Human Epididymis Protein-4 (HE-4): A Novel Cross-Class Protease Inhibitor

Nirmal Chhikara1,*, Mayank Saraswat1,*, Anil Kumar Tomar1, Sharmistha Dey1, Sarman Singh2, Savita Yadav1*

1 Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India, 2Department of Lab Medicine, All India Institute of Medical Sciences, New Delhi, India

Abstract

Epididymal proteins represent the factors necessary for maturation of sperm and play a crucial role in sperm maturation. HE-4, an epididymal protein, is a member of whey acidic protein four-disulfide core (WFDC) family with no known function. A WFDC protein has a conserved WFDC domain of 50 amino acids with eight conserved cystine residue. HE-4 is a 124 amino acid long polypeptide with two WFDC domains. Here, we show that HE-4 is secreted in the human seminal fluid as a disulfide-bonded homo-trimer and is a cross-class protease inhibitor inhibits some of the serine, aspartyl and cysteine proteases tested using hemoglobin as a substrate. Using SPR we have also observed that HE-4 shows a significant binding with all these proteases. Disulfide linkages are essential for this activity. Moreover, HE-4 is N-glycosylated and highly stable on a wide range of pH and temperature. Taken together this suggests that HE-4 is a cross-class protease inhibitor which might confer protection against microbial virulence factors of proteolytic nature.

Introduction

Human and chimpanzee genomes are similar to the extent of nearly 90% but, there are differences which impart uniqueness to the human species as well as chimpanzees and reveal the intrinsic genetic differences in the expressions of genes. A sudden or adaptive evolution in some of the genes or genomic regions might be playing some essential part in these differences. Recent comparative analysis of human and chimpanzee genome sequences identified some 16 regions with high density of rapidly evolving genes [1]. One such region contains genes encoding whey acidic protein (WAP) domain proteins. This region on human chromosome 20q13 is called WAP four-disulfide core domain (WFDC) locus containing 14 genes encoding WFDC type proteinase inhibitors [2]. Apart from these 14 genes from the same locus there are at least four other proteins having WFDC domain but are present at different chromosomes (Ch. 16, 17 and X chromosome) [3–6]. Amplification of the 20q12-13 region has been found [7–8]. Consistent with these studies, SLPI and elafin are known to be expressed in various carcinomas and implicated in initiation or progression of tumorigenesis [9–13].

Most of the members of the family with the exception of SLPI, elafin, KAL1, EPPIN and ps20 have not been examined at the protein level. SLPI, elafin and ps20 have been reported to be expressed in different cell types including airway epithelium and mucosal secretions from tissues including male reproductive tract, respiratory tract as well as in inflammatory cells like T-cells and macrophages [14–16]. Two types of functions attributed to this family of proteins are regulation of proinflammatory mediators and anti-bacterial or anti-fungal activity [17–18]. Anti-infective activity of SLPI, elafin, and pro-infection attributes of ps20 regarding HIV have also been uncovered [19].

Another member of the family, WFDC-2 or HE-4 (human epididymis protein-4) was found to be the most frequently upregulated in ovarian carcinomas [20]. This protein is also known as Epididymal secretory protein E4, Major epididymus-specific protein E4 and putative protease inhibitor WAP5. WFDC-2 gene product was originally thought to be a protein specifically expressed in the epididymis and was dubbed as a tissue marker for the same [21]. Later, it was found to be expressed in the oral cavity, respiratory tract, female genital tract and distal renal tubules. Evidence of its expression in the colonic mucosa has also been found [22]. Evidence of HE-4 expression in various tumor types of the lung including lung adenocarcinoma has also been reported [16]. HE-4 levels in the serum have been suggested to be a sensitive marker for ovarian cancer and protein used as a histological marker for ovarian cancer [23–25].

On the basis of structural and sequential similarity of HE-4 with other WAP proteins like SLPI and elafin it was suggested that the protein might have antiprotease activity within the male reproductive, oral and respiratory tract. Other than this suggestion of...
HE-4 working as an antiprotease, no work has been done to elucidate the role and/or structure of the protein. Even this claim projecting it as a protease inhibitor has never been tested. It seems plausible that HE-4 will have as yet undetermined pivotal roles in human physiology. An investigation of this demand detailed studies at the protein level. A prerequisite for the above said objective is a rapid and efficient method to generate sufficient amounts of protein. This paper reports an accessible and efficient purification of HE-4 protein from human seminal fluid. We have further characterized HE-4, and found it to be a cross-class protease inhibitor and show a significant binding with all the proteases tested.

**Results**

**HE-4 purified from seminal fluid exists as a trimer**

HE-4 was purified from normozoospermic pooled human seminal plasma. Heparin-sepharose unbound fraction was further fractionated on DEAcephacel. Seminal plasma fractionated into four main peaks on DEAE-sephacel (Fig. 1 A, B). All the peaks were further fractionated by sephadex G-75 size exclusion chromatography (data shown only for peak I of DEAE) and HE-4 immunoactivity was found in peak III of G-75 fractionation (Fig. 1E, lane II). Western blot of purified protein and whole seminal plasma showed a band of approximately 14 kDa (Fig. 1E) which was confirmed by SDS-PAGE of the same in reducing conditions (Fig. 1D lane III). Under non-reducing conditions, HE-4 was found to migrate as a single band of approximately 42 kDa (Fig. 1D lane II). This suggests that HE-4 exists as a disulphide bonded trimer in human seminal plasma and disulphide bond reduction provide a monomer of 14 kDa. Identity of HE-4 was also confirmed by LC-MS/MS (Table 1). Mascot score of identification for HE-4 was 123 and four peptides were matched. One previous study has reported HE-4 to be N-glycosylated in saliva [26]. We analyzed the glycosylation of HE-4 with PAS staining method (Fig. 2A). We treated HE-4 with PNGase F under denaturing conditions and compared it with the native protein. An approximate shift of 2.5 kDa in the migration of deglycosylated HE-4 in SDS-PAGE was observed compared to native protein (Fig. 2B). This confirmed that HE-4 in seminal fluid is N-glycosylated.

