM, O’Byrne MK 2007 Cysteine-rich secretory protein 2 binds to mitogen-activated protein kinase kinase 11 in mouse sperm. *Biol Reprod* 77:108–114
151. Mochca-Morales J, Martin BM, Possani LD 1990 Isolation and characterization of helothermine, a novel toxin from *Heloderma horridum horridum* (Mexican beaded lizard) venom. *Toxicon* 28:299–309
152. Nobile M, Magnelli V, Lagostena L, Mochca-Morales J, Possani LD, Prestipino G 1994 The toxin helothermine affects potassium currents in newborn rat cerebellar granule cells. *J Membr Biol* 139:49–55
153. Nobile M, Noceti F, Prestipino G, Possani LD 1996 Helothermine,

- a lizard venom toxin, inhibits calcium current in cerebellar granules. *Exp Brain Res* 110:15–20
154. Morrisette J, Kratzschmar J, Haendler B, el-Hayek R, Mochca-Morales J, Martin BM, Patel JR, Moss RL, Schleuning WD, Coronado R, Possani LD 1995 Primary structure and properties of helothermine, a peptide toxin that blocks ryanodine receptors. *Biophys J* 68:2280–2288
155. Brown RL, Lynch LL, Haley TL, Arsanjani R 2003 Pseudochetoxin binds to the pore turret of cyclic nucleotide-gated ion channels. *J Gen Physiol* 122:749–760
156. Tudor JE, Pallaghy PK, Pennington MW, Norton RS 1996 Solution structure of ShK toxin, a novel potassium channel inhibitor from a sea anemone. *Nat Struct Biol* 3:317–320
157. Pennington MW, Mahnir VM, Khaytin I, Zaydenberg I, Byrnes ME, Kem WR 1996 An essential binding surface for ShK toxin interaction with rat brain potassium channels. *Biochemistry* 35:16407–16411
158. Dauplais M, Lecoq A, Song J, Cotton J, Jamin N, Gilquin B, Roumestand C, Vita C, de Medeiros CL, Rowan EG, Harvey AL, Menez A 1997 On the convergent evolution of animal toxins. Conservation of a diad of functional residues in potassium channel-blocking toxins with unrelated structures. *J Biol Chem* 272:4302–4309
159. Ito N, Mita M, Takahashi Y, Matsushima A, Watanabe YG, Hirano S, Odani S 2007 Novel cysteine-rich secretory protein in the buccal gland secretion of the parasitic lamprey, *Lethenteron japonicum*. *Biochem Biophys Res Commun* 358:35–40
160. Hopla CE, Durden LA, Keirans JE 1994 Ectoparasites and classification. *Rev Sci Tech* 13:985–1017
161. Cambi A, Figdor CG 2003 Dual function of C-type lectin-like receptors in the immune system. *Curr Opin Cell Biol* 15:539–546
162. Gupta G, Surolia A 2007 Collectins: sentinels of innate immunity. *Bioessays* 29:452–464
163. Cummings RD, Smith DF 1992 The selectin family of carbohydrate-binding proteins: structure and importance of carbohydrate ligands for cell adhesion. *Bioessays* 14:849–856
164. Barthel SR, Gavino JD, Descheny L, Dimitroff CJ 2007 Targeting selectins and selectin ligands in inflammation and cancer. *Expert Opin Ther Targets* 11:1473–1491
165. Muta T, Miyata T, Misumi Y, Tokunaga F, Nakamura T, Toh Y, Ikehara Y, Iwanaga S 1991 Limulus factor C. An endotoxin-sensitive serine protease zymogen with a mosaic structure of complement-like, epidermal growth factor-like, and lectin-like domains. *J Biol Chem* 266:6554–6561
166. Whittaker CA, Hynes RO 2002 Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol Biol Cell* 13:3369–3387
167. Ahsan M, Ohta K, Kuriyama S, Tanaka H 2005 Novel soluble molecule, Akhirin, is expressed in the embryonic chick eyes and exhibits heterophilic cell-adhesion activity. *Dev Dyn* 233:95–104
168. Delrieu I, Waller CC, Mota MM, Grainger M, Langhorne J, Holder AA 2002 PSLAP, a protein with multiple adhesive motifs, is expressed in *Plasmodium falciparum* gametocytes. *Mol Biochem Parasitol* 121:11–20
169. Robertson NG, Skvorak AB, Yin Y, Weremowicz S, Johnson KR, Kovatch KA, Battey JF, Bieber FR, Morton CC 1997 Mapping and characterization of a novel cochlear gene in human and in mouse: a positional candidate gene for a deafness disorder, DFNA9. *Genomics* 46:345–354
170. de Kok YJ, Bom SJ, Brunt TM, Kemperman MH, van Beusekom E, van der Velde-Visser SD, Robertson NG, Morton CC, Huygen PL, Verhagen WI, Brunner HG, Cremers CW, Cremers FP 1999 A Pro51Ser mutation in the COCH gene is associated with late onset autosomal dominant progressive sensorineural hearing loss with vestibular defects. *Hum Mol Genet* 8:361–366
