Protein Expression and Purification 102 (2014) 52–62

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Expression and biochemical characterization of recombinant human epididymis protein 4

Ling Hua [a,b], Yunhui Liu [b], Shuai Zhen [b], Deyou Wan [b], Jiyue Cao [a,*], Xin Gao [b,*]

[a] Laboratory of Veterinary Pharmacology and Toxicology, College of Veterinary Medicine, Huazhong Agricultural University, 1 Shi-zi-shan Street, Wuhan 430070, Hubei, PR China
[b] Department of Pharmacology and Toxicology, Beijing Institute of Radiation Medicine, Beijing 100850, PR China

ARTICLE INFO

Article history:
Received 9 May 2014
and in revised form 1 August 2014
Available online 13 August 2014

Keywords:
Recombinant human epididymis protein 4
Antiprotease activity
Antimicrobial activity
N-glycosylation

ABSTRACT

Whey acidic proteins (WAP) belong to a large gene family of antibacterial peptides that perform critical immune system functions. The function of human epididymis protein 4 (HE4), a 124-amino acid long polypeptide that has two whey acidic protein four-disulfide core (WFDC) domains, is not well studied. Here, a fusion gene encoding the HE4 protein fused to an IgG1 Fc domain was constructed. The recombinant HE4 protein was expressed as a secretory protein in *Pichia pastoris* and mammalian HEK293-F cells and was subsequently purified. Our data suggested that the HE4 protein produced by these two expression systems bound to both gram-negative and gram-positive bacteria, but demonstrated slightly inhibitory activity towards the growth of *Staphylococcus aureus*. Moreover, HE4 exhibited proteinase inhibitory activity towards trypsin, elastase, matrix metallopeptidase 9, and the secretory proteinases from *Bacillus subtilis*. The effects of glycosylation on the biochemical characterization of HE4 were also investigated. LC–ESI–MS glycosylation analysis showed that the high-mannose glycosylated form of HE4 expressed by *P. pastoris* has lower biological activity when compared to its complex-glycosylated form produced from HEK293-F cells. The implications of this are discussed, which may be provide theoretical basis for its important role in the development of cancer and innate immune system.

Introduction

Human epididymis protein 4 (HE4)[1] is a secreted glycoprotein that was first identified in the epithelium of the distal epididymis and initially predicted to be a protease inhibitor involved in sperm maturation [1–3]. HE4 is reportedly expressed in many normal tissues including the epithelia of respiratory and reproductive tissues, as well as in several tumor cell lines including ovarian, lung, colon, and breast cancer cell lines [4,5]. HE4 levels in the serum were proposed to be a sensitive histological marker for ovarian cancer [1,6], as well as a potential biomarker and therapeutic target for renal fibrosis and lung cancer [7,8].

HE4 is also referred to as WFDC2 because it contains two whey acidic protein four disulfide core (WFDC) domains, which contain WFDC motifs that are made up of eight cysteine residues [9]. Some members of the four-disulfide core family of proteins act as protease inhibitors, because the WFDC domain can form a looped structure that binds to the active site and inhibit serine proteases [10]. These include secretory leukocyte protease inhibitor (SLPI) and elafin, which inhibit endogenous proteases such as elastase and trypsin [11,12]. They also inhibit exogenous proteases secreted by microorganisms, as they display antibacterial, antifungal, and antiviral properties [13].

Based on the structural and sequence similarity of HE4 to other whey acidic proteins (WAPs) such as SLPI and elafin, the protein is suggested to have antiprotease and antimicrobial activities within the male reproductive, oral, and respiratory tracts. However, the activity of HE4 has not been well characterized in vitro. To achieve this, a rapid and efficient method to generate sufficient amounts of protein is required.

The expression of disulfide-rich proteins such as SLPI in *Escherichia coli* is reportedly time-consuming because it can result in non-native, inactive products that require reductive denaturation and oxidative refolding to recover the native, active conformation [14,15]. Mammalian cells have the proper requisites enabling

* Corresponding authors. Tel.: +86 2787281593 (J. Cao), +86 1066930252 (X. Gao).
E-mail addresses: caojiyue@mail.hzau.edu.cn (J. Cao), gaoxinamms@163.com (X. Gao).

