THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 5, Issue of February 1, pp. 3632–3639, 2002
Printed in U.S.A.

Localization of the Transglutaminase Cross-linking Site in SVS III, a Novel Glycoprotein Secreted from Mouse Seminal Vesicle*

Received for publication, August 8, 2001, and in revised form, November 19, 2001
Published, JBC Papers in Press, November 26, 2001, DOI 10.1074/jbc.M107578200

Han-Jia Lin‡§, Ching-Wei Luo¶, and Yee-Hsiung Chen‡¶‖

From the ‡Institute of Biochemical Sciences, College of Science, National Taiwan University and ¶Institute of Biological Chemistry, Academia Sinica, Taipei 106, Taiwan

The nucleotide sequence of *MpSv-1*, a novel androgen-regulated gene exclusively expressed in mouse seminal vesicle, was analyzed to establish a 5′-flanking region of 2123 bp, three exons of 95, 765, and 330 bp, and two introns of 222 and 811 bp. The transcription unit is 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 rapid evolving substrates of transglutaminase gene family. The protein sequence deduced from this gene contains 265 amino acid residues in which the central part, residues 116–145, is a region composed of five short tandem repeats, consisting of four amino acid residues, Q*X*K(S/T), where *X* is an aliphatic amino acid residue. Among the mouse seminal vesicle secretory proteins that could be resolved by SDS-PAGE into seven major components, SVS I–VII, the antiserum against residues 77–109 of the *MpSv-1*-translated protein only reacted with SVS III. Matrix-assisted laser desorption/ionization-time of flight mass spectral analysis from a trypsin digest of SVS III supported this protein as derived from *MpSv-1*. SVS III was immunolocalized to the epithelium of both the primary and secondary folds of the seminal vesicle and the copulatory plug. All of mouse SVS I–III were proven to be substrates of transglutaminase and could be cross-linked readily after the enzyme reaction. The transglutaminase cross-linking site of SVS III was identified to be the tandem repeats of Q*X*K(S/T) in the central part of this protein molecule.

Seminal vesicles are present in most male mammals. After puberty, seminal vesicle secretion (SVS)[1] accumulates in the lumen of this reproductive gland. Upon ejaculation, SVS is discharged and makes up the major portion of the seminal plasma that is the complex biological fluid formed from the mixing of various fluids in the male reproductive tract. Because rodents have proven to be good experimental animals for the molecular study of mammalian reproduction, attempts have been made to study the structure and function of mouse SVS proteins. They consist of several minor components such as SVA (1), p12 (2) etc., and seven well defined major proteins designated SVS I–VII in decreasing order of M_r = 95,000–8,000 according to their mobilities on SDS-PAGE (3). The structure and function of SVS II (4), SVS IV (3), and SVS VII (5) have been reported.

Semen is coagulated in a substantial number of mammalian species, such as many myomorphic rodents, certain moles, hedgehogs, marsupials, rabbits, stallions, boars, and several primates (6). The deposition of semen coagulum of such animals, including rodent, into the vagina at coitus results in the formation of a copulatory plug that occludes the vaginal barrel close to the uterine cervix. Thus extirpation of the seminal vesicle and coagulating gland (anterior prostate) from a mouse prevents formation of a vaginal plug and greatly reduces fertility (7). Therefore, the importance of male accessory sexual glands in the physiological utilization of semen coagulum in mammalian reproduction, although still not well defined, should not be overlooked. In this regard, it has been demonstrated that human plasma factor XIII, a secretory transglutaminase (8, 9), is able to catalyze the cross-links of both semenogelin I/II (10) and mouse SVS I/II (4). Because these transglutaminase protein substrates are exclusively expressed in seminal vesicle and have their transcription units in a similar arrangement, they are referred to as rapidly evolving substrate of transglutaminase (REST) (11, 12). Transglutaminase is secreted from some male accessory sexual glands. For instance, the RNA message of a secretory transglutaminase has been identified in the human prostate (13), and a transglutaminase has been purified from rat coagulating gland fluid (14). Apparently, the transglutaminase-catalyzed protein cross-linking, subsequent to the discharge of luminal fluid of male accessory sexual glands, may result in the formation of the semisolid gelatinous mass in human semen or seminal plasma clotting in rodent semen. In this work, we present data to support mouse SVS III as a novel REST that participates in the formation of the copulatory plug. Our results conclude that the five tandem repeats of a tetrapeptide Q*X*K(S/T) in the central part of this protein molecule act as the transglutaminase cross-linking sites.

EXPERIMENTAL PROCEDURES

Proteins and Chemicals—Guinea pig liver transglutaminase (EC 2.3.2.13), thrombin, trypsin, Schiff's reagent, monodansylcadaverine and 3,3′-diaminobenzidine were purchased from Sigma. Human plasma coagulation factor XIII was purchased from Calbiochem. Goat anti-rabbit IgG conjugated with alkaline phosphatase or horseradish perox-

* This work was supported in part by Grants NSC90–2311-B002-023 and NSC90–2311-B001-049) from the National Science Council, Taipei, Taiwan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF323459.

§ Some of the work described in this paper form parts of a dissertation submitted in partial fulfillment of the requirements of Ph.D. at the National Taiwan University.

‖ To whom correspondence should be addressed: Institute of Biochemical Sciences, College of Science, National Taiwan University, P.O. Box 23-106, Taipei 106, Taiwan. Fax: 886-2-23635038; E-mail: bc304@gate.sinica.edu.tw.

[1] The abbreviations used are: SVS, seminal vesicle secretion; REST, rapidly evolving substrate of transglutaminase; RT, reverse transcription; RACE, rapid amplification of cDNA ends; GST, glutathione *S*-transferase; PBS, phosphate-buffered saline; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight.

idase and ECL Western blotting detection reagent were purchased from Amersham Biosciences, Inc. (Buckinghamshire, United Kingdom). Nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate stock solution was purchased from Roche Diagnostics GmbH (Mannheim, Germany). The BCA protein assay reagent was purchased from Pierce, and α-cyano-4-hydroxycinnamic acid was purchased from Aldrich (Gillingham, UK).

Animals and Hormone Treatment—Outbred ICR mice 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. Animals were treated according to the institutional guidelines for the care and use of experimental animals. They were housed under controlled lighting (14 h of light, 10 h of dark) at 21–22 °C and were provided with water and NIH 31 laboratory mouse chow *ad libitum*. Animals were humanely killed by cervical dislocation.

For investigation of the androgenic effect, adult male mice (8 weeks), which had been castrated for 3 weeks, received a daily subcutaneous injection of testosterone propionate in corn oil (5 mg/kg body weight) for 8 consecutive days. The control animals received corn oil only. Seminal vesicles were removed from the animals 12 h after the last injection.

cDNA and Genomic Cloning—Total cellular RNA was prepared from the seminal vesicles of adult mice using an Ultraspec-II RNA isolation kit (Biotex, Houston, TX). A *MpSv-1* cDNA fragment (nucleotides 55~95 and 318~873 of Fig. 1) was obtained from amplification of the total RNA using the Access RT-PCR kit (Promega, Madison, WI) with a primer designed for the DNA. The polyadenylated fraction of the total RNA was isolated according to the instructions of the Oligotex mRNA mini kit (Qiagen GmbH, Hilden, Germany). Single-stranded cDNA was prepared from the polyadenylated RNAs, using a ZAP-cDNA synthesis kit (Stratagene, La Jolla, CA). Preparation of double-stranded cDNAs, construction of orientation-specific cDNA in the Uni-ZAP XR vector with *Eco*RI/*Xho*I terminus, and packaging of the constructed vectors into phage followed the manufacturer's instructions (Stratagene). The recombinant phages were transfected into the bacterial host XL1-Blue MRF′ strain. The cDNA library efficiency was 2.5×10^5 plaque-forming units/μg of cDNA.