171. Collin RW, Pauw RJ, Schoots J, Huygen PL, Hoefsloot LH, Cremers CW, Kremer H 2006 Identification of a novel COCH mutation, G87W, causing autosomal dominant hearing impairment (DFNA9). *Am J Med Genet A* 140:1791–1794
172. Pauw RJ, Huygen PL, Collin RW, Cruysberg JR, Hoefsloot LH, Kremer H, Cremers CW 2007 Phenotype description of a novel DFNA9/COCH mutation, I109T. *Ann Otol Rhinol Laryngol* 116:349–357
173. Fransen E, Van Camp G 1999 The COCH gene: a frequent cause of hearing impairment and vestibular dysfunction? *Br J Audiol* 33:297–302
174. Bhattacharya SK, Annangudi SP, Salomon RG, Kuchtey RW, Peachey NS, Crabb JW 2005 Cochlin deposits in the trabecular meshwork of the glaucomatous DBA/2J mouse. *Exp Eye Res* 80:741–744
175. Picciani R, Desai K, Guduric-Fuchs J, Cogliati T, Morton CC, Bhattacharya SK 2007 Cochlin in the eye: functional implications. *Prog Retin Eye Res* 26:453–469
176. Hardy DM, Oda MN, Friend DS, Huang Jr TT 1991 A mechanism for differential release of acrosomal enzymes during the acrosome reaction. *Biochem J* 275:759–766
177. Salicioni AM, Platt MD, Wertheimer EV, Arcelay E, Allaire A, Sosnik J, Visconti PE 2007 Signalling pathways involved in sperm capacitation. *Soc Reprod Fertil Suppl* 65:245–259
178. De Blas G, Michaut M, Trevino CL, Tomes CN, Yunes R, Darszon A, Mayorga LS 2002 The intraacrosomal calcium pool plays a direct role in acrosomal exocytosis. *J Biol Chem* 277:49326–49331
179. Breitbart H 2002 Intracellular calcium regulation in sperm capacitation and acrosomal reaction. *Mol Cell Endocrinol* 187:139–144
180. Carlson AE, Westenbroek RE, Quill T, Ren D, Clapham DE, Hille B, Garbers DL, Babcock DF 2003 CatSper1 required for evoked Ca²⁺ entry and control of flagellar function in sperm. *Proc Natl Acad Sci USA* 100:14864–14868
181. Marquez B, Ignatz G, Suarez SS 2007 Contributions of extracellular and intracellular Ca²⁺ to regulation of sperm motility: release of intracellular stores can hyperactivate CatSper1 and CatSper2 null sperm. *Dev Biol* 303:214–221
182. Darszon A, Acevedo JJ, Galindo BE, Hernandez-Gonzalez EO, Nishigaki T, Trevino CL, Wood C, Beltran C 2006 Sperm channel diversity and functional multiplicity. *Reproduction* 131:977–988
183. Harper CV, Barratt CL, Publicover SJ 2004 Stimulation of human spermatozoa with progesterone gradients to simulate approach to the oocyte. Induction of [Ca²⁺] (i) oscillations and cyclical transitions in flagellar beating. *J Biol Chem* 279:46315–46325
184. Darszon A, Nishigaki T, Wood C, Trevino CL, Felix R, Beltran C 2005 Calcium channels and Ca²⁺ fluctuations in sperm physiology. *Int Rev Cytol* 243:79–172
185. Teramoto H, Coso OA, Miyata H, Igishi T, Miki T, Gutkind JS 1996 Signaling from the small GTP-binding proteins Rac1 and Cdc42 to the c-Jun N-terminal kinase/stress-activated protein kinase pathway. A role for mixed lineage kinase 3/protein-tyrosine kinase 1, a novel member of the mixed lineage kinase family. *J Biol Chem* 271:27225–27228
186. Gallo KA, Johnson GL 2002 Mixed-lineage kinase control of JNK and p38 MAPK pathways. *Nat Rev Mol Cell Biol* 3:663–672
187. Cha H, Smith BL, Gallo K, Machamer CE, Shapiro P 2004 Phosphorylation of golgin-160 by mixed lineage kinase 3. *J Cell Sci* 117:751–760
188. Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, Wagner L 2003 Database resources of the National Center for Biotechnology. *Nucleic Acids Res* 31:28–33
189. Urayama S, Harada Y, Nakagawa Y, Ban S, Akasaka M, Kawasaki N, Sawada H 2008 Ascidian sperm GPI-anchored crisp-like protein as a binding partner for an allorecognizable sperm receptor on the vitelline coat. *J Biol Chem* 283:21725–21733
190. Yudin AI, Li MW, Robertson KR, Tollner T, Cherr GN, Overstreet JW 2002 Identification of a novel GPI-anchored CRISP glycoprotein, MAK248, located on the posterior head and equatorial segment of cynomolgus macaque sperm. *Mol Reprod Dev* 63:488–499
191. Cosentino MJ, Cockett AT 1986 Structure and function of the epididymis. *Urol Res* 14:229–240
192. Brooks DE 1983 Epididymal functions and their hormonal regulation. *Aust J Biol Sci* 36:205–221
193. Dacheux JL, Paquignon M 1980 Relations between the fertilizing ability, motility and metabolism of epididymal spermatozoa. *Reprod Nutr Dev* 20:1085–1099