[1] *Abbreviations used:* HE4, human epididymis protein 4; WFDC, whey acidic protein four-disulfide core; SLPI, secretory leukocyte protease inhibitor; WAPs, whey acidic proteins; LB, Luria–Bertani; BMGY, buffered complex glycerol mediums; YPD, yeast extract peptone dextrose; BMMY, buffered complex methanol mediums; PCR, polymerase chain reaction; Fc, fragment crystallizable; OD_{600}, optical density at 600 nm.

http://dx.doi.org/10.1016/j.pep.2014.08.004
1046-5928/

native folding and modification, which is a suitable expression system for studying the biochemical characterization of such a small protein. Meanwhile, yeast-based expression systems offer a variety of advantages including shorter fermentation times, cheaper operating costs, chemically defined media, lack of viral contamination, higher protein titers, and significantly shorter development times from gene to protein, which also has the ability to perform post-translational modification. Therefore, we selected two different expression systems: the methylotrophic yeast *Pichia pastoris* and mammalian HEK293-F cells in this study.

Glycosylation is by far the most common and complex post-translational modification, and is crucial for correct folding, stability, and bioactivity of proteins [16–19]. In higher eukaryotes, cells synthesize complex oligosaccharides, and the oligosaccharide moieties are implicated in a wide range of cell–cell and cell-matrix recognition events that are required for biological processes ranging from immune recognition to cancer development [20]. In lower eukaryotes such as yeast, glycosylation is the high-mannose type. This hyper-mannosylation of secreted glycoproteins hampers the downstream processing of heterologously expressed glycoproteins and leads to a short in vivo half-life to the protein [21], which may render it less efficacious or even immunogenic [22]. Generally, glycoproteins of various origins have variable glycosylation patterns that might affect their physical and biochemical properties [23,24]. Owing to the different glycosylation between *P. pastoris* and mammalian cells, the role of glycosylation in biochemical functions of HE4 has also been elucidated.

Here, we report the data obtained regarding expression, purification, and biophysical characterization of HE4 expressed in *P. pastoris* and mammalian cells. Further, we have characterized the effect of glycosylation on HE4 biological activity, which can provide theoretical basis for its important role in the development of cancer and innate immune system.

Materials and methods

Strains, cell culture, and enzymes

P. pastoris GS115 (His4-deficient) was used as the expression host. The TOP-10 *E. coli* strain was purchased from Biomed Corp. (Beijing, China). The *E. coli* cells transformed with plasmids were cultured at 37 °C in Luria-Bertani (LB) medium (5 g/L yeast extract, 10 g/L tryptone, and 10 g/L NaCl) containing 100 μg/mL of ampicillin or low LB medium containing 25 μg/mL of Zeocin (Invitrogen, Carlsbad, CA, USA). Minimal dextrose, buffered complex glycerol mediums (BMGY), yeast extract peptone dextrose (YPD), and buffered complex methanol mediums (BMMY), were prepared according to the manufacturer's instructions. Restriction enzymes and cloning kits were obtained from TaKaRa Bio Inc. (Dalian, China). FreeStyle™HEK293-F cells, FreeStyle™293 Expression Medium, and expression vector pHXT-Fc [25] were obtained from Invitrogen (Carlsbad, CA, USA). *N*-Glycosidase purified from *Flavobacterium menigosepticum* (PNGase F) was obtained from New England Biolabs (Ipswich, MA, USA).