Effect of treatment with various chemicals on solution behavior and possible aggregation pattern of native HE-4 was analyzed with dynamic light scattering (DLS). Purified HE-4 was incubated with solutions of varying pH (from acidic to alkaline) and different chemicals including various salts and reducing agents at different concentrations. The hydrodynamic radius ($R_h$) of purified HE-4 (Tris-HCl buffer pH 8.0 in which protein was purified) was found to be 3.18±0.07 nm. This corresponds to approximately 46 kDa, and in complete agreement with migration of HE-4 in SDS-PAGE. Different pH had varied effect on HE-4 $R_h$ values (Fig. 3A). At low pH, the $R_h$ was higher which returned to native values as pH increased. Metal ions can contribute to aggregation of HE-4 purified from seminal fluid exists as a trimer.

**Sequence analysis of HE-4**

Multiple sequence alignments (MSA) of HE-4 across multiple species were performed using ClustalW2 on EBI web server (Fig. 3). WFDC-2 protein sequences from different species show high similarity (Fig. 3C). Each domain consists of eight Cys residues which are fully conserved. Besides, there are other many conserved residues a close examination of alignment reveals that these conserved residues also follow a common pattern of spacing. Comparison of the spacing between common conserved residues in each domain finds the following pattern: K1/G1/C2/7/11/C3/3/C2/D2/2/C3/4/KC2/2/GC3/3/4/C2/2/P, where the number in bracket represents the length of spacing. Both WAP domains of HE-4 were separately further aligned with WAP domains of other WFDC family members having anti-protease activities assigned to them (pig elafin, human elafin and human SLPI-WAP2) to gain information about function of HE-4 (Fig. 3A, B). It, reveals that in addition to C residues, K-121/72/85, P-122/73/86, G-123/74/87, P-126/77/90, M-134/85/98, N-136/87/100, P-137/88/101, P-138/89/102, N-139/90/103, D-142/95/108, K-150/103/116, G-154/107/120, G-157/110/123 and P-163/116/129 [Representations, Amino acid-position in pig elafin sequence/position in human elafin/position in human SLPI] are fully conserved, and all these may be crucial for anti-protease activity of these domains. Alignment of these domains with HE-4-WAP1 shows that out of 22 conserved residues, 3 residues including second cysteine are not conserved (C-132/83/96, M-134/85/98 and P-137/88/101), and five others are replaced by similar amino acid residues (as indicated by ‘:’ & ‘.’ in alignment). Similarly, some amino acids are not conserved, and some are substituted by similar amino acids in HE-4-WAP2. When we adjusted the optimum alignments, of HE-4-WAP1 and WAP2 domains with known antiprotease activity WAP domains, in such a way that all eight cysteine residues are perfectly conserved, and the similar pattern was observed in both alignments. Overall, six amino acid residues were found fully conserved (For HE-4-WAP1: K32, G34, P37, D52, K60, P73 and For HE-4 WAP 2: K76, G78, P81, D100, K108, P122).

**HE-4 is a cross-class protease inhibitor**

Protease inhibition activity of HE-4 was tested according to the method of Lee and Lin [28]. Hemoglobin was used as a substrate and percentage inhibition was calculated as described in material and methods. Trypsin, chymotrysin, prostate specific antigen (PSA), protease K, pepsin and papain were employed for the experiment, and HE-4 showed dose-dependent inhibition of all of these proteases (Fig. 4A). The dose dependence was more for papain and pepsin and less so for other proteases. Blue native electrophoresis was performed to determine the physical complex formation of HE-4 with all the proteases (Fig. 5). All the proteases tested showed a higher molecular weight complex with HE-4 compared to HE-4 or proteases alone.

Effect of pH, temperature and other chemicals on HE-4 activity was tested using trypsin as a model protease. HE-4 had slightly better trypsin inhibitory activity after incubation at pH 2 and 10 while, at pH 4–6 activity was more or less similar (Fig. 4C). Effect of temperature on HE-4 inhibition of trypsin was undertaken, and activity was measured after incubation of HE-4 at different temperatures. HE-4 inhibited trypsin the most at 25°C and inhibitory activity decreased to 81% at 60°C but even after boiling at 100°C for 10 min, HE-4 retained 29% inhibition (Fig. 4B).
suggesting a highly stable tertiary structure resistant to heat. This is further supported by the observation that HE-4 retained 64% inhibition of trypsin in the presence of SDS. Only 5% SDS itself showed minor inhibition of trypsin activity but much less compared to SDS-treated HE-4 (figure 4D). This suggests that SDS does not denature HE-4 significantly up to 5% concentration. Presence of β-mercaptoethanol (10% v/v) abolished the activity of HE-4 (Figure 4D). A study of effect of disulfide bond reduction on HE-4 activity against trypsin was undertaken using different concentrations of DTT. There was a significant decline in the inhibitory activity of HE-4 against trypsin (29.12%) already at 0.25 mM DTT (Figure 4F). In presence of 1 mM DTT, HE-4 completely lost its inhibitory activity (Figure 4F). This suggests that disulfide bonds based structure stabilization is essential for the protease inhibition.

EDTA had no effect on inhibition while 2 mM ZnCl₂ reduced the activity to 83% and curiously, introduction of twice the molar ratio of EDTA to zinc supplemented sample restored the activity a
Interaction of HE-4 with various proteases as seen by surface plasmon resonance

HE-4 interactions with serine proteases were confirmed with surface plasmon resonance (SPR) and kinetic constants were calculated. HE-4 was immobilized on the chip using EDC/NHS chemistry as described in methods. Different proteases were flowed over the chip in varying concentrations, ranging from 75–300 nM. Concentration of proteases above this range was shown to reduce the signal; possibly due to “Hook effect”. Among all the proteases tested, the highest affinity (K_D) and association constant (K_a) was found to be for proteinase K followed by chymotrysin, PSA and then trypsin (Table 2). Trypsin had the lowest association and dissociation constants among all the proteases tested.