194. Asquith KL, Baleato RM, McLaughlin EA, Nixon B, Aitken RJ 2004 Tyrosine phosphorylation activates surface chaperones facilitating sperm-zona recognition. *J Cell Sci* 117:3645–3657
195. Aitken RJ, Nixon B, Lin M, Koppers AJ, Lee YH, Baker MA 2007

- Proteomic changes in mammalian spermatozoa during epididymal maturation. *Asian J Androl* 9:554–564
196. **Bremner WJ, de Kretser DM** 1976 The prospects for new, reversible male contraceptives. *N Engl J Med* 295:1111–1117
 197. **Hinton BT** 1980 The epididymal microenvironment: a site of attack for a male contraceptive? *Invest Urol* 18:1–10
 198. **Austin CR** 1967 Capacitation of spermatozoa. *Int J Fertil* 12:25–31
 199. **Belton Jr RJ, Adams NL, Foltz KR** 2001 Isolation and characterization of sea urchin egg lipid rafts and their possible function during fertilization. *Mol Reprod Dev* 59:294–305
 200. **Cross NL** 2004 Reorganization of lipid rafts during capacitation of human sperm. *Biol Reprod* 71:1367–1373
 201. **Shadan S, James PS, Howes EA, Jones R** 2004 Cholesterol efflux alters lipid raft stability and distribution during capacitation of boar spermatozoa. *Biol Reprod* 71:253–265
 202. **Tanphaichitr N, Carmona E, Bou Khalil M, Xu H, Berger T, Gerton GL** 2007 New insights into sperm-zona pellucida interaction: involvement of sperm lipid rafts. *Front Biosci* 12:1748–1766
 203. **Visconti PE, Moore GD, Bailey JL, Leclerc P, Connors SA, Pan D, Olds-Clarke P, Kopf GS** 1995 Capacitation of mouse spermatozoa. II. Protein tyrosine phosphorylation and capacitation are regulated by a cAMP-dependent pathway. *Development* 121:1139–1150
 204. **Demarco IA, Espinosa F, Edwards J, Sosnik J, De La Vega-Beltran JL, Hockensmith JW, Kopf GS, Darszon A, Visconti PE** 2003 Involvement of a Na⁺/HCO₃⁻ cotransporter in mouse sperm capacitation. *J Biol Chem* 278:7001–7009
 205. **Gadella BM, Van Gestel RA** 2004 Bicarbonate and its role in mammalian sperm function. *Anim Reprod Sci* 82–83:307–319
 206. **Visconti PE, Westbrook VA, Chertihin O, Demarco I, Sleight S, Diekman AB** 2002 Novel signaling pathways involved in sperm acquisition of fertilizing capacity. *J Reprod Immunol* 53:133–150
 207. **Visconti PE, Bailey JL, Moore GD, Pan D, Olds-Clarke P, Kopf GS** 1995 Capacitation of mouse spermatozoa. I. Correlation between the capacitation state and protein tyrosine phosphorylation. *Development* 121:1129–1137
 208. **Suarez SS, Osman RA** 1987 Initiation of hyperactivated flagellar bending in mouse sperm within the female reproductive tract. *Biol Reprod* 36:1191–1198
 209. **Burkman LJ** 1984 Characterization of hyperactivated motility by human spermatozoa during capacitation: comparison of fertile and oligozoospermic sperm populations. *Arch Androl* 13:153–165
 210. **Florman HM, First NL** 1988 Regulation of acrosomal exocytosis. II. The zona pellucida-induced acrosome reaction of bovine spermatozoa is controlled by extrinsic positive regulatory elements. *Dev Biol* 128:464–473
 211. **Florman HM, First NL** 1988 The regulation of acrosomal exocytosis. I. Sperm capacitation is required for the induction of acrosome reactions by the bovine zona pellucida in vitro. *Dev Biol* 128:453–463
 212. **Oliphant G, Reynolds AB, Thomas TS** 1985 Sperm surface components involved in the control of the acrosome reaction. *Am J Nat* 174:269–283
 213. **Fraser LR, Harrison RA, Herod JE** 1990 Characterization of a decapacitation factor associated with epididymal mouse spermatozoa. *J Reprod Fertil* 89:135–148
 214. **Fraser LR** 1998 Interactions between a decapacitation factor and mouse spermatozoa appear to involve fucose residues and a GPI-anchored receptor. *Mol Reprod Dev* 51:193–202
 215. **Nixon B, MacIntyre DA, Mitchell LA, Gibbs GM, O'Bryan M, Aitken RJ** 2006 The identification of mouse sperm-surface-associated proteins and characterization of their ability to act as decapacitation factors. *Biol Reprod* 74:275–287
 216. **Moffitt JS, Boekelheide K, Sedivy JM, Klysik J** 2007 Mice lacking Raf kinase inhibitor protein-1 (RKIP-1) have altered sperm capacitation and reduced reproduction rates with a normal response to testicular injury. *J Androl* 28:883–890
