Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 326 (2005) 355–363
BBRC
www.elsevier.com/locate/ybbrc

Functional preservation of duplicated pair for RSVS III gene in the REST locus of rat 3q42☆

Han-Jia Lin[a,*], Chin-Mei Lee[b], Ching-Wei Luo[a], Yee-Hsiung Chen[a,b,*]

[a] *Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan*
[b] *Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taipei, Taiwan*

Received 7 November 2004
Available online 18 November 2004

Abstract

Among the major proteins in the rat seminal vesicle secretion, transglutaminase catalyzed the cross-links among RSVS I–III. Six peptide sequences determined from the trypsin digests of RSVS III were confirmed in the protein sequence derived from two paralogs, RSVS III_{α} and RSVS III_{β} gene, in rat 3q42. Their transcription units are organized with the first exon encoding a signal peptide, and the second a secreted protein, whereas the third encompasses a 3′-non-translated nucleotide that shares common features of rapidly evolving substrates of transglutaminase (REST) gene family. These two genes have 99% identity in their coding region and both express in adult rats with the same transglutaminase cross-linking site, manifesting functional preservation. All of REST genes reported thus far for human and Muridae were mapped in a chromosomal locus between *KCNS1* and *SLPI* suggesting the locus as an active evolving region. The molecular evolution of this gene family is discussed.

Keywords: Copulatory plug; Gene duplication; Mass spectrum; Rapidly evolving substrates of transglutaminase; Rat; Seminal vesicle

Seminal vesicle is a male mammalian accessory apparatus observed in many but not all mammalian species. Upon ejaculation, seminal vesicle secretions (SVS) stored in the central lumen squirt into urethra to make up the major part of the liquid portion of semen. Since the fertility of mice as well as rat is greatly reduced if their seminal vesicles are extirpated or removed [1,2], this accessory sexual gland has great importance in reproduction. This has prompted studies of the structure and function of protein components of SVS.

The formation of the semi-solid gelatinous mass of the human semen or the seminal plasma clotting in the rodent semen arises subsequent to the discharge of luminal fluid from the male sexual glands and is generally thought to involve transglutaminase-catalyzed cross-linking of proteins in the SVS [3]. In human, the protein substrates have been identified as human semenogelins I and II (HSg I and HSg II), two of the predominant protein components of human SVS [4]. The HSg I/II-like proteins have also been identified in the seminal vesicles of other primates such as the rhesus monkey [5], the cotton-top tamarin [6], and the marmoset [7]. In addition, guinea pig SVP-1 [8], mouse SVS II [9], and rat SVS II [10] in the rodent SVS have all been shown to be good protein substrates of transglutaminase. The genes encoding these protein substrates have similar organization and comprise a characteristic three-exon transcription unit in which the first exon encodes a short signal peptide, the second exon encodes the protein in its entirety, including a termination codon, and the third exon is not translated. Moreover, there is a clear strong similarity between their 5′-flanking regions, first exons, and

☆ This work was supported in part by Grants NSC92-2311-B-002-019 and NSC92-2311-B-001-089. The nucleotide sequence reported in this paper has been submitted to GenBank with Accession No. AF323459.
* Corresponding authors. Fax: +886 2 23635038.
E-mail addresses: hanjia@gen-info.osaka-u.ac.jp (H.-J. Lin), bc304@gate.sinica.edu.tw (Y.-H. Chen).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.11.030

3′UTRs, while the protein coding regions in their second exons have experienced rapidly evolutionary change. As a result, the protein substrates in the SVS of different animals show great differences in their primary structure. Accordingly, they are referred to as the rapidly evolving substrates for transglutaminase (REST) gene family [11].

The evolutionary steps undergone by the gene family as well as the physiological significance of their encoded proteins are still obscure. Our previous works have established the genomic structure of mouse SVS III and demonstrated it as a transglutaminase protein substrate involved in copulatory plug formation [12]. Here, we report the duplication of the Muridae SVS III gene in both the rat 3q42 and the mouse 2H3 regions, and the preservation of these paralogous pair with regard to their transglutaminase substrate activity.

Materials and methods

Animals, proteins, and chemicals. Guinea pig liver transglutaminase (EC 2.3.2.13), trypsin, monodansyl-cadaverine, and 3,3′-diaminobenzidine were obtained from Sigma, St. Louis, MO, and α-cyano-4-hydroxycinnamic acid was from Aldrich, Gillingham, England.

Outbred Wistar rats were from the Charles River Laboratories (Wilmington, MA.) and were maintained and bred in the animal center at the College of Medicine, National Taiwan University. The seminal vesicles and coagulating glands of adult rats were carefully dissected free from each other. SVS was collected by squeezing and was diluted with 50 volumes of 50 mM Tris/HCl containing 150 mM NaCl at pH 7.5.

Analytical methods. Protein concentration was determined using the BCA protein assay [13]. According to the method of Laemmli [14], SDS/PAGE was conducted. Assay of the transglutaminase-catalyzed protein cross-links [15] and trypsin digests of protein in the rat SVS followed the previous method [12]. The molecular mass and peptide sequence were determined by a Micromass Q-Tof Ultima MALDI (Waters, Milford, MA). To carry out mass spectrometric analysis, the peptide solution was mixed with an equal volume of the matrix solution (1% α-cyano-4-hydroxycinnamic acid, 50% (v/v) acetonitrile, and 0.1% trifluoroacetic acid) and allowed to air-dry on a sample target. The peptide sequences were identified from the spectral peaks by the Mascot search program (http://www.matrixscience.com/).