Converse study was also performed with SPR, where all serine proteases were on the chip and HE-4 was flowed. No significant changes were observed in binding affinity of the proteases and HE-4 (Table 2). Unfortunately, we could not determine the kinetic constants for papain and pepsin as we faced an unexpected problem of negative sensograms (Fig. 6A, B). Although, negative sensograms are not uncommon they are usually ascribed to the differences in buffer composition pre and post injection [29]. In the present study, there were no differences in the buffer compositions therefore we sought to further investigate the reason of signal dropping below baseline. For this, papain was incubated with increased concentration of HE-4 (HE-4: papain; 1:1 to 5:1) at pH 8.5 for 1 hr, and as a control HE-4 and papain was incubated separately for 1 hr at room temperature. Later mixtures were resolved on 14% SDS-PAGE under reducing conditions. HE-4 and papain, when incubated alone, showed band at their own molecular weight, but when they are incubated together two new bands appeared in the mixture below HE-4 band as seen in fig. 7A. In SDS-PAGE, we observed two bands that are probably the cleavage product of HE-4 by papain because with increasing concentrations of HE-4, the band at approximately 20 kDa and with increasing concentration of HE-4, the decrease in intensity of pepsin (full size) band. This band at 20 kDa was definitely of papain as western blot using HE-4 antibodies (Figure 7C Lane 3) revealed that there was no band of HE-4 at that position. This only partially explains the sensogram of pepsin in fig 6A and does not explain why the signal went down below the baseline unless there was also the minor cleavage of HE-4 which was not detectable in SDS-PAGE or WB because SPR would be more sensitive to even small amount of cleavage. The sensograms of other proteases were characteristic of a normal protein-protein interaction (Figure 8).

Discussion

HE-4 is upregulated in ovarian carcinoma and also expressed in a variety of lung tumors. There are several reports in the literature which suggest important functions of HE-4: over-expression in ovarian carcinoma, over-expression in prostate cancer model mice with PTEN inactivation, interaction with pleiotrophin (PTN) which regulates angiogenesis and is involved in tumor formation as well as the upregulation of HE-4 during the expected window of receptivity in the endometrium under the control of progesterone in primates [20,30–32]. All of these evidences pointed towards an important role of HE-4 in human physiology but to date no study has tried to investigate the role or structure of HE-4 in normal or pathological conditions. We have developed a rapid and efficient method for purification of native HE-4 from human seminal fluid, which will help study the function and structure of HE-4.

HE-4 is glycosylated and highly stable protein

HE-4 exists as a disulfide bonded trimer in human seminal fluid as inferred from SDS-PAGE in reducing and non-reducing conditions. This is not unheard of in proteins and many human and some viral proteins employ intermolecular disulfide linkages to form tertiary structure [33–34]. HE-4 has eight predicted disulfide bonds per monomer of protein as it is a small protein, these bonds and intermolecular disulfide linkages would be expected to give it a compact structure resistant to denaturing agents. Accordingly, it found that HE-4 is resistant to pH, heat and even SDS in protease inhibition assay taking trypsin as a model protease. Seminal fluid has a high concentration of zinc and it regulates the function of several seminal fluid proteins like PSA and sememngelin [35]. Therefore, we measured R_h value and activity of HE-4 in...
the presence of zinc and we observed lower R_h than that of purified native protein in Tris buffer. Introduction of EDTA in twice the molar ratio of zinc increased the R_h value a little bringing it closer to native HE-4. Trypsin inhibition activity of HE-4 reduced slightly in the presence of zinc (mean: 83%) while the addition of EDTA (twice the molar ratio) to the Zn supplemented HE-4 recovered the activity a little (mean: 85.6%) as shown in figure 4D. The activity lost by 2 mM ZnCl_2 could be completely rescued by addition of increasing concentrations of EDTA (Figure 4E). At this stage evidence for mechanism of effect of zinc on activity and structure of HE-4 is inconclusive. However, this idea is being pursued further in our laboratory. Disulfide bonds play an important role in maintaining the native conformation of a protein, which in turn provide stability/resistance towards pH and temperature treatments. HE-4 contains 8 disulfide bonds. Therefore, it was of interest to evaluate the effect of DTT reduction on the trypsin inhibitory activity of HE-4. Even at .05 mM there was a significant reduction in trypsin inhibitory activity of HE-4 and at 1 mM HE-4 lost all its activity against trypsin (Figure 4F). Disulfide bond reduced HE-4 was oxidized and refolded (as described in methods) but HE-4 was unable to restore the activity and at 16 hr. time-point only 3% activity was restored. PNGase F treatment of HE-4 produced a shift of approximately 2.5 kDa which is considerable given the molecular weight of the monomer (Figure 2B). This also partly explains the observed heat and pH resistance of the protein. Asn44 has been reported to be glycosylated in salivary HE-4 previously [36], and no other glycosylation site has been either predicted or reported. This implies that Asn44 which lies in the N-terminal WAP domain is heavily glycosylated. One previous study has reported HE-4 to be a secreted glycosylated protein of approximately 25 kDa in two ovarian carcinoma cell lines [32]. This difference with seminal fluid form of protein could be cell line specific, or it could be another isoform of the protein resulting from alternative splicing [36]. HE-4 can undergo alternative splicing to yield four different isoforms other than full length protein. Some of these isoforms has only N-terminal WAP domain and some C-terminal WAP domain while three of these isoforms own unique sequence not found in others. So this 25 kDa isoform found in ovarian cancer lines could be different from the presumably full length (based on molecular weight from SDS-PAGE) protein we have in seminal fluid, and it might even lack trimerization property, or have a different glycosylation pattern. Importance of glycosylation in functions of HE-4 (seminal fluid or other sites) requires further studies which are underway in the laboratory.