According to the partial rat pSv-1 (*RpSv-1*) cDNA sequence reported by Izawa (15), one oligonucleotide, 5′-CCTCCTTCTGGAGAAGCAAGC-3′, which represents 1–21 nucleotides of the cDNA sequence and another oligonucleotide, 5′-AATGCTACTGTTTTACTGGAC-3′, which is complementary to nucleotides 577–597, were employed as a primer pair for RT-PCR amplification of the total RNA of mouse seminal vesicle to yield a DNA fragment (~600 bp). The DNA fragment, which was tentatively designated as the *MpSv-1* cDNA fragment, was sequenced to confirm 81% identity with the corresponding nucleotide sequence of *RpSv-1*. Plaques from the cDNA library were screened by DNA hybridization using a random-primed probe to the *MpSv-1* cDNA fragment, which was prepared using a Prime-a-gene kit (Promega, Madison, WI). Randomly chosen positives from different pools were plaque-purified. They were excised into a phagemid using the ExAssist interference-resistant helper phage and transformed into the bacterial SOLR™ strain for DNA sequencing.

Based on the *MpSv-1* cDNA sequence (Fig. 1), four oligonucleotides, GSV I–IV, were synthesized. GSV II/III represents nucleotides 30–58/59–79. GSV I is complementary to GSV II, and GSV IV is complementary to nucleotides 2195- 2223. PCR amplified DF II/DF IV from an adaptor-*Dra*I library and DF I/DF III from the *Ssp*I and *Pvu*II libraries of the GenomeWalker kit (CLONTECH, Palo Alto, CA) using the adaptor primer (AP-1) and each of GSV I–IV as the primer pair (Fig. 1). The PCR-amplified DNA fragment was then ligated into the pGEM-T-easy vector (Promega) via TA cloning, and the recombinant plasmid was introduced into *Escherichia coli* JM109 strain by transformation.

5′-Rapid Amplification of cDNA Ends (RACE) and Northern Blotting—Using 5′-RACE to map the transcription initiation site followed the report of Gong and Ge (16) by using the SMART™ RACE system (CLONTECH). The 5′-RACE cDNA synthesis primer and SMART II oligonucleotide were mixed with 500 ng of polyadenylated RNA freshly prepared from mouse seminal vesicle, and the first-strand cDNA synthesis was achieved by Superscript II reverse transcriptase (Invitrogen, Groningen, The Netherlands). An antisense oligonucleotide, which is complementary to TTGTGCAAACTAAGCACCATGCAT in the *MpSv-1* cDNA (nucleotides 435–458 of Fig. 1), and Universal Primer Mix were added to the 5′-RACE-Ready cDNA to amplify the 5′-end of *MpSv-1* cDNA using Advantage 2 polymerase mixture (CLONTECH). Finally, a nested primer pair including the nucleotide ACCAGTTCCTCTATGGGCAC in *MpSv-1* cDNA in antisense (nucleotides 378–397 of Fig. 1) and the nested universal primer were used for the nested PCR. The amplified DNA fragments were cloned and sequenced to establish the transcriptional initiation site.

A *MpSv-1* cDNA fragment inserted into the pGEM-T-easy vector and a cDNA fragment of the mouse glyceraldehyde-3-phosphate dehydrogenase gene (1233-bp) inserted into the pGEM3 vector were used as a template to prepare the ^{32}P-labeled cDNA probes using a Promega random-priming kit. The general procedures of Northern analysis were followed (17). RNA samples were separated in denaturing 1.5% agarose/formaldehyde gel and then blotted onto a nylon membrane filter by capillary transfer. The membranes were hybridized with one ^{32}P-labeled probe, and the RNA messages on the filter were visualized by autoradiography. The probe was removed from the membrane, and the same membrane was then hybridized with another ^{32}P-labeled probe. Thus, hybridization with two probes was performed on the same filter membrane.

Construction of Expression Vectors for Recombinant Polypeptides—The cDNA fragment-encoded residues 21~105, 77~109, 116~145, 147~196, or 197~237 of the *MpSv-1*-translated protein (Fig. 1) were amplified from the *MpSv-1* cDNA as a template by PCR, using a primer pair of the 15-nucleotide sequences at the 5′ end of each cDNA with a *Bam*HI site linked to the sense oligonucleotide and a *Eco*RI site linked to the antisense oligonucleotide. We synthesized the sense oligonucleotide with a *Bam*HI site and the antisense oligonucleotides with an *Eco*RI site for residues 109–115, 116–121, 122–127, 129–134, 136–141, and 142–147 and annealed each oligonucleotide pair to obtain the short DNA fragments. Each cDNA fragment was ligated to pGEX-6P-1 expression vector via *Bam*HI and *Eco*RI site, and the ligated mixture was used to transform *E. coli* strain BL21. By colony PCR (18) with an antisense oligonucleotide of each cDNA and a sense pGEX-5′-sequencing primer in pGEX-6P-1 vector, positive clones that were able to produce a chimeric polypeptide were screened and confirmed by DNA sequencing. The glutathione *S*-transferase (GST)-fused polypeptides from the cell homogenates were purified by affinity chromatography on a column of glutathione-agarose beads, and the purity of each chimeric polypeptide was identified by SDS-PAGE.

Protein Blotting and Proteolysis in the Polyacrylamide—New Zealand White rabbits were immunized with a chimeric polypeptide of GST fused to residues 77–109 of the *MpSv-1*-translated protein. Mouse SVS proteins were resolved by SDS-PAGE on a gel slab ($6.5 \times 10.5 \times 0.075$ cm) according to the method of Laemmli (19). The proteins on the gel were stained with Coomassie Brilliant Blue or transferred to a nitrocellulose membrane. After transfer, the protein blots were immunodetected by the Western blot procedure, using the antiserum as the primary antibody diluted to 1:10000 in a blocking solution (5% nonfat skimmed milk in PBS), and goat anti-rabbit IgG was conjugated with horseradish peroxidase as the secondary antibody diluted to 1:10000 in the blocking solution. The enzyme-staining bands were enhanced by chemiluminescence detection using an ECL kit according to the manufacturer's instruction.

The SVS III protein band from the polyacrylamide gel was cut out, transferred to a microcentrifuge tube, washed three times with 0.5 ml of 25 mM NH_4HCO_3, 50% acetonitrile, and dried. The protein samples in the gel were incubated with 0.2 unit of trypsin in 25 mM NH_4HCO_3 at 37 °C overnight. The supernatant was dried, redissolved in 30 μl of 0.1% formic acid, sonicated for 30 s, heated to 40 °C for 10 min, and centrifuged before mass spectral analysis.