 217. **Gibbons R, Adeoya-Osiguwa SA, Fraser LR** 2005 A mouse sperm decapacitation factor receptor is phosphatidylethanolamine-binding protein 1. *Reproduction* 130:497–508
 218. **Martins SG, Miranda PV, Brandelli A** 2003 Acrosome reaction inhibitor released during in vitro sperm capacitation. *Int J Androl* 26:296–304
 219. **Huang YH, Chen YH, Lin CM, Ciou YY, Kuo SP, Chen CT, Shih CM, Chang EE** 2007 Suppression effect of seminal vesicle autoantigen on platelet-activating factor-induced mouse sperm capacitation. *J Cell Biochem* 100:941–951
 220. **Suarez SS, Varosi SM, Dai X** 1993 Intracellular calcium increases with hyperactivation in intact, moving hamster sperm and oscillates with the flagellar beat cycle. *Proc Natl Acad Sci USA* 90:4660–4664
 221. **Ho HC, Suarez SS** 2001 Hyperactivation of mammalian spermatozoa: function and regulation. *Reproduction* 122:519–526
 222. **Darszon A, Lopez-Martinez P, Acevedo JJ, Hernandez-Cruz A, Trevino CL** 2006 T-type Ca²⁺ channels in sperm function. *Cell Calcium* 40:241–252
 223. **Jimenez-Gonzalez C, Michelangeli F, Harper CV, Barratt CL, Publicover SJ** 2006 Calcium signalling in human spermatozoa: a specialized 'toolkit' of channels, transporters and stores. *Hum Reprod Update* 12:253–267
 224. **Quill TA, Ren D, Clapham DE, Garbers DL** 2001 A voltage-gated ion channel expressed specifically in spermatozoa. *Proc Natl Acad Sci USA* 98:12527–12531
 225. **Ren D, Navarro B, Perez G, Jackson AC, Hsu S, Shi Q, Tilly JL, Clapham DE** 2001 A sperm ion channel required for sperm motility and male fertility. *Nature* 413:603–609
 226. **Jin J, Jin N, Zheng H, Ro S, Tafolla D, Sanders KM, Yan W** 2007 Catsper3 and Catsper4 are essential for sperm hyperactivated motility and male fertility in the mouse. *Biol Reprod* 77:37–44
 227. **Qi H, Moran MM, Navarro B, Chong JA, Krapivinsky G, Krapivinsky L, Kirichok Y, Ramsey IS, Quill TA, Clapham DE** 2007 All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. *Proc Natl Acad Sci USA* 104:1219–1223
 228. **Ho HC, Suarez SS** 2003 Characterization of the intracellular calcium store at the base of the sperm flagellum that regulates hyperactivated motility. *Biol Reprod* 68:1590–1596
 229. **Ho HC, Suarez SS** 2001 An inositol 1,4,5-trisphosphate receptor-gated intracellular Ca(2+) store is involved in regulating sperm hyperactivated motility. *Biol Reprod* 65:1606–1615
 230. **Aoki N, Sakiyama A, Kuroki K, Maenaka K, Kohda D, Deshimaru M, Terada S** 2008 Serotiflin, a CRISP family protein with binding affinity for small serum protein-2 in snake serum. *Biochim Biophys Acta* 1784:621–628
 231. **Aoki N, Sakiyama A, Deshimaru M, Terada S** 2007 Identification of novel serum proteins in a Japanese viper: homologs of mammalian PSP94. *Biochem Biophys Res Commun* 359:330–334
 232. **Mbikay M, Nolet S, Fournier S, Benjannet S, Chapdelaine P, Paradis G, Dube JY, Tremblay R, Lazure C, Seidah NG, Chretien M** 1987 Molecular cloning and sequence of the cDNA for a 94-amino-acid seminal plasma protein secreted by the human prostate. *DNA* 6:23–29
 233. **Jeng H, Liu KM, Chang WC** 1993 Purification and characterization of reversible sperm motility inhibitors from porcine seminal plasma. *Biochem Biophys Res Commun* 191:435–440
 234. **Chao CF, Chiou ST, Jeng H, Chang WC** 1996 The porcine sperm motility inhibitor is identical to β -microseminoprotein and is a competitive inhibitor of Na⁺,K⁺-ATPase. *Biochem Biophys Res Commun* 218:623–628
 235. **Tollner TL, Yudin AI, Treece CA, Overstreet JW, Cherr GN** 2004 Macaque sperm release ESP13.2 and PSP94 during capacitation: the absence of ESP13.2 is linked to sperm-zona recognition and binding. *Mol Reprod Dev* 69:325–337
 236. **Anahí Franchi N, Avendano C, Molina R, Tissera A, Maldonado C, Oehninger S, Coronel C** 2008 β -Microseminoprotein in human spermatozoa and its potential role in male fertility. *Reproduction* 136:157–166
 237. **Cuasnicu PS, Gonzalez Echeverria F, Piazza AD, Cameo MS, Blaquier JA** 1984 Antibodies against epididymal glycoproteins block fertilizing ability in rat. *J Reprod Fertil* 72:467–471
 238. **Hall JC, Tubbs CE** 1997 Purification and characterization of protein D/E, a putative sperm-binding protein involved in fertilization. *Prep Biochem Biotechnol* 27:239–251