HE-4 is a disulfide-bonded trimer and bioinformatic analysis alone does not correctly predict its function

A consensus sequence was obtained by aligning HE-4 sequences from different species (Fig. 3C). Though the high degree of similar conservation of residues and spacing between them points towards a similarity in functional aspects of both domains, but we cannot deny the fact that length difference between first and second cysteine residues may contribute to the different functions of these domains or the ability to covalently oligomerise, which requires inter-molecular disulfide bridges. WFDG domain of elafin and one
of the two WFDC domains of SLPI is known to impart anti-proteinase activity to both of these proteins. Therefore, it is a widely held notion that all of the members of the family will have anti-proteinase activity. Sequence analysis by Bingle et al showed that only these two members have the same spacing between cysteines essential for protease inhibition [19]. Authors of the paper argue that no other member of the WFDC family contains the same spacing, so it is not necessary that they will have anti-proteinase activity. A separate study by Hu et al show that both wild type and WAP mutated KAL-1 enhanced amidolytic activity of uPA (urokinase-type plasminogen activator) [37] which are consistent with the prediction made by Bingle et al. This study adds a new dimension in this discussion about function of WFDC family members as we have observed antiprotease activity in HE-4. The possible reason for this discrepancy might be this spacing playing some role in inhibitory activity of SLPI and elafin. Both of these proteins are monomers while, we found HE-4 to be a disulfide bonded trimer so this rearrangement of the structure might give HE-4 unusual properties not predicted by sequence analysis. Moreover, we found that disulfide bond reduction abolishes the protease inhibition by HE-4, so we suggest that, in case of HE-4, this protease inhibition does not depend upon any of the two domains (N- and C-terminal WAP domain) alone as whole trimer seems to be necessary for the inhibition. Although one can speculate that HE-4 might be an example of inhibitors where single domains are repeated and linked together (as in ovomucoid) and this new single chain inhibitor can inhibit many different proteases [38]. The possibility of three monomers linked together by disulfide bridges point towards a compact, more rigid structure, and might resemble mechanism of ecotin in which, a homodimer is active, and both monomers provide the protease binding surface [39]. Further studies are needed to understand the mechanism of inhibition of a wide range of proteases by HE-4 which will highlight the residues crucial for inhibition; however, proximity of a variety of residues induced by trimerization might be necessary. Although we do not exclude the possibility that monomer folding pattern of HE-4 as compared to SLPI and elafin is not markedly different, so some common residues might be beneficial for protease inhibition by these proteins. Multiple sequence alignment was performed with both the WAP domains of HE-4 with WAP domains of WFDC family members having known protease inhibition activity (Fig. 3A, B). Some key amino acids which are suggested to be necessary for antiprotease activity of these domains are not conserved in HE-4, however, there is more than 70% conservation overall. More importantly, it is not always feasible to justify function of HE-4 simply on the basis of conservation of

Figure 4. Antiprotease activity of HE-4 at different concentration and at different condition against various serine, cystine and aspartic proteases. (A) Percentage inhibition of various proteases like: trypsin, chymotrypsin, PSA, proteinase K, pepsin and papain by HE-4. X-axis is the HE-4 concentration in µg/ml and Y-axis is the % inhibition of respective proteases. (B) Effect of temperature (25 °C–100 °C) on the protease inhibition activity of HE-4 against trypsin. (C) Effect of pH (2–10) on the protease inhibition activity of HE-4 against trypsin. (D) Effect of different chemicals like: ZnCl2, SDS and β-mercaptoethanol on the protease inhibition activity of HE-4 using trypsin as model protease. EDTA was added to ZnCl2 pre-treated HE-4 to check whether it can rescue the decrease in activity induced by ZnCl2. SDS control is the % inhibition of trypsin by 5% SDS and 5% SDS point is the activity of SDS-pretreated HE-4 against trypsin. +ve control is HE-4 inhibition of trypsin without any treatment and −ve control is without HE-4, therefore 0% inhibition. (E) Effect of varying concentrations of EDTA on rescue of trypsin % inhibition decrease by ZnCl2 treatment of HE-4. (F) Effect of varying concentration of DTT on trypsin inhibitory activity of HE-4.

doi:10.1371/journal.pone.0047672.g004
critical residues in primary sequence. Factors such as the overall structure of the domain, exposure of some specific amino acids and types of residues lining the active site may be more vital, and contribute to the activity as may be true in this study.

HE-4 is cross-class protease inhibitor which is cleaved by papain and induces autolysis of pepsin in vitro

HE-4 inhibited a range of serine proteases like trypsin, chymotrypsin, PSA and proteinase K as well as cysteine proteases like papain and aspartyl proteases like pepsin. The physical complex formation of HE-4 with all these proteases was confirmed with blue native electrophoresis where complexes were migrating less compared to HE-4 or proteases alone (Figure 5). Blue native electrophoresis was chosen because some of these protease like trypsin are basic proteins therefore they do not migrate towards anode and are lost. Fig. 4A shows the protease inhibition assay results starting at 5 μg/ml concentration of HE-4 up to 50 μg/ml, and at 50 μg/ml concentration, it inhibits almost completely all the tested proteases. SLPI inhibits trypsin and chymotrypsin with K_{d} which is not markedly different for both of these proteases [40] while HE-4 has low affinity towards trypsin and considerably higher affinity towards chymotrypsin (Table 2) as determined by SPR. This highlights different inhibition profile for HE-4 and SLPI. SLPI is thought to neutralize the excess neutrophil protease activity in upper airways. HE-4 is also expressed in sub-mucosal glands of respiratory tissues, but although SLPI and HE-4 both are found in the same tissue, the precise cells expressing both these proteins are mutually exclusive [16]. This suggests them to be under different regulatory control which point towards slightly different functions of both these proteins. HE-4 has a strong affinity towards PSA at pH 5.0 using 75–300 nm PSA (K_{D} = 1.06 × 10^{-7} M & K_{A} = 9.40 × 10^{4} M^{-1}) and suggests that this protein might be involved in regulation of kallikrein activity