Immunohistochemical Staining—Tissues or copulatory plugs were fixed in freshly prepared Bouin's solution (0.2% picric acid, 2% (v/v) paraformaldehyde in PBS) overnight, dehydrated in ethanol, infiltrated, and embedded in paraffin. Each section (6 μm) was mounted on a slide that had been precoated with gelatin. Sections were dried at 45 °C, deparaffined in xylene, and rehydrated through a gradient from alcohol to distilled water. The rehydrated sections were immersed in the blocking solution in a moisture chamber at 25 °C for 1 h and washed with PBS containing 0.05% (v/v) Tween 20 (PBST) four times, each for 15 min. The sections on slides were incubated with the antiserum diluted to 1:500 in the blocking solution. After the slides had been gently agitated in four changes of PBST for 15 min each, the primary antibody was immunodetected with alkaline phosphatase- or horseradish peroxidase-conjugated anti-rabbit IgG diluted to 1:1000 in the blocking solution. Slides were incubated in 10 ml of substrate solution (0.06% 3,3′-diaminobenzidine and 0.03% $NiCl_2$ in 50 mM Tris at pH 7.6) in the presence of 10 μl of 30% H_2O_2 for 15 min at room temperature to reveal the enzyme activity staining due to horseradish peroxidase. To visualize the alkaline phosphatase activity, slides were incubated in 100× dilution of nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate stock in 100 mM Tris-HCl, 100 mM NaCl, and 5 mM $MgCl_2$ at pH 9.0 for 15 min.

Analytical Methods—Seminal vesicles and coagulating glands of adult mice were carefully dissected free from each other. Secretion of

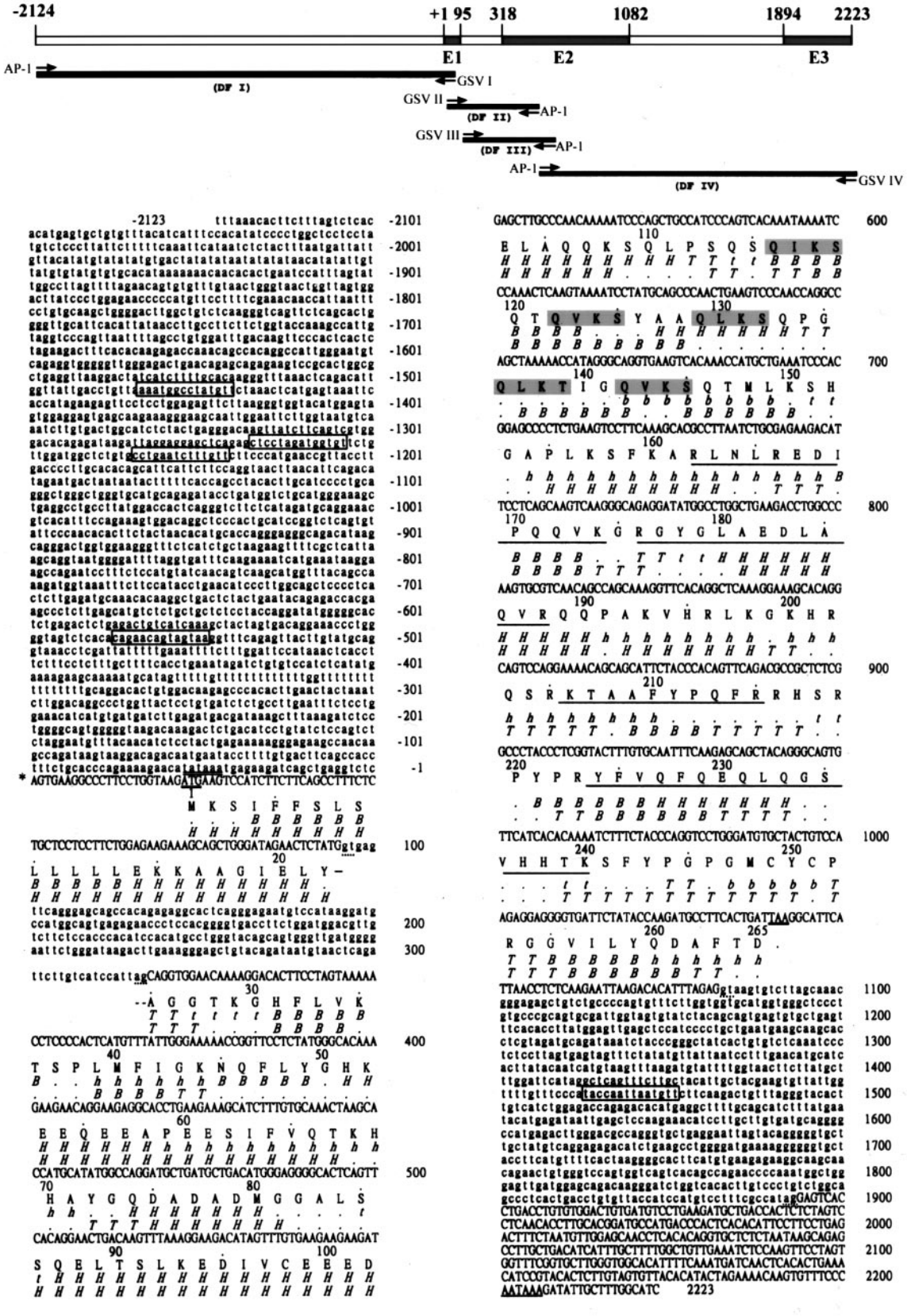

FIG. 1. **The *MpSv-1* genomic structure and the predicted secondary structure from the deduced protein sequence.** The genomic sequence, which consists of 4346 nucleotides, was constructed by aligning the nucleotide sequences of DF I–IV shown in a schematic diagram. The genomic structure includes three exons denoted by *E1–3* (*gray bars*), which were established from a 1174-bp cDNA, 5′-flanking region, and two introns (*open bars*). Nucleotide numbers are relative to the transcription initiation site marked by *. The translation start, stop codon, and the polyadenylated signal are *underlined with solid lines*. The 5′-flanking region and introns are in *lowercase letters*. Exon regions are in *bold* and *capital letters*. The consensus sequences between the exon/intron boundaries are *underlined with dashed lines*. The canonical TATA box at position −28 bp is *underlined with double solid lines*. The five ARE-like sequences, four in the promoter region and one in the second intron, are denoted in *open boxes*. The deduced protein sequences are given in one-letter code numbered from the translational start. The peptide sequences assigned

either accessory sexual gland was performed by squeezing. Coagulating gland fluid was diluted with 20 volumes of 50 mM Tris/HCl containing 150 mM NaCl at pH 7.5, and SVS was diluted with 50 volumes of the same buffer. The coagulating gland solution was centrifuged at 12,000 × *g* for 10 min. According to a modified method of Lundwall *et al.* (4), the transglutaminase-catalyzed protein cross-link was measured in a reaction buffer of 50 mM Tris, 7.5 mM Ca^{2+}, and 700 nM monodansylcadaverine at pH 7.5. To 20 μl of SVS solution, 1 μg of thrombin reaction-activated human factor XIII, 1 μg of guinea pig liver transglutaminase in 20 μl of the reaction buffer, or 20 μl of coagulating gland fluid diluted with 20 volumes of the reaction buffer was added and incubated in the presence of or without a final concentration of 50 mM EDTA at 37 °C for 1 h. The reaction solution was then mixed with an equal volume of 2× SDS-PAGE sample buffer and boiled before the electrophoresis for the identification of protein cross-links.

Protein concentrations were determined using the BCA protein assay (20) according to the manufacturer's instructions. The amino acid sequence was determined using automated Edman degradation with a gas phase sequenator (492 protein sequencer with on line 140 C analyzer, PerkinElmer Life Sciences). To carry out mass spectrometric analyses on a Voyager Elite mass spectrometer (PerSeptive Biosystems, Framingham, MA), 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% (v/v) trifluoroacetic acid) and allowed to air-dry on a sample target. The peptide sequences were searched from the spectral peaks by a MS-fit search program.