 239. **Fournier-Delpech S, Courtois M, Dubois MP** 1985 Decreased fer-

- tility and motility of spermatozoa from rats immunized with a prealbumin epididymal-specific glycoprotein. *J Androl* 6: 246–250
240. Cohen DJ, Munuce MJ, Cuasnicu PS 1996 Mammalian sperm-egg fusion: the development of rat oolemma fusibility during oogenesis involves the appearance of binding sites for sperm protein “DE”. *Biol Reprod* 55:200–206
 241. Cohen DJ, Ellerman DA, Cuasnicu PS 2000 Mammalian sperm-egg fusion: evidence that epididymal protein DE plays a role in mouse gamete fusion. *Biol Reprod* 63:462–468
 242. Cohen DJ, Ellerman DA, Busso D, Morgenfeld MM, Piazza AD, Hayashi M, Young ET, Kasahara M, Cuasnicu PS 2001 Evidence that human epididymal protein ARP plays a role in gamete fusion through complementary sites on the surface of the human egg. *Biol Reprod* 65:1000–1005
 243. Ellerman DA, Brantua VS, Martinez SP, Cohen DJ, Conesa D, Cuasnicu PS 1998 Potential contraceptive use of epididymal proteins: immunization of male rats with epididymal protein DE inhibits sperm fusion ability. *Biol Reprod* 59:1029–1036
 244. Perez Martinez S, Conesa D, Cuasnicu PS 1995 Potential contraceptive use of epididymal proteins: evidence for the participation of specific antibodies against rat epididymal protein DE in male and female fertility inhibition. *J Reprod Immunol* 29:31–45
 245. Inoue N, Ikawa M, Isotani A, Okabe M 2005 The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. *Nature* 434:234–238
 246. Kaupp UB, Kashikar ND, Weyand I 2008 Mechanisms of sperm chemotaxis. *Annu Rev Physiol* 70:93–117
 247. Bonnell BS, Reinhart D, Chandler DE 1996 *Xenopus laevis* egg jelly coats consist of small diffusible proteins bound to a complex system of structurally stable networks composed of high-molecular-weight glycoconjugates. *Dev Biol* 174:32–42
 248. al-Anzi B, Chandler DE 1998 A sperm chemoattractant is released from *Xenopus* egg jelly during spawning. *Dev Biol* 198:366–375
 249. Xiang X, Burnett L, Rawls A, Bieber A, Chandler D 2004 The sperm chemoattractant “allurin” is expressed and secreted from the *Xenopus* oviduct in a hormone-regulated manner. *Dev Biol* 275: 343–355
 250. Xiang X, Kittelson A, Olson J, Bieber A, Chandler D 2005 Allurin, a 21 kD sperm chemoattractant, is rapidly released from the outermost jelly layer of the *Xenopus* egg by diffusion and medium convection. *Mol Reprod Dev* 70:344–360
 251. Burnett LA, Boyles S, Spencer C, Bieber AL, Chandler DE 2008 *Xenopus tropicalis* allurin: expression, purification, and characterization of a sperm chemoattractant that exhibits cross-species activity. *Dev Biol* 316:408–416
 252. Olson JH, Xiang X, Ziegert T, Kittelson A, Rawls A, Bieber AL, Chandler DE 2001 Allurin, a 21-kDa sperm chemoattractant from *Xenopus* egg jelly, is related to mammalian sperm-binding proteins. *Proc Natl Acad Sci USA* 98:11205–11210
 253. Hamann H, Jude R, Sieme H, Mertens U, Topfer-Petersen E, Distl O, Leeb T 2007 A polymorphism within the equine CRISP3 gene is associated with stallion fertility in Hanoverian warmblood horses. *Anim Genet* 38:259–264
 254. Paoloni-Giacobino A, Kern I, Rumpler Y, Djelati R, Morris MA, Dahoun SP 2000 Familial t(6;21)(p21.1;p13) translocation associated with male-only sterility. *Clin Genet* 58:324–328
 255. Olesen C, Hansen C, Bendsen E, Byskov AG, Schwinger E, Lopez-Pajares I, Jensen PK, Kristoffersson U, Schubert R, Van Assche E, Wahlstroem J, Lespinasse J, Tommerup N 2001 Identification of human candidate genes for male infertility by digital differential display. *Mol Hum Reprod* 7:11–20
 256. Ernst T, Hergenroth M, Kenzelmann M, Cohen CD, Bonrouhi M, Weninger A, Klaren R, Grone EF, Wiesel M, Gudemann C, Kuster J, Schott W, Staehler G, Kretzler M, Hollstein M, Grone HJ 2002 Decrease and gain of gene expression are equally discriminatory markers for prostate carcinoma: a gene expression analysis on total and microdissected prostate tissue. *Am J Pathol* 160:2169–2180
 257. Chan PS, Chan LW, Xuan JW, Chin JL, Choi HL, Chan FL 1999 In situ hybridization study of PSP94 (prostatic secretory protein of 94 amino acids) expression in human prostates. *Prostate* 41:99–109