**Table 2. Binding affinity of HE-4 with different proteases.**

| Protease | HE-4 | Protease | HE-4 |
|----------|------|----------|------|
| Proteinase K | 4.05 × 10^{-4} | PSA | 4.05 × 10^{-4} |
| Chymotrypsin | 7.56 × 10^{-4} | Trypsin | 4.06 × 10^{-4} |
| PSA | 4.46 × 10^{-5} | | 9.90 × 10^{-5} |

In separate experiments HE-4 or proteases were immobilized on CMS chip and various proteases or HE-4 was flowed over it.

doi:10.1371/journal.pone.0047672.t002

**Figure 5. Blue native gel of HE-4 after 2 hr. incubation with different proteases.** Proteases and HE-4 alone were also run to compare with the complexes formed with them. 1: Marker 45 kDa Ovalbumin, A: HE-4, B: Trypsin, C: Trypsin+HE-4, D: Chymotrypsin, E: Chymotrypsin+HE-4, F: Proteinase K, G: Proteinase K+HE-4, H: PSA, I: PSA+HE-4, J: Pepsin, K: Pepsin+HE-4, L: Papain, M: Papain+HE-4.
doi:10.1371/journal.pone.0047672.g005
cascades. Protein C inhibitor (PCI), α-2 macroglobulin (A2M) and α1-anti chymotrypsin (ACT) are three main inhibitors of PSA reported in human seminal fluid. Concentrations of these inhibitors in seminal fluid are low compared to PSA [41–43] suggesting there must be other inhibitors, which serve as main inhibitors. HE-4 seems to be one such inhibitor. PSA has been implicated in liquefaction of human seminal fluid upon ejaculation and HE-4 probably regulates the activity of PSA in semen and might protect sperm against excessive PSA activity. HE-4 inhibited proteinase K (75–300 nm) with highest affinity among the proteases tested as apparent by highest \( K_a = 2.15 \times 10^7 \text{ M}^{-1} \) & \( K_D = 4.65 \times 10^{-8} \text{ M} \) at pH5.0 (Table 2). Proteinase K is a member

Figure 6. SPR sensograms of HE-4 interaction with papain and pepsin. Concentrations of HE-4 were fixed at 10 µM. Concentrations of papain and pepsin were 75 nM. Injection time of papain and pepsin is indicated by an arrow. (A) Pepsin-HE-4 interaction. (B) Papain-HE-4 interaction. doi:10.1371/journal.pone.0047672.g006

Figure 7. 14% SDS-PAGE to resolve HE-4 and pepsin, papain incubated alone or together for the same duration. (A) 14% SDS-PAGE (silver stained) Lane1: Molecular weight marker. Lane2: Fresh HE-4. Lane3: HE-4 after 1 hr incubation. Lane 4: Fresh papain. Lane5: papain after 1 hr incubation. Lane 6- Lane 10 papain and HE-4 in 1:1–1:5 (10 µg:10 µg–10 µg:50 µg) ratio after 1 hr incubation (in 50 mM tris-HCl buffer, pH 8.5). (B) 14% SDS-PAGE (silver stained) Lane1: Molecular weight marker. Lane2: Fresh HE-4 (10 µg). Lane3: HE-4 after 1 hr incubation (10 µg). Lane 4: Fresh pepsin (10 µg). Lane5: pepsin (10 µg) after 1 hr incubation. Lane 6- Lane 10 pepsin and HE-4 in 1:1–1:5 ratio (10 µg:10 µg–10 µg:50 µg) after 1 hr incubation (in 50 mM sodium acetate buffer, pH 5.0). (C) Immunodetection of HE-4 after incubating HE-4 alone and with pepsin and papain for the same duration. Lane1: HE-4 (10 µg) after 1 hr incubation. Lane 2: HE-4 (10 µg) after 1 hr incubation with papain (10 µg) and Lane3: HE-4 (10 µg) 1 hr incubation with pepsin (10 µg). doi:10.1371/journal.pone.0047672.g007
of subtilisin like protease family, and it has been shown that structures of most members is conserved as a core with insertions and deletion confined to surface loops [44]. This suggests that HE-4 might be a broad spectrum inhibitor of microbial subtilisin like proteases although they need. Many pathogenic fungi utilize subtilisin-like serine proteases as their virulence factors [45–48] and HE-4 as a part of the host response might be conferring protection in the sites of its expression one of which is the respiratory tract, others being oral cavity and urogenital tract.