The DNA insert in the pGEM-T-easy vector or phagmid was sequenced by the dideoxynucleotide chain termination method using a primer designed for each DNA concerned. Each base was determined at least three times by an ABI PRISM 377–96 DNA sequencer using an ABI PRISM BigDye™ terminator cycle sequencing ready reaction kit (Applied Biosystems, Foster City, CA).

RESULTS

The Genomic Structure of MpSv-1—The *MpSv-1* cDNA fragment (see "Experimental Procedures") was used as a probe to screen a cDNA library of the mouse seminal vesicle. The cDNA of one clone in the constructed phagmids from randomly chosen positive clones (27 positives over 10,000 phages) was sequenced to establish a *MpSv-1* cDNA sequence of 1,174 bp that includes a 5′-untranslated region of 6 bp, an open reading frame of 798 bp, which encodes the full length of a polypeptide chain consisting of 265 amino acid residues, and a 3′-untranslated region of 370 bp, which ends with a polyadenylated region (Fig. 1). Determination of the 5′-end sequence of the mRNA by 5′-RACE analysis supported the transcription initiation site at 22 bp upstream from the ATG start codon.

Four DNA fragments, DF I–IV of the *MpSv-1* gene (Fig. 1), were amplified (see "Experimental Procedures"). The nucleotide sequences of the 2182, 490, 514, and 1704 bp determined from DF I–IV were aligned to establish the *MpSv-1* genomic sequence consisting of 4346 bp. The cDNA sequence was confirmed in the genomic sequence. Alignment of the genomic sequences with the cDNA sequence identified a 5′-flanking region up to −2123 bp from the transcription initiation site, three exons of 95, 765, and 330 bp, and two introns of 222 and 811 bp (Fig. 1). All of the exon/intron boundaries conform to the GT-AG consensus rule. A TATA box appears at 22∼28 bp upstream of the transcription initiation site, and no CAAT box element is present within about 100 bp of the 5′-flanking region. The consensus binding sequence for several transcription factors including PEA3 (nucleotides −1007 ∼ −1002, −804 ∼ −799, and −562 ∼ −557) and GATA (nucleotides −97 ∼ −92) were found. Exon 1 encodes a short peptide of 24 amino acid residues with the general features of a signal sequence that contains a 7- to 11-residue hydrophobic core preceded by one or two basic residues at the N terminus of some eukaryotic secretory proteins (Fig. 1), suggesting that the translational product of *MpSv-1* is a secretory protein. A termination codon of TAA appears in exon 2 that encodes 241 amino acid residues of the protein in its entirety. Downstream of the termination codon, including exon 3, is not translated. The deduced protein sequence was predicted to have a higher β-sheet/α-helix potential and a very low turn-structure potential in residues 1–145, whereas considerable turn structure in addition to α-helix and β-sheet was predicted in residues 146–265 (Fig. 1).

SVS III Is a Novel 34-kDa Glycoprotein Derived from MpSv-1—Because residues 77–109 in the deduced protein sequence share a very low degree of similarity with the peptide sequence of mouse seminal vesicle secretory proteins that have been reported thus far and they were estimated to have a high antigenic index by the method of Jameson-Wolf (21), GST fused to this peptide segment was used to immunize rabbits. Among the mouse SVS proteins on the gel slab of SDS-PAGE (Fig. 2, *lane 1*), the antiserum was immunoreactive only to SVS III that was identified as a 34-kDa glycoprotein (*cf. lanes 2* and *3* of Fig. 2), indicating the high specificity of the antibody against SVS III in the antiserum. The consensus Asn-Xaa-(Ser/Thr) for a *N*-linked carbohydrate (22, 23) is not present in the protein sequence, suggesting *O*-glycoconjugate in SVS III. For some unknown reason, we failed to determine the NH_2-terminal residues of SVS III by automated Edman degradation after the protein had been transferred from the polyacrylamide gel to a polyvinylidene difluoride membrane. Instead, we digested SVS III in the polyacrylamide gel with trypsin and determined the molecular mass of each trypsin digest from mass spectral analysis. The MS-fit search result confirmed the peptide fragments of residues 206–215, 177–188, 163–174, and 224–240, because they were matched with the theoretical molecular mass estimated for each peptide fragment that was predicted from trypsin digestion of the cDNA-deduced protein sequence (Fig. 1). This substantiated SVS III as derived from *MpSv-1*.

Androgen Stimulates SVS III Expression in the Luminal Epithelium of Seminal Vesicle—The tissue distribution of SVS III RNA message was examined in the reproductive tracts of adult mice, including seminal vesicle, epididymis, testis, coagulating gland, vas deferens, uterus, ovary, prostate, and vagina, and in the non-reproductive organs, including lung, kidney, brain, spleen, liver, pancreas, and heart. The mRNA was abundant in seminal vesicle but was not detectable in the other tissues (not shown). The immunohistochemical staining patterns of SVS III in a tissue slice of the adult seminal vesicles revealed that the immunoreactivity was mainly localized to the luminal epithelium of the primary and the secondary mucosa folds and that the immunochemical staining was very weak in the smooth muscle layer (Fig. 3). The strong immunochemical staining intensity of SVS III in the lumen supports the idea that SVS III accumulates in the lumen as a result of its secretion from the luminal epithelium.

The SVS III mRNA level in the seminal vesicles at different ages was compared in order to assess whether SVS III mRNA expression coincided with the development of this accessory sexual gland. A very low level of the mRNA message was detected in mice younger than 3 weeks old, appeared strongly when they were 4 weeks old, increased to its highest level when they were 6 weeks old, and remained at this higher level thereafter (Fig. 4). The positive correlation of SVS III mRNA

from a MALDI-TOF mass spectrum of the trypsin-digested SVS III are *underlined*. Five tandem repeats of the short peptide sequence Q*X*K(S/T) are denoted in *gray boxes*. The secondary structures listed in the first and the second line below the protein sequence were predicted from the Chou-Fasman algorithm (33, 34) and the Garnier-Osguthorpe-Robson algorithm (35), respectively. *B* or *b*, strong or weak β form former; *T* or *t*, strong or weak β turn former; *H* or *h*, strong or weak helix former.

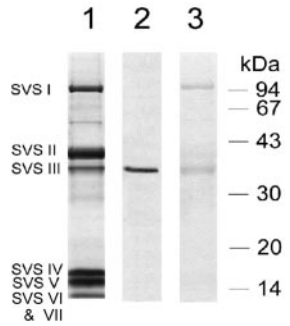

FIG. 2. **Demonstration of the *MpSv-1*-translated protein in mouse SVS.** Mouse SVS proteins (15 μg) were subjected to SDS-PAGE on a 14% gel slab. The gel was stained with Coomassie Brilliant Blue to reveal the protein bands denoted SVS I–VII (*lane 1*) or was reacted with periodic acid-Schiff reagent for the detection of glycoprotein bands (*lane 3*). Meanwhile, the proteins in the gel were transferred to a polyvinylidene difluoride filter and immunodetected by Western blot procedure using rabbit antiserum against residues 77–109 in the *MpSv-1* cDNA-deduced protein (*lane 2*).