 258. Bjartell AS, Al-Ahmadie H, Serio AM, Eastham JA, Eggen SE, Fine SW, Udby L, Gerald WL, Vickers AJ, Lilja H, Reuter VE, Scardino PT 2007 Association of cysteine-rich secretory protein 3 and β -microseminoprotein with outcome after radical prostatectomy. *Clin Cancer Res* 13:4130–4138
 259. Satoh T, Timme TL, Saika T, Ebara S, Yang G, Wang J, Ren C, Kusaka N, Mouraviev V, Thompson TC 2003 Adenoviral vector-mediated mRTPV-1 gene therapy for prostate cancer. *Hum Gene Ther* 14:91–101
 260. Naruishi K, Timme TL, Kusaka N, Fujita T, Yang G, Goltsov A, Satoh T, Ji X, Tian W, Abdelfattah E, Men T, Watanabe M, Tabata K, Thompson TC 2006 Adenoviral vector-mediated RTPV-1 gene-modified tumor cell-based vaccine suppresses the development of experimental prostate cancer. *Cancer Gene Ther* 13: 658–663
 261. Li L, Abdelfattah E, Cao G, Ren C, Yang G, Goltsov AA, Chinault AC, Cai WW, Timme TL, Thompson TC 2008 Glioma pathogenesis-related protein 1 exerts tumor suppressor activities through proapoptotic reactive oxygen species-c-Jun-NH2 kinase signaling. *Cancer Res* 68:434–443
 262. Call KM, Glaser T, Ito CY, Buckler AJ, Pelletier J, Haber DA, Rose EA, Kral A, Yeager H, Lewis WH, Jones C, Housman DE 1990 Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms’ tumor locus. *Cell* 60:509–520
 263. Rivera MN, Haber DA 2005 Wilms’ tumour: connecting tumorigenesis and organ development in the kidney. *Nat Rev Cancer* 5:699–712
 264. Rivera MN, Kim WJ, Wells J, Driscoll DR, Brannigan BW, Han M, Kim JC, Feinberg AP, Gerald WL, Vargas SO, Chin L, Iafate AJ, Bell DW, Haber DA 2007 An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. *Science* 315:642–645
 265. Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP 1993 Relaxation of imprinted genes in human cancer. *Nature* 362:747–749
 266. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, Smith PJ, Reeve AE 1993 Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms’ tumour. *Nature* 362: 749–751
 267. Chilukamarri L, Hancock AL, Malik S, Zabkiewicz J, Baker JA, Greenhough A, Dallosso AR, Huang TH, Royer-Pokora B, Brown KW, Malik K 2007 Hypomethylation and aberrant expression of the glioma pathogenesis-related 1 gene in Wilms tumors. *Neoplasia* 9:970–978
 268. Hawdon JM, Hotez PJ 1996 Hookworm: developmental biology of the infectious process. *Curr Opin Genet Dev* 6:618–623
 269. Moyle M, Foster DL, McGrath DE, Brown SM, Laroche Y, De Meutter J, Stanssens P, Bogowitz CA, Fried VA, Ely JA, Soule HR, Vlasuk GP 1994 A hookworm glycoprotein that inhibits neutrophil function is a ligand of the integrin CD11b/CD18. *J Biol Chem* 269:10008–10015
 270. Del Valle A, Jones BF, Harrison LM, Chadderdon RC, Cappello M 2003 Isolation and molecular cloning of a secreted hookworm platelet inhibitor from adult *Ancylostoma caninum*. *Mol Biochem Parasitol* 129:167–177
 271. Hawdon JM, Jones BF, Hoffman DR, Hotez PJ 1996 Cloning and characterization of *Ancylostoma* secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae. *J Biol Chem* 271:6672–6678
 272. Hawdon JM, Narasimhan S, Hotez PJ 1999 *Ancylostoma* secreted protein 2: cloning and characterization of a second member of a family of nematode secreted proteins from *Ancylostoma caninum*. *Mol Biochem Parasitol* 99:149–165
 273. Muchowski PJ, Zhang L, Chang ER, Soule HR, Plow EF, Moyle M 1994 Functional interaction between the integrin antagonist neutrophil inhibitory factor and the I domain of CD11b/CD18. *J Biol Chem* 269:26419–26423
 274. Takada Y, Ye X, Simon S 2007 The integrins. *Genome Biol* 8:215
 275. Solovjov DA, Pluskota E, Plow EF 2005 Distinct roles for the α and β subunits in the functions of integrin α M β 2. *J Biol Chem* 280: 1336–1345
 276. Bin Z, Hawdon J, Qiang S, Hainan R, Huiqing Q, Wei H, Shu-Hua X, Tiehua L, Xing G, Zheng F, Hotez P 1999 *Ancylostoma* secreted protein 1 (ASP-1) homologues in human hookworms. *Mol Biochem Parasitol* 98:143–149

277. Zhan B, Liu Y, Badamchian M, Williamson A, Feng J, Loukas A, Hawdon JM, Hotez PJ 2003 Molecular characterisation of the Ancylostoma-secreted protein family from the adult stage of *Ancylostoma caninum*. *Int J Parasitol* 33:897–907