We could not determine the kinetics constants for HE-4 with pepsin and papain due to the negative sensogram obtained. On further investigation, we found that papain when incubated with HE-4, presents two more bands in SDS-PAGE below the molecular weight of HE-4. With increasing concentration of HE-4, the HE-4 band remains more or less same in intensity while one of the bands below it (approx 8 kDa) seems to increase in intensity and intensity of papain band also remains the same. This band of cleaved HE-4 was confirmed by western blot (Figure 7C) using HE-4 antibodies. This leads us to propose that papain cleaves HE-4 to generate low molecular weight fragments. The sensogram of papain injection over immobilized HE-4 on CM5 chip shows initial interaction and subsequent dropping of signal below baseline which is in accordance with papain cleavage of HE-4. In protease inhibition assay, HE-4 shows considerable inhibition of papain which leads us to suggest that even after cleavage, the fragment(s) of HE-4 remain(s) bound to papain inhibiting its proteolytic activity. This was confirmed with blue native electrophoresis as papain incubated with HE-4 presented only one band of high-molecular weight (Figure 5). Contrastingly, when pepsin was incubated with HE-4, it seemingly underwent self-cleavage, as evidenced by the appearance of approximately 20 kDa band which increased in intensity with increasing concentration of HE-4 with simultaneous reduction in intensity of full length pepsin band. Pepsinogen undergoes autolysis, and yield active pepsin (approximately 35 kDa) at acidic pH, while we observed a 20 kDa fragment upon incubation with HE-4. Moreover, in both the incubation experiments, we had taken control of incubating the papain and pepsin alone at appropriate pH 8.5 and 5 respectively, and these fragments of HE-4 cleavage by papain and self-cleavage of pepsin did not appear when these proteases were incubated alone. This led us to conclude that HE-4 induces the self cleavage of pepsin and papain cleaves HE-4 and the fragment of HE-4 remains active. In blue native electrophoresis we observed only one band of HE-4 (full length) when pepsin and HE-4 were incubated together (Figure 5). This suggests that pepsin cleavage also does not dissociate HE-4 from pepsin. Cysteine proteases are implicated in a variety of processes in mammalian physiology. This inhibition of papain by HE-4 fragment has greater implications for cell biology. Papain like cysteine proteases has a conserved core structure [49], and highly conserved catalytic site formed by three residues Cys25, His159, and Asn175 (papain numbering) [50]. HE-4 might inhibit other members as well, or they like papain, might be involved in cleavage of HE-4, and producing fragment, which seems to be active, although it remains speculative until further studies. IA3, PI-3 and equistatin, all inhibitors of cysteine proteases like papain with different specificities and targets have been isolated from different animal and microbial sources [51]. Cruzain is a papain like protease of trypanosoma cruzi, causative agent of chagas disease and its inhibitors are being looked into as possible drug development leads [52]. Structure elucidation of HE-4 and its inhibition mechanism of papain despite cleavage might help design potent inhibitors. While, on the other hand, discovery of self cleavage of pepsin when combined with protease inhibition of pepsin by HE-4 in protease inhibition assay point that HE-4 renders pepsin inactive by inducing the self-cleavage. Induced self-cleavage of a protease by a protease inhibitor has not been reported in literature previously to the best of our knowledge. This is an exceedingly curious instance, and further characterization of this phenomenon is underway in the laboratory. To confirm whether HE-4 preparation is contaminated with any endogenous protease co-purifying with HE-4, we assayed HE-4 alone with hemoglobin and BAPNA and compared to other proteases used in this study. No protease activity was found in the purified HE-4.
confirming that this self-cleavage of pepsin is due to HE-4. Aspartic proteases are extremely valuable drug targets including the retroviral family and fungal aspartic proteases. Although there are structural differences between retroviral and eukaryotic pepsin-like proteases, there are similarities as well; the cleavage site loops are homologous, the Asp dyad is located in the interface region and N-terminal lobe of pepsin like enzymes are structurally similar to viral subunits [53–54]. HE-4 might help design better aspartic protease inhibitors in the future.

Finally, we can say this is the first report to establish that HE-4 is a highly stable protein which shows cross-class protease inhibition. A broad spectrum protease inhibition points towards a role in innate immunity conferring protection against microbial virulence factor of proteolytic nature. Seminal fluid HE-4 might be different from HE-4 found in other tissues like ovarian cancer cells because of alternative splicing or different glycosylation. Its ability to trimerize might not be present in all the isoforms.

**Materials and Methods**

**Sample collection**

Human semen samples, only normospermic (sperm count >20 million/ml, sperm motility >50%), were collected from Department of Laboratory Medicine, All India Institute of Medical Sciences (AIIMS), New Delhi, after written informed consent and the approval of the study protocol from ethics sub-committee/ ethics committee of AIIMS (permit number T-03/01-04-2009). Semen samples were first subjected to liquefaction at room temperature (RT) for 30 min. Semen sample was centrifuged at 1,500 g for 15 min at 4°C to separate sperm from seminal plasma. Later, for further clarification of seminal plasma supernatant was centrifuged at 7,000 g for 15 min. at 4°C.

**Isolation and purification of HE-4**

The supernatant, diluted with 50 mM Tris–HCl, pH 7.5, containing 150 mM NaCl, was loaded on heparin–sepharose CL-6B (GE-Healthcare, Uppsala, Sweden) column. The unbound fraction was pooled separately and applied on DEAE-Sephacel column. 50 mM Tris–HCl (pH 8.0) was used as equilibration/ binding buffer. After extensive washing, DEAE-sephacel bound proteins were eluted with NaCl linear gradient (0–0.5 M) in equilibration buffer. The first peak, obtained at 0.1 M NaCl was pooled and concentrated up to 20 mg/ml by ultrafiltration (Millipore USA). Final purification of protein was achieved by size exclusion chromatography on sephadex G-75 (Sigma–Aldrich, USA) column, pre-equilibrated with 50 mM Tris–HCl (pH 8.0, containing 150 mM NaCl). Fractions eluted at the flow rate of 6 ml/hr were measured at 280 nm, and pooled separately for each peak and concentrated by ultrafiltration using 3 kDa membrane cut-off. 12.5% SDS-PAGE was performed as previously described [55] to analyze approximate molecular weight and purity of the protein. Finally, gel was stained with colloidal coomassie brilliant blue (CBB).