Construction of expression plasmids

HE4 (GenBank accession number AY212888.1) was amplified by polymerase chain reaction (PCR) using SKOV3 cell (human ovarian cancer cell) genomic DNA as template. To generate the *P. pastoris* gene construct pPICZaA-HFc [6], HE4 lacking its intrinsic signal peptide-coding sequence (282 bp) was fused to the IgG1 fragment crystallizable (Fc) fragment (690 bp) by using splicing with overlap extension PCR [6]. The PCR product, amplified using the P1/P2 and P3/P4 primer pairs, was ligated into the pSURE-T plasmid (Biomed, Beijing, China), generating pHFc-T. The HFc gene fragment was released from the pHFc-T vector by endonuclease digestion using *Xho*I/*Sac*II and was subsequently subcloned into the pPICZaA expression vector [26], yielding pPICZaA-HFc. The coding sequence was subcloned along with the open reading frame of a-factor signal that is under the control of the AOXI promoter. The HEK293-F cell expression vector pHXT-HFc was constructed by amplifying the full-length HE4 (372 bp) coding sequence by using the primers P5/P6. The primers are shown in Table. 1. The PCR product was digested with *Eco*RI/*Sal*I and then ligated into the pHXT-Fc vector directly. Both of the two expression vectors were transformed and amplified in *E. coli* Top-10 competent cells. Insert-containing clones were confirmed using clone PCR and nucleotide sequence analysis.

Expression of recombinant HE4 in P. pastoris

The recombinant plasmid pPICZaA-HFc was linearized with *Sac*I, and with some modifications, was electroporated into *P. pastoris* GS115 strain cells as previously described [27]. pPICZaA-HFc-positive transformants were screened on YPD plates containing 0.1 mg/mL Zeocin and then incubated at 30 °C until colonies appeared. Using sterile toothpicks, transformants were patched onto a nitrocellulose membrane, and the membrane was then laid on a BMGY plate to allow colonies to grow for 2 days. Protein expression was then induced for another 2 days on BMMY plates supplied with 100 μL 100% methanol every 12 h. The colonies were washed off the nitrocellulose membranes under a stream of distilled water, and the recombinant proteins secreted by the colonies that were bound to the nitrocellulose membranes were detected by immunoblotting using an anti-WFDC2 antibody (1:500; Abclonal, USA) and then a goat anti-rabbit horseradish peroxidase secondary antibody (1:5000; Sigma, USA); the protein dots were visualized by using a chemiluminescence kit (GE Healthcare Life Sciences, China) and ChemiScope 3400 mini imaging system (ClinxScience Instruments Co., Ltd., China). A single positive transformant was inoculated in flasks containing 600 mL of BMGY at 30 °C with constant shaking (250 rpm) until the optical density at 600 nm (OD_{600}) reached 6.0. The cells were harvested by centrifugation at 3000×*g* for 5 min at 4 °C and then resuspended in 400 mL BMMY to induce expression. The flasks were induced at 28 °C and methanol was added to a final concentration of 0.5% (v/v) every 12 h to sustain the induction. At different time intervals, the supernatants were collected by centrifugation (1500 rpm, 5 min, at 4 °C and analyzed directly by Western blot. A negative control using the empty pPICZaA vector was carried out in parallel. After 48 h of induction, the culture supernatants were harvested by centrifugation at 8000×*g* for 15 min at 4 °C.

Expression of recombinant HE4 in HEK293-F cells

pHXT-HFc vector was prepared using an Endotoxins Removal Plasmid DNA Purification System (Qiagen, Germany). On the day of transfection, the cells were counted (viability should be over 90%). The required volume of cells was aspirated and suspended

Table 1
Oligonucleotiedes used to amplify the three endoglucanase ORFs.

5′–3′ Primer sequences	
P1	5′-GACTGGTTCC AATTGACAAG C-3′
P2	5′-GTGTCCTGTGTCACTCCCAATTTCGACAAGACTCACACTTGTCCAC-3′
P3	5′-GTGGACAAGTGTGAGTCTTGTCGAAATTGGGAGTGACACAGGACAC-3′
P4	5′-TCCCCGCGGT CACTTACCTG GAGACAAAGA C-3′
P5	5′-GGGAATTCATGCTGCTTGTCGCCTAG-3′
P6	5′-CGGTCGACGAAATTGGGAG TGACACAG-3′