278. Yatsuda AP, Eysker M, Vieira-Bressan MC, De Vries E 2002 A family of activation associated secreted protein (ASP) homologues of *Cooperia punctata*. *Res Vet Sci* 73:297–306
279. Ding X, Shields J, Allen R, Hussey RS 2000 Molecular cloning and characterisation of a venom allergen AG5-like cDNA from *Meloidogyne incognita*. *Int J Parasitol* 30:77–81
280. Gao B, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS 2001 Molecular characterisation and expression of two venom allergen-like protein genes in *Heterodera glycines*. *Int J Parasitol* 31:1617–1625
281. Tawe W, Pearlman E, Unnasch TR, Lustigman S 2000 Angiogenic activity of *Onchocerca volvulus* recombinant proteins similar to vespoid venom antigen 5. *Mol Biochem Parasitol* 109:91–99
282. Daub J, Loukas A, Pritchard DI, Blaxter M 2000 A survey of genes expressed in adults of the human hookworm, *Necator americanus*. *Parasitology* 120:171–184
283. Kovalick GE, Schreiber MC, Dickason AK, Cunningham RA 1998 Structure and expression of the antigen 5-related gene of *Drosophila melanogaster*. *Insect Biochem Mol Biol* 28:491–500
284. Mans BJ, Andersen JF, Francischetti IM, Valenzuela JG, Schwan TG, Pham VM, Garfield MK, Hammer CK, Ribeiro JM 2008 Comparative sialomics between hard and soft ticks: implications for the evolution of blood-feeding behavior. *Insect Biochem Mol Biol* 38:42–58
285. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM 1999 Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc Natl Acad Sci USA* 96:15155–15160
286. Ameri M, Wang X, Wilkerson MJ, Kanost MR, Broce AB 2008 An immunoglobulin binding protein (antigen 5) of the stable fly (*Diptera: Muscidae*) salivary gland stimulates bovine immune responses. *J Med Entomol* 45:94–101
287. Calvo E, Dao A, Pham VM, Ribeiro JM 2007 An insight into the sialome of *Anopheles funestus* reveals an emerging pattern in anopheline salivary protein families. *Insect Biochem Mol Biol* 37:164–175
288. Valenzuela JG, Pham VM, Garfield MK, Francischetti IM, Ribeiro JM 2002 Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. *Insect Biochem Mol Biol* 32:1101–1122
289. King TP, Sobotka AK, Alagon A, Kochoumian L, Lichtenstein LM 1978 Protein allergens of white-faced hornet, yellow hornet, and yellow jacket venoms. *Biochemistry* 17:5165–5174
290. Hoffman DR 1985 Allergens in *Hymenoptera* venom XIV: IgE binding activities of venom proteins from three species of vespids. *J Allergy Clin Immunol* 75:606–610
291. King TP, Alagon AC, Kuan J, Sobotka AK, Lichtenstein LM 1983 Immunochemical studies of yellow jacket venom proteins. *Mol Immunol* 20:297–308
292. Hoffman DR 1985 Allergens in *Hymenoptera* venom XV: The immunologic basis of vespid venom cross-reactivity. *J Allergy Clin Immunol* 75:611–613
293. Lu G, Villalba M, Coscia MR, Hoffman DR, King TP 1993 Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets. *J Immunol* 150:2823–2830
294. King TP, Lu G 1997 Hornet venom allergen antigen 5, Dol m 5: its T-cell epitopes in mice and its antigenic cross-reactivity with a mammalian testis protein. *J Allergy Clin Immunol* 99:630–639
295. Bohle B, Zwolfer B, Fischer GF, Seppala U, Kinaciyan T, Bolwig C, Spangfort MD, Ebner C 2005 Characterization of the human T cell response to antigen 5 from *Vespula vulgaris* (Ves v 5). *Clin Exp Allergy* 35:367–373
296. Ribeiro JM, Francischetti IM 2003 Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annu Rev Entomol* 48:73–88
297. Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M, Fritig B 1993 Plant ‘pathogenesis-related’ proteins and their role in defense against pathogens. *Biochimie* 75:687–706
298. van Loon LC, van Strien EA 1999 The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol Mol Plant Pathol* 55:85–97
299. Kitajima S, Sato F 1999 Plant pathogenesis-related proteins: molecular mechanisms of gene expression and protein function. *J Biochem (Tokyo)* 125:1–8
300. Antoniow JF, Pierpoint WS 1978 Purification of a tobacco leaf protein associated with resistance to virus infection. *Biochem Soc Trans* 6:248–250
301. Antoniow JF, Ritter CE, Pierpoint WS, Van Loon LC 1980 Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV. *J Gen Virol* 47:79–87