**Kinetic analysis**

**Antiprotease activity of HE-4 against different proteases and effect of various treatments.** Inhibitory activity of different serine proteases like trypsin, chymotrypsin, PSA, proteinase K and cysteine proteases like papain and asperyl proteases like pepsin were measured with the modified method described by Lee and Lin [28]. The sample assay was as the following: 250 μl of HE-4 trimer purified from human seminal plasma was pre-incubated, at 37°C, with the same volume of proteases dissolved in 0.1 M glycine–NaOH buffer (pH 9.5) for 20 min. All the enzymes were purchased from Sigma, except PSA which was purified manually as previously described [56]. Papain and pepsin were dissolved in 50 mM ammonium acetate buffer (pH 6.5) and 50 mM sodium acetate buffer (pH 5.0) respectively. 500 μl of 1% solution of hemoglobin dissolved in the same buffer (Sigma- Aldrich, USA) was added to it, and the mixture was incubated for 40 min. The reaction was stopped by adding 2 ml of 5% trichloroacetic acid. Samples were centrifuged at 15,000 g for 10 min, and the absorbance of the supernatant was measured at 280 nm. The enzyme standard assay was as follows: 250 μl of a sample was replaced by distilled water. The chosen enzyme concentrations gave an increase of absorbance at λ 280 of approximately 0.005 OD U/min. With purified protein (HE-4), a control assay to detect the activity of endogenous proteases was also performed: 250 μl of an enzyme solution was replaced by distilled water. The percentage of inhibition was calculated as follow:

$$\frac{\text{io280 (Enzyme+Sample)} - \text{io280 (Control+Sample)}}{\text{io280 (Enzyme+Control)}} \times 100$$

Additionally, effect of temperature, pH and different chemicals on trypsin inhibition activity of HE-4 was also analyzed. For that we first incubated the 50 μg/ml of protein (HE-4) at different temperature in the range of 25°C–100°C, pH(2–10) and with different chemicals like SDS, β-mercaptoethanol, EDTA, ZnCl₂, and MnCl₂+EDTA for 1 hr and the following experiments were performed as described above. In case of SDS HE-4 was incubated with 5% SDS for 2 hours before being added to trypsin and checking the activity as described above. For SDS control, inhibition of trypsin by only 5% SDS without the HE-4 was also checked as shown in fig. 4D as SDS-control. To determine the effect of zinc on HE-4 activity, HE-4 was pre-incubated with 2 mM ZnCl₂ for 2 hr, and then activity was checked. Then in the aliquots of the same sample, different concentrations of EDTA were added and activity was checked. The effect of DTT reduction on inhibitory activity of HE-4 was examined after incubation of HE-4 with different concentrations of DTT (0.05–1.0 mM) in 25 mM NH₂HCO₃ for 15 min at 56°C. The reaction was terminated by adding iodoacetamide at twice the amount of each DTT concentration and the residual inhibitory activity against trypsin was determined as described above. Refolding and oxidation assay was performed as described previously [57] to check whether activity can be restored after reduction. Briefly, HE-4 was incubated at 37°C for 4 hr with 0.5 M phosphate buffer and 10 mM DTT then reaction was stopped by adding iodoacetamide [IAA] so that final concentration of IAA is 20 mM. It was dialyzed against 0.1 M KCl-HCl buffer (pH 2.0) for 3 hr. Then it was followed by dialysis against 0.01 M of the same buffer for 16 hr. at 4°C. The dialysed protein was rapidly diluted 100times with bufferA (100 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA (pH 8.5), 1 mM GSH and 0.5 mM GSSG). The mixture was kept at 25°C for refolding to occur. Aliquots were withdrawn at different time intervals (0, 1, 2, 4, 8, 16 hours) and trypsin inhibitory activity was determined as described above.

**Blue Native gel electrophoresis**

Blue Native electrophoresis was performed as described previously [38]. Briefly, all proteases were incubated with equal amount of HE-4 for 2 hours at RT. Then loading Buffer (15% glycerol, 50 mM Bis/tris/HCl, pH 7.0) was added to complex mixtures. Different proteases and HE-4 alone were also run in separate lane to make the comparison with complex with HE-4 and protease alone. 5–10% acrylamide gradient gel was used for...
separation of complexes and 50 mM tricine, 15 mM bis-tris/HCl pH 7.0 was used as anode buffer and it remained the same for the whole run. Blue native gel was performed at 4–7°C. Electrophoresis was started at 100 Volts until samples were within stacking gel. When sample reached resolving gel 15–17 mA current was applied. Gel was run for total 3–4 hr.

Evaluation of binding potential and binding constants using SPR binding studies

The protein sensor chip was prepared by immobilization of HE-4 on a research grade CM5 chip (Biosensor AB, Uppsala, Sweden) according to the manufacturer’s recommendations. SPR binding studies were performed using a BIACore 2000TM (Biacore International, AB, and Uppsala, Sweden). 60 μl of HE-4 (30 μg/ml) in 10 mM sodium acetate buffer (pH 5.0) was injected to the flow cell for 4 min at the rate of 5 μl/min and washed with 60 min at the rate of 20 μl/min. The analytes at different concentrations were injected for 4 min at a rate of 10 μl/min.

The protein samples were diluted in sodium acetate buffer pH 5.0. Typically, 90 μl of different dilutions of protein were injected at a flow rate of 30 μl/min. The same buffer was passed over the sensor surface at the end of sample injection, to allow dissociation. After a 3 min dissociation phase, the sensor surface was regenerated by 30 μl of 50 mM NaOH and 1 M NaCl. The response was monitored as a function of time (sensogram) at 25°C. Binding studies were done by plotting the changes in response unit (RU) values with time (Sensograms). Kinetic constants of different proteases were calculated from the association and dissociation phases with BIACore evaluation software version 3.0. Similarly, the negative control was also performed with SPR where all serine proteases were immobilized on the chip and HE-4 was flow.