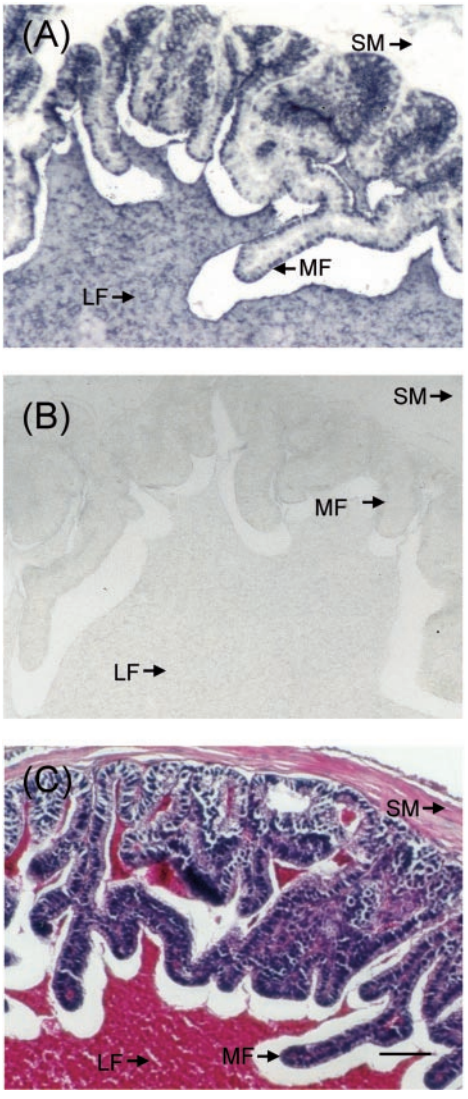

FIG. 3. **Immunolocalization of SVS III in sections of seminal vesicles of adult mice.** Cross-sections of seminal vesicle were immunoreacted with the SVS III antibody (*A*) or with normal serum (*B*) and stained further with horseradish-peroxidase-conjugated anti-rabbit IgG as secondary antibodies (see "Experimental Procedures"). The specimens were stained with hematoxylin and eosin for contrast (*C*). Photos were taken with bright field illumination. *MF*, mucosa fold; *SM*, smooth muscle; *LF*, luminal fluid. The *scale bar* represents 100 μm. The *blue deposit* represents the presence of SVS III as it appears in *MF* and *LF*.

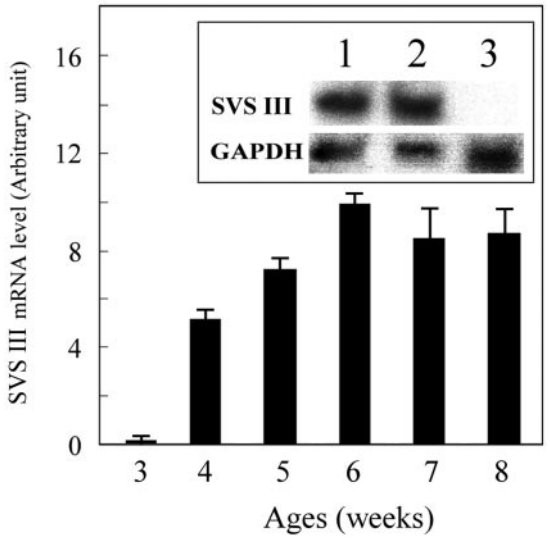

FIG. 4. **Androgen dependence of SVS III mRNA expression in mouse seminal vesicle.** SVS III mRNA in the total RNA (15 μg) prepared from seminal vesicle homogenates was detected by a Northern blot procedure. SVS III mRNA and glyceraldehyde-3-phosphate dehydrogenase mRNA in the tissue are displayed in the *inset*: *lane 1*, normal adult; *lane 2*, adult castrated 3 weeks previously and receiving testosterone propionate in corn oil for 8 consecutive days; *lane 3*, adult castrated 3 weeks previously and receiving corn oil for 8 consecutive days. In the development of animals aged 3–8 weeks, the relative amount of SVS III mRNA in the tissue was determined by densitometer scanning of the autoradiogram and adjusted with respect to the glyceraldehyde-3-phosphate dehydrogenase mRNA level. Data represent the means of three experiments, and error bars represent S.D.

expression with the development of this accessory sexual gland prompted the assessment of the androgenic effect on the gene expression *in vivo*. As shown in the Northern blot analysis (*inset* of Fig. 4), the SVS III mRNA disappeared in the tissues of adult mice that had been castrated 3 weeks earlier and received a daily injection of corn oil only. The RNA message reappeared in the tissues of castrated mice after they had received daily injections of testosterone propionate in corn oil for 8 consecutive days (subcutaneously, 5 μg/g body weight per day). As compared with the SVS III mRNA level in seminal vesicles of normal adults, the absence of SVS III mRNA in the control castrates and its recovery soon after the hormonal administration revealed the androgenic stimulation of gene expression in seminal vesicle. According to the consensus palindrome of ARE, AGAACA*nnn*TGTTCT (24), five ARE-like sequences with 50–67% identity to the consensus ARE were identified in the *MpSv-1* genomic sequence. Four of these potential AREs are found spread over the 5′-flanking region and one in the second intron (Fig. 1).

SVS III in Copulatory Plug and Its Transglutaminase Cross-linking Site—Following the standard procedures for the immunochemical staining, 6-μm cross-sections of a copulatory plug were treated with the antiserum against SVS III and alkaline phosphatase-conjugated goat anti-rabbit IgG. The SVS III immunoreactivity reflected by the intensity of enzyme activity staining was very high in contrast to the control specimen treated with normal serum (Fig. 5). Apparently, SVS III is one of the seminal vesicle secretory proteins involved in the formation of the copulatory plug.

Transglutaminase is a Ca^{2+}-dependent enzyme that catalyzes the formation of an isopeptide bond between its protein substrate through ϵ-(γ-glutamyl)lysine cross-bridges (8). To determine what protein components in SVS serve as the substrate of this enzyme, SVS was incubated with an extract of mouse coagulating gland, human plasma factor XIII, or guinea

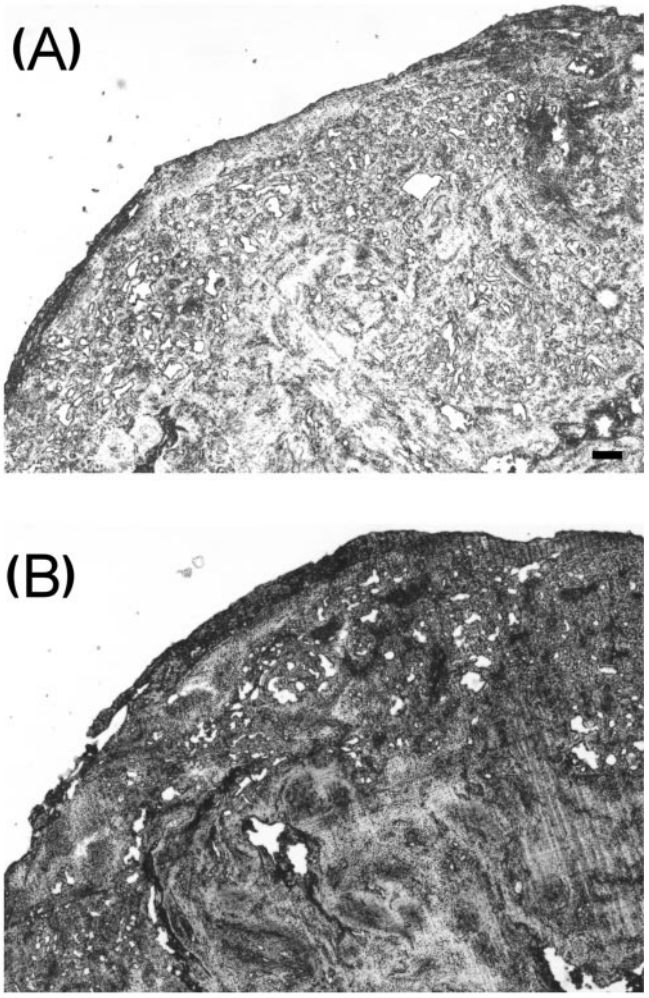

FIG. 5. **Demonstration of SVS III in the mouse copulatory plug.** Cross-sections of the copulatory plug were incubated with normal serum (*A*) or antiserum against SVS III, and the antibodies were immunodetected with alkaline phosphatase-conjugated anti-rabbit IgG (see "Experimental Procedures"). The *dark blue deposit* after the enzyme activity staining indicates the presence of SVS III. The *scale bar* represents 100 μm.