Rapid amplification of cDNA ends and genomic cloning. According to the partial cDNA sequence of *pSV-1*, reported by Izawa [16], 5′-rapid amplification of cDNA ends (RACE) was applied to map the transcription initiation site [17]. Total cellular RNA was prepared from the seminal vesicles of adult rats using an Ultraspec-II RNA isolation kit (Biotex, Houston, TX). The 5′-RACE cDNA Synthesis Primer and SMART II oligo (Clontech, Palo Alto, CA) were mixed with 1 μg of freshly prepared rat seminal vesicle total RNA and the first-strand cDNA synthesis was achieved by Superscript II reverse transcriptase (Gibco BRL, Gaithersburg, MD). An oligonucleotide, 5′-GGGG CCTCTTCCTGTTCTTCTCTGTGTCC-3′ which is complementary to nucleotides 167–195 in the RSVS III_α cDNA (Fig. 3A), and Universal Primer Mix were added to the 5′-RACE-ready cDNA to amplify the cDNA fragments which were sequenced to establish the transcriptional initiation site using Adventage 2 Polymerase Mix (Clontech, Palo Alto, CA). Two oligonucleotides, 5′-ACACCACTTTCC TTTGACTG-3′ and 5′-GAGCCTTCCTGACATCAT-3′, were used to amplify the 3′ ends of RSVS III_α and RSVS III_β, respectively, by using the same SMART RACE system.

According to the cDNA sequence (Fig. 3A), three oligonucleotides (GSV1-3) were synthesized. GSV1 (5′-CCCGGCTGCTTGCTTCT CCAGAAGGA-3′) and GSV3 (5′-AACACTTGTTTTCTATTCTG TGTAAC-3′) are complementary to nucleotides 57–82 and 1135–1160, respectively. GSV2 (5′-GTGGAGGCCCTTCCTGGTAAG-3′) is complementary to nucleotides 2–27 in the RSVS III_α cDNA (Fig. 3A). PCR was used to amplify DF I from an adaptor-*Dra*I library, DF II/DF III from the *Pvu*II libraries, and DF IV from the *Dra*I library of GenomeWalker kit (Clontech, Pola Alto, CA) using the adaptor primer (AP-1) and each of GSV1-3 as the primer pair (Fig. 3B). The PCR-amplified DNA fragment was then ligated into the pGEM-T-easy vector (Promega, Madison, WI) via TA cloning and the recombinant plasmid was introduced into *Escherichia coli* JM109 strain by transformation.

To identify and quantify the RNA messages of RSVS III_α and RSVS III_β. Nucleotides 420–435 in the sequence shown in Fig. 3A and another complementary to nucleotides 947–964 were prepared and used as a primer pair to perform RT-PCR with the total RNA (0.2 μg) of rat seminal vesicle. The PCR amplification program was: 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s, and cycling 15, 20 or 25 times. The PCR fragment was digested with 0.1 U/μl of *Sph*I at 37 °C for 2 h, and the digested sample was resolved by agarose gel electrophoresis. Each DNA fragment was cloned and sequenced to confirm the expression of both. For quantitative analysis, the gel bands were scanned and analyzed by the computer program, NIH image 1.62 (http://rsb.info.nih.gov/nih-image/Default.html). The intensity of 545 bp gel band was counted for the expression of *RSVS III_α*, the 372 bp plus 173 bp bands were counted for *RSVS III_β*, and reaction products from different amplification cycles were normalized to compare the relative amount of each mRNA.

Northern blotting. A RSVS III_α cDNA fragment (nucleotides 55–435) inserted into the pGEM-T-easy vector or a cDNA fragment of the rat glyceraldehyde-3-phosphated dehydrogenase (GAPDH) gene (nucleotides 1209–1966) inserted into the pGEM3 vector was used as a template to prepare ^{32}P-labeled cDNA probes using a Promega random-priming kit. The general procedures of Northern analysis were followed [18]. The total RNA samples prepared from tissue homogenates were separated on a denaturing 1.5% agarose/formaldehyde gel and then blotted onto a nylon membrane filter by capillary transfer. The membrane was hybridized with one of the two ^{32}P-labeled probes and the RNA messages on the filter were visualized by autoradiography. The probe was then stripped from the membrane and the same membrane was then hybridized with the other ^{32}P-labeled probe. Thus, hybridization with the two probes was performed on the same filter membrane.

The comparison of genomic sequences. The nucleotide sequence of RSVS III_α gene established from this work was mapped on the entire rat genome using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). The cognate REST genes and the two marker genes, *KCNS1* and *SLPI*, were identified by Unigene (http://www.ncbi.nih.gov/entrez/query.fcgi?db=unigene). The chromosome structures of human, mouse, and rat were obtained from Mapviewer (http://www.ncbi.nlm.nih.gov/mapview/). The nucleotide sequences of the two genes were matched by dot plot, using the program, Compare, which was provided by SeqWeb (http://gcg.nhri.org.tw/). The nucleotide sequences of the various genes concerned in this work were collected from EMBL or GenBank and the accession numbers are listed in parentheses for each gene: *HSg I* and *HSg II* (Z47556); rhesus *Sg II* (X92589); tamarin *Sg I* (AJ002153); guinea pig *SVP* (U59711); and mouse *semenoclotin*/*MSVS II* (XX91270); *MSVS III_α* (AF323459); *MSVS III_β* (NT039210); *MSVS IV* (NM009300); *MSVS V* (NM009301); *MSVS VI* (NM013679); *RSVS II* (J05443); *RSVS III_α* (this work); *RSVS III_β* (XM215936); *RSVS IV* (NM012662); *RSVS V* (NM133516); and *RSVS VI* (XM342575).