Supporting Information

Figure S1 Effect of different reagents and conditions on the oligomerization of HE-4 was measured by DLS. Plot was drawn as a function of concentration of reagents and conditions as a function of X axis with hydrodynamic radii of protein on Y axis. (A) Effect of pH 2–10 (empty circle) and different salt: MgCl2, FeCl2, CuCl2, CaCl2, ZnCl2 and ZnCl2+EDTA (filled circle). (B) Effect of DTT (filled circle) and CHAPS (empty circle) in varying concentration (1%–5%). (C) Effect of SDS (filled circle) and Tween-20 (empty circle) in varying concentration (2.5%–1%) and (D) Effect of Triton X-100 (empty circle) and β-mercaptoethanol (filled circle) in varying concentration (0.1%–1% and the 5%–10%) respectively. (TIF)

Supplemental Methods S1

Author Contributions

Conceived and designed the experiments: NC MS SY. Performed the experiments: NC AKT. Analyzed the data: NC MS SDY. Contributed reagents/materials/analysis tools: SD SS. Wrote the paper: NC AKT MS SY.

References

1. Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437: 69–87.
2. Hurle B, Swanson W, NISC Comparative Sequencing Program, Green ED (2007) Green Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WIFDC locus in the primate lineage. Genome Res 17: 276–296.
3. Larsen M, Ressler SJ, Gerdes MJ, Lu B, Byron M, et al. (2000) The WFDC1 gene encoding a novel secreted protein with multiple types of protease-inhibitory modules. Proc Natl Acad Sci U S A 97: 13047–13052.
4. Kirchhoff C, Halbhuber K, Ertl R, Kraft N (1991) A major human epididymis protein 4 (HE-4) is a secreted glycoprotein that is overexpressed by human kallikrein 7 (hK7/SCCE) and its inhibitor antileukoprotease (ALP/SLPI) in uterine endocervical glands and in cervical adenocarcinomas. Oncol Rep 11: 507–512.
5. Tian X, Shigemasa K, Hirata E, Gu L, Uebaba Y, et al. (2004) Expression of human normal mammary epithelial cells and carcinomas is regulated at the transcriptional level. Cancer Res 65: 2537–2541.
6. Schalkwijk J, Wielow O, Hirose S (1999) The trappin gene family: proteins defined by an N-terminal conserved tandem protease substrate domain and a C-terminal four-disulphide core. Biochem J 340: 569–577.
7. Trexler M, Bányai L, Patthy L (2002) Distinct expression pattern of two related human proteins containing multiple types of protease-inhibitory modules. Biochem Biophys Res Commun 293: 223–228.
8. del Castillo I, Cohen-Salmon M, Blanchard S, Luffalla G, Petit C (1992) Structure of the X-linked Kallmann 452 syndrome gene and its homologous pseudogene on the Y chromosome. Nat Genet 2: 305–310.
9. Williams SE, Brown TI, Roghanian A, Sallenave JM (2006) SLPI and elafin: one glove, many fingers. Clin Sci (Lond.) 110: 21–35.
10. Larramendy ML, Lushnikova T, Bjerkjæst AM, Wistuba II, Virmani AK, et al. (2000) Comparative genomic hybridization reveals complex genetic changes in primary breast cancer tumors and their cell lines. Cancer Genet Cytogenet 119: 173–180.
11. Schlosser N, Hassanzadeh Ghassabeh G, Zhang J, Brens L, De Baetselier P, et al. (2003) Secretory leukocyte protease inhibitor promotes the tumorigenic and metastatic potential of cancer cells. Proc Natl Acad Sci USA 100: 3705–3709.
12. Trexler M, Bányai L, Pathy L (2002) Distinct expression pattern of two related human proteins containing multiple types of protease-inhibitory modules. Biol Chem 383: 223–228.
13. Bingle L, Cross SS, Wallace WA, Raasd D, et al. (2006) WIFDC2-HE4: a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. Respir Res 7: 61.
14. Williams SE, Brown TI, Roghanian A, Sallenave JM (2006) SLPI and elafin: one glove, many fingers. Clin Sci (Lond.) 110: 21–35.
15. Sallenave JM, Shulmann J, Crossley J, Jordana M, Gauldie J (1994) Regulation of secretory leukocyte protease inhibitor (SLPI) and elastase-specific inhibitor (ESI/elafin) in human airway epithelial cells by cytokines and neutrophilic enzymes. Am J Respir Cell Mol Biol 11: 733–741.
16. Bingle L, Cross SS, Wallace WA, Raasd D, et al. (2006) WIFDC2-HE4: a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. Respir Res 7: 61.
17. Kirchhoff C, Halbhuber K, Ertl R, Kraft N (1991) A major human epithelinspecific cDNA encodes a protein with sequence homology to extracellular protease inhibitors. Biol Reprod 45: 530–537.
18. Liu X, Shigemasa K, Hirata E, Gu L, Uebaba Y, et al. (1996) Expression of human kallikrein 7 (hK7/SCCE) and its inhibitor antileukoprotease (ALP/SLPI) in uterine endocervical glands and in cervical adenocarcinomas. Oncol Rep 1: 1033–1038.
19. Bingle CD, Vyakarnam A (2008) Novel innate immune functions of the whey acidic protein family. Trends Immunol 29: 444–453.
20. Israeli O, Goldring-Aviram A, Rienstein S, Ben-Barguch G, Korach J, et al. (2005) In silico chromosomal clustering of genes displaying altered expression patterns in ovarian cancer. Cancer Genet Cytoenet 160: 35–42.
21. Schalkwijk J, Wielow O, Hirose S (1999) The trappin gene family: proteins defined by an N-terminal conserved tandem protease substrate domain and a C-terminal four-disulphide core. Biochem J 340: 569–577.
22. Schaner ME, Ross DT, Ciaravino G, Sorlie T, Troyanskaya O, et al. (2003) Gene expression patterns in ovarian carcinomas. Mol Biol Cell 14: 4376–4386.
23. Bingle L, Cross SS, Wallace WA, Raasd D, et al. (2006) WIFDC2-HE4: a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. Respir Res 7: 61.
HE4 Inhibits a Wide Range of Proteases