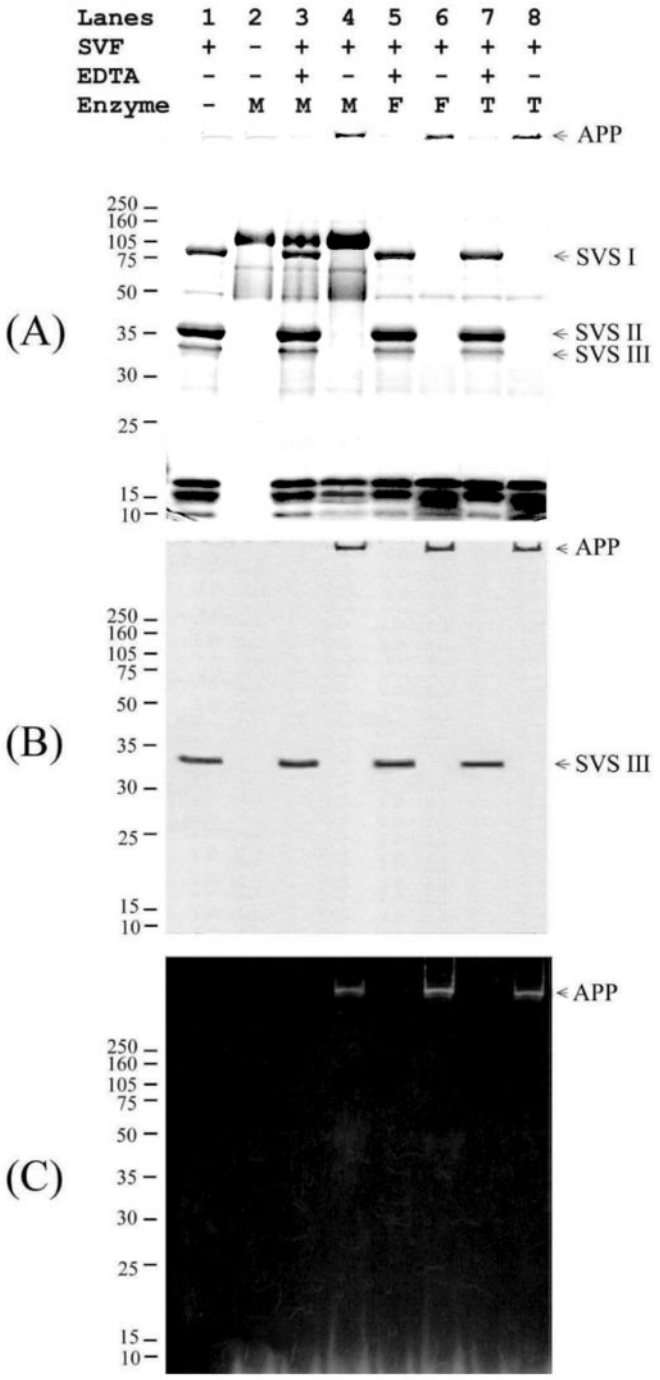

FIG. 6. **Involvement of SVS III in the transglutaminase-catalyzed protein cross-link among mouse SVS proteins.** Thrombin-activated factor XIII (*F*), guinea pig liver transglutaminase (*T*), or mouse coagulating gland fluid (*M*) was incubated with mouse SVS proteins (15 μg) at 37 °C for 1 h, and the reaction mixture was resolved by SDS-PAGE on a 14% gel slab (see "Experimental Procedures"). The gel was stained with Coomassie Brilliant Blue (*A*), immunodetected by Western blot analysis using the rabbit antiserum against SVS III (*B*), or observed under ultraviolet light to show the fluorescence arising from monodansylcadaverine (*C*). *APP* is the site of sample application.

pig liver transglutaminase in the presence of monodansylcadaverine in a reaction buffer. Monodansylcadaverine was used as a fluorescent substrate of transglutaminase (25). The protein components after incubation were identified by SDS-PAGE (Fig. 6). Only a major protein band over 105 kDa and a minor 55-kDa band were identified in the coagulating gland extract (Fig. 6*A*, *lane 2*). This caused no confusion with the identification of mouse SVS protein (*cf. lanes 1 and 2* of Fig. 6*A*). All of the SVS I–III were cross-linked by the enzymes from different origins to form high molecular weight complexes, which incorporated monodansylcadaverine and appeared at the application site of polyacrylamide gel (*cf. lanes 4*, *6*, and *8* in Fig. 6, *A* and *C*). SVS I~III were not cross-linked when the enzyme activity was inhibited by the addition of EDTA to the reaction buffer (*cf. lanes 3*, *5*, and *7* in Fig. 6, *A* and *C*). SVS III in the cross-linked complex or its free form in the incubation was also confirmed by Western blot analysis (Fig. 6*B*). For the time being, we were unable to explain whether each of SVS I–III was cross-linked itself or they were cross-linked to each other in the formation of high molecular weight complexes.

We measured the transglutaminase substrate activity of recombinant polypeptides including GST fused to residues 21–105 (F1), 77–109 (F2), 109–115 (F3), 116–121 (F4), 116–145 (F5), 147–196 (F6), and 197–236 (F7) of SVS III (Fig. 7). GST, F1, F2, F5, F6, and F7 were respectively identified as a 28.4-, 37.8-, 33.8-, 31.5-, 34.1-, and 33.3-kDa band by SDS-PAGE. The enzyme did not cross-link either GST or F1 (*lanes 1* and *2* of Fig. 7), whereas the enzyme-catalyzed cross-links of F5 were very striking (*cf. lanes 7* and *8* of Fig. 7). There remained almost no free form of F5 after the enzyme reaction. Based on

the relation between molecular size and protein mobility on SDS-PAGE, dimers, trimers, tetramers, pentamers, and hexamers of F5 were clearly identified. Homopolymers even larger than hexamers were also detected. Apparently, F5 were intermolecularly cross-linked by the enzyme reaction. Unlike F5, almost all of F2, F6, or F7 remained in the free form, and only trace amounts of dimers/trimers were formed from each of these three chimeric polypeptides after the enzyme treatment, indicating that ^{106}QQKS^{109} in F2, ^{171}QQVKG^{175} in F6, a glutamine-rich region of ^{227}QFQEQLQG^{235}, and a lysine-rich region of ^{198}KGKHRQSRK^{206} in F7 are not good transglutaminase substrates. F3 was not cross-linked by the enzyme, indicating that ^{109}SQLPSQS^{115} is not a transglutaminase substrate either (*lane 4* of Fig. 7). F4 was cross-linked to dimer by the enzyme reaction (*cf. lanes 5* and *6* of Fig. 7), manifesting the transglutaminase substrate activity of the short peptide ^{116}QIKSQT^{121}. Likewise, the enzyme was able to cross-link a chimeric peptide of GST fused to ^{122}QVKSYA^{127}, ^{129}QLKSQP^{134}, ^{136}QLKTIG^{141}, or ^{142}QVKSQT^{147} (data not shown). Apparently, one segment of Q*X*K(S/T) is sufficient for the transglutaminase-catalyzed protein cross-linking. These data clearly demonstrated that the five segments of peptide sequence Q*X*K(S/T) in residues 116–145 act as the transglutaminase cross-linking sites.