302. Cutt JR, Harpster MH, Dixon DC, Carr JP, Dunsmuir P, Klessig DF 1989 Disease response to tobacco mosaic virus in transgenic tobacco plants that constitutively express the pathogenesis-related PR1b gene. *Virology* 173:89–97
303. Alexander D, Goodman RM, Gut-Rella M, Glascock C, Weymann K, Friedrich L, Maddox D, Ahl-Goy P, Luntz T, Ward E, Ryals J 1993 Increased tolerance to two oomycete pathogens in transgenic tobacco expressing pathogenesis-related protein 1a. *Proc Natl Acad Sci USA* 90:7327–7331
304. Niderman T, Genetet I, Bruyere T, Gees R, Stintzi A, Legrand M, Fritig B, Mosinger E 1995 Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans*. *Plant Physiol* 108:17–27
305. Riviere MP, Marais A, Ponchet M, Willats W, Galiana E 2008 Silencing of acidic pathogenesis-related PR-1 genes increases extracellular β -(1 $\rightarrow$ 3)-glucanase activity at the onset of tobacco defence reactions. *J Exp Bot* 59:1225–1239
306. Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD 1996 Systemic acquired resistance. *Plant Cell* 8:1809–1819
307. Fraser RSS 1981 Evidence for the occurrence of the ‘‘pathogenesis-related’’ proteins in the leaves of healthy tobacco plants during flowering. *Physiol Plant Pathol* 19:69–76
308. Dewald B, Bretz U, Baggiolini M 1982 Release of gelatinase from a novel secretory compartment of human neutrophils. *J Clin Invest* 70:518–525
309. Kjeldsen L, Sengelov H, Lollike K, Nielsen MH, Borregaard N 1994 Isolation and characterization of gelatinase granules from human neutrophils. *Blood* 83:1640–1649
310. Weiss SJ, Peppin GJ 1986 Collagenolytic metalloenzymes of the human neutrophil. Characteristics, regulation and potential function in vivo. *Biochem Pharmacol* 35:3189–3197
311. Oyewumi L, Kaplan F, Sweezey NB 2003 Lgl1, a mesenchymal modulator of early lung branching morphogenesis, is a secreted glycoprotein imported by late gestation lung epithelial cells. *Biochem J* 376:61–69
312. Gingras MC, Margolin JF 2000 Differential expression of multiple unexpected genes during U937 cell and macrophage differentiation detected by suppressive subtractive hybridization. *Exp Hematol* 28:350
313. Friess H, Ding J, Kleeff J, Liao Q, Berberat PO, Hammer J, Buchler MW 2001 Identification of disease-specific genes in chronic pancreatitis using DNA array technology. *Ann Surg* 234:769–778; discussion 778–769
314. Sarles H, Bernard JP, Johnson C 1989 Pathogenesis and epidemiology of chronic pancreatitis. *Annu Rev Med* 40:453–468
315. Seltzer Z, Dubner R, Shir Y 1990 A novel behavioral model of neuropathic pain disorders produced in rats by partial sciatic nerve injury. *Pain* 43:205–218
316. Millan MJ 1999 The induction of pain: an integrative review. *Prog Neurobiol* 57:1–164
317. Vaillancourt PD, Langevin HM 1999 Painful peripheral neuropathies. *Med Clin North Am* 83:627–642, vi
318. Newton RA, Bingham S, Davey PD, Medhurst AD, Piercy V, Raval P, Parsons AA, Sanger GJ, Case CP, Lawson SN 2000 Identification of differentially expressed genes in dorsal root ganglia following partial sciatic nerve injury. *Neuroscience* 95:1111–1120
319. Keast JR, Gleeson RJ 1998 Androgen receptor immunoreactivity

- is present in primary sensory neurons of male rats. *Neuroreport* 9:4137–4140
320. **Spinale FG** 2002 Matrix metalloproteinases: regulation and dysregulation in the failing heart. *Circ Res* 90:520–530
321. **Chess PR, D'Angio CT, Pryhuber GS, Maniscalco WM** 2006 Pathogenesis of bronchopulmonary dysplasia. *Semin Perinatol* 30:171–178
322. **Fry BG** 2005 From genome to “venome”: molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. *Genome Res* 15:403–420
323. **Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SF, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes R, Kochva E** 2006 Early evolution of the venom system in lizards and snakes. *Nature* 439:584–588
324. **Guex N, Peitsch MC** 1997 SWISS-MODEL and the Swiss-Pdb-Viewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723
325. **Ronquist F, Huelsenbeck JP** 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574

Endocrine Reviews is published by The Endocrine Society (<http://www.endo-society.org>), the foremost professional society serving the endocrine community.