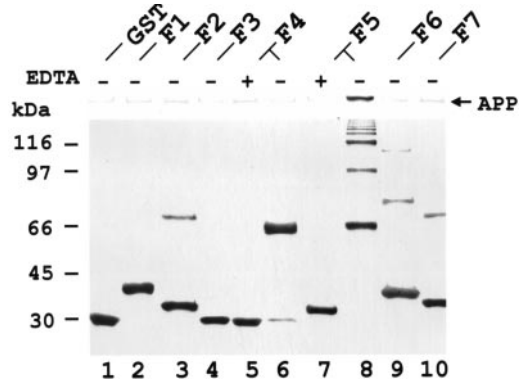

FIG. 7. **Identification of the transglutaminase-catalyzed site in SVS III.** The transglutaminase protein substrate activity was tested for a chimeric polypeptide of GST fused to residues 21–105 (F1), 77–109 (F2), 109–115 (F3), 116–121 (F4), 116–145 (F5), 147–196 (F6), or 197–237 (F7) of SVS III. Each chimeric polypeptide (2.5 μg) was incubated with guinea pig liver transglutaminase (0.1 μg) in the same experimental condition described in Fig. 6. The reaction mixture was analyzed by SDS-PAGE on a 10% gel slab, and the gel was stained with Coomassie Brilliant Blue to reveal the protein bands. *APP* indicates the sample application site.

DISCUSSION

REST genes have their transcription unit split into three exons, and most of the coding nucleotides are present in their second exons. The nucleotide sequences are highly conserved in their first exons, first introns, second introns, and third exons, but their second exons may undergo rapid evolution, which give rise to proteins that are not similar in their primary structure (11). A pairwise comparison between *MpSv-1* and the human semenogelin I gene (26) or *MpSv-1* and the mouse SVS II gene (12) showed nucleotide sequence conservation in their first exons, first introns, second introns, and third exons and a significant difference in the nucleotide sequences of their second exons (Fig. 8). Such a characteristic arrangement of the transcript unit of the *MpSv-1* gene together with the demonstration of its translated protein as a transglutaminase substrate suggests it as a new member of the REST gene family.

The protein sequence deduced from *MpSv-1* cDNA (Fig. 1) is not identically aligned with the primary structures reported for mouse SVS proteins such as SVA (27), SVS II (4), SVS IV (3), SVS VII (5), and P12 (28). Our results support SVS III, which differs from SVS I and SVS II, as a novel 34-kDa glycoprotein derived from *MpSv-1*. Lundwall *et al.* (4) have demonstrated transglutaminase-catalyzed protein cross-linking of mouse SVS I and II. In this work, we have demonstrated that, in addition to these two SVS proteins, SVS III is also cross-linked by transglutaminase from different origins. This agrees with the strong immunoreactivity of SVS III detected in a cross-section of a copulatory plug (Fig. 5), which manifests an important role for SVS III as a new semenoclotin involved in the protein cross-linking in the copulatory plug.

The transcription of *MpSv-1* is exclusive to the seminal vesicle and SVS III is predominantly secreted from the luminal epithelium of mucosa folds in this accessory sexual gland into the lumen (Fig. 3). The castration of male adults results in the disappearance of the RNA message in seminal vesicle, and the gene expression is restored in the accessory sexual gland after administration of testosterone to the castrated males. This is consistent with the lack of *MpSv-1* mRNA in the seminal vesicle before puberty, when serum androgen is at a low level, and represents a positive correlation of androgen-stimulated *MpSv-1* expression with the animal maturation. Since castration causes considerable involution of the seminal vesicle, the

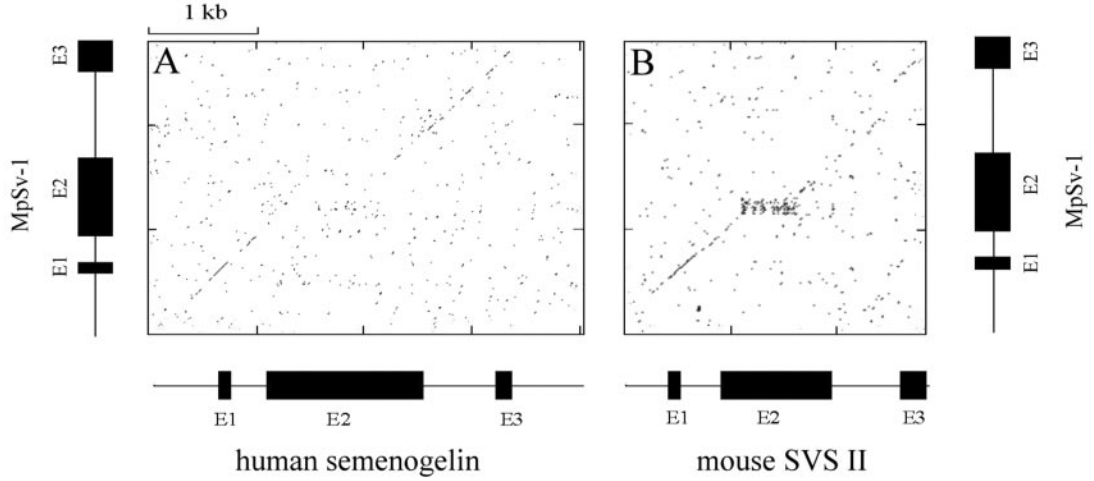

FIG. 8. **Similar arrangement of the transcription units in *MpSv-1*, MSVS II, and human semenogelin I genes.** The location of exon (*E*) of each gene is illustrated on the axis. The nucleotide sequences of two genes were matched using the program, Compare, provided by SeqWeb (gcg.nhri.org.tw/). One *dot* represents an identical sequence along the polynucleotides of two genes. A *continuation of dots* along a straight line gives a stretch of nucleotide sequence similarity in the two genes.

increase in *MpSv-1* mRNA after testosterone administration to castrates may result from either androgen-induced cell proliferation or from androgen-dependent gene expression. Whether the ARE-like sequences in *MpSv-1* are functional awaits future study.

Until now, the protein substrate of transglutaminase in mouse SVS, namely SVS I–III, have not been purified. This hinders the characterization of transglutaminase-activated sites in the protein molecules by direct protein analysis. In this work, we tackled the problem by an alternative approach. As is well known, transglutaminase catalyzes the protein cross-links through an acyl transfer reaction in which the γ-carboxamide group of a glutamine residue on one protein molecule serves as the primary substrate of an acyl donor, and the ϵ-amino group of a lysine residue on the other protein molecule acts as an acceptor substrate. Previous studies indicate that a hydrophobic residue, adjacent to and on either the carboxyl side of the glutamine residue or the amino side of the lysine residue in a polypeptide, enhances significantly their substrate activity (29, 30). The opposite effect prevails when a hydrophobic residue is on the amino side of the glutamine residue or the carboxyl side of the lysine residue. Mouse SVS III has a high content of glutamines and lysines, which together amount to 20% of all amino acid residues of the protein. There are 29 glutamines and 24 lysines in one SVS III molecule. The 7 glutamines and 5 lysines in the central part, residues 116–147, of this protein are in a cluster of five tandem repeats of the short peptide Q*X*K(S/T) that is characteristic of a transglutaminase substrate acyl donor as well as an acyl acceptor. No such peptide sequence is present in either the NH_2-terminal region (residues 1–115) that has 9 glutamine residues and 10 lysine residues or the COOH-terminal region (residues 146–265) that contains 13 glutamine residues and 9 lysine residues. Using the recombinant polypeptide of GST fused to different parts of the SVS III molecule, we have demonstrated that neither the NH_2-terminal nor the COOH-terminal region but the central region of SVS III is the transglutaminase cross-linking site (Fig. 7). The lack of a predicted turn structure in the central region (Fig. 1) makes the five tandem repeats unlikely to juxtapose in an antiparallel arrangement. Rather the secondary structure of this region may be extended to expose the transglutaminase substrate residues for the glutamine-lysine cross-linking.

The characteristic arrangement of the transcription unit in the REST gene has also been found in the rat SVS II gene (31) and guinea pig SVP genes (32), but no demonstration about the cross-linking of these proteins by transglutaminase has yet been reported, despite suggestions that they are transglutaminase substrates. Like mouse SVS III, a high content of glutamine and lysine residues makes up 22, 27, 30, and 20% of total amino acid residues in human semenogelin I, mouse SVS II, rat SVS II, and guinea pig SVP-1, respectively. The peptide sequence Q*X*K(S/T) is tandem-repeated in mouse SVS II and rat SVS II but is absent in human semenogelin and guinea pig SVP. Possibly, the REST genes may evolve to species-specific transglutaminase-activated sites. Mouse SVS II contains 375 amino acid residues with 20 segments of Q*X*K(S/T) in one region, which comprises tandem repeats of the sequence QVKSSGS that extends from residue 82 to approximately residue 250 (4). Such a dense Q*X*K(S/T) in the central parts of mouse SVS II and SVS III is characteristic in the nucleotide sequence similarity shown in a rectangular box on the dot plot of these two genes (Fig. 8). The same situation occurs in the rat SVS II, which is made up of 414 amino acid residues in which over 50% of the protein exhibits a 13-residue repeat with the consensus sequence QSQLKSFGQVKSS that spans residues 86–298 of the protein molecule. Mouse SVS II and rat SVS II were proposed to have their internal repetitive structure in an antiparallel arrangement by spatial juxtaposition of the transglutaminase substrate residues for glutamine-lysine cross-linking (31). It raises a question of whether the conformation formed by such an internal structure is necessary for a transglutaminase substrate. This necessity may be ruled out in light of our demonstration that one segment of Q*X*K(S/T) as a basic unit is sufficient for the transglutaminase-catalyzed cross-linking. The internal antiparallel structure in mouse SVS II and rat SVS II is unlikely in the central part of the SVS III molecule, suggesting that SVS II and SVS III may have different roles in copulatory plug formation.

Acknowledgments—We thank Thai-Yen Ling, Ph.D., Institute of Biomedical Sciences in Academia Sinica, for the help with the MALDI-TOF mass spectrometer analysis.

REFERENCES

1. Huang, Y. H., Chu, S. T., and Chen, Y. H. (2000) *Biol. Reprod.* **63,** 1562–1566
2. Chen, L. Y., Lin, Y. H., Lai, M. L., and Chen, Y. H. (1998) *Biol. Reprod.* **59,** 1498–1505
3. Chen, Y. H., Pentecost, B. T., McLachlan, J. A., and Teng, C. T. (1987) *Mol. Endocrinol.* **1,** 707–716
4. Lundwall, A., Peter, A., Lovgren, J., Lilja, H., and Malm, J. (1997) *Eur. J. Biochem.* **249,** 39–44
5. Luo, C. W., Lin, H. J., and Chen, Y. H. (2001) *J. Biol. Chem.* **276,** 6913–6921
6. Williams-Ashman, H. G. (1984) *Mol. Cell Biochem.* **58,** 51–61
7. Peitz, B., and Olds-Clarke, P. (1986) *Biol. Reprod.* **35,** 608–617
8. Folk, J. E. (1980) *Annu. Rev. Biochem.* **49,** 517–531
9. Lorand, L., and Conrad, S. M. (1984) *Mol. Cell. Biochem.* **58,** 9–35
10. Peter, A., Lilja, H., Lundwall, A., and Malm, J. (1998) *Eur. J. Biochem.* **252,** 216–221
11. Lundwall, A., and Lazure, C. (1995) *FEBS Lett.* **374,** 53–56
12. Lundwall, A. (1996) *Eur. J. Biochem.* **235,** 424–430
13. Dubbink, H. J., Verkaik, N. S., Faber, P. W., Trapman, J., Schröder, F. H., and Romijn, J. C. (1996) *Biochem. J.* **315,** 901–908
14. Esposito, C., Pucci, P., Amoresano, A., Marino, G., Cozzolino, A., and Porta, R. (1996) *J. Biol. Chem.* **271,** 27416–27423
15. Izawa, M. (1990) *Endocrinol. Jpn.* **37,** 233–238
16. Gong, B., and Ge, R. (2000) *BioTechniques* **28,** 846–848, 850, 852
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual,* 2nd Ed., pp. 7.19–7.50 and 9.58, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
18. Güssow, D., and Clackson, T. (1989) *Nucleic Acids Res.* **17,** 4000–4001
19. Laemmli, U. K. (1970) *Nature* **227,** 680–685
20. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) *Anal. Biochem.* **150,** 76–85
21. Jameson, B. A., and Wolf, H. (1988) *Comput. Appl. Biosci.* **4,** 181–186
22. Marshall, R. D. (1972) *Annu. Rev. Biochem.* **41,** 673–702
23. Gavel, Y., and von Heijne, G. (1990) *Protein Eng.* **3,** 433–442
24. Roche, P. J., Hoare, S. A., and Parker, M. G. (1992) *Mol. Endocrinol.* **6,** 2229–2235
25. Velasco, P. T., Karush, F., and Lorand, L. (1988) *Biochem. Biophys. Res. Commun.* **152,** 505–511
26. Ulvsback, M., Lazure, C., Lilja, H., Spurr, N. K., Rao, V. V., Loffler, C., Hansmann, I., and Lundwall, A. (1992) *J. Biol. Chem.* **267,** 18080–18084
27. Yu, L. C., Chen, J. L., Tsai, W. B., and Chen, Y. H. (1993) *Biochem. J.* **296,** 571–576
28. Lai, M. L., Chen, S. W., and Chen, Y. H. (1991) *Arch. Biochem. Biophys.* **290,** 265–271
29. Gorman, J. J., and Folk, J. E. (1980) *J. Biol. Chem.* **255,** 419–427
30. Schrode, J., and Folk, J. E. (1979) *J. Biol. Chem.* **254,** 653–661
31. Harris, S. E., Harris, M. A., Johnson, C. M., Bean, M. F., Dodd, J. G., Matusik, R. J., Carr, S. A., and Crabb, J. W. (1990) *J. Biol. Chem.* **265,** 9896–9903
32. Hagstrom, J. E., Fautsch, M. P., Perdok, M., Vrabel, A., and Wieben, E. D. (1996) *J. Biol. Chem.* **271,** 21114–21119
33. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry* **13,** 222–245
34. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276
35. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol.* **120,** 97–120