

26712

Isolation and characterization of the gene expressing the major salivary gland protein of the female mosquito, *Aedes aegypti*

Anthony A. James, Karen Blackmer^{*1}, Osvaldo Marinotti^{*2}, Corine R. Ghosn, and Jeffrey V. Racioppi^{*3}

Department of Molecular Biology and Biochemistry, University of California, Irvine, CA, U.S.A.

(Received 13 July 1990; accepted 27 August 1990)

We have undertaken a molecular analysis of the salivary glands of hematophagous insects in order to better understand their role in blood feeding and in the transmission of infectious diseases. To that end, genomic and cDNA clones of a gene designated D7, expressed abundantly in the adult female salivary glands of the vector mosquito *Aedes aegypti*, have been isolated and characterized. This gene encodes a mRNA shown by Northern analysis and in situ hybridization to tissue sections to be specifically transcribed in the distal-lateral and medial lobes of the glands, regions that are highly differentiated in females. The deduced gene product, a protein of approximately 37 kDa appears to be novel. Polyclonal antibodies made to a recombinant D7 product recognize a protein with the proper molecular weight in female salivary glands and saliva. These studies indicate that the D7 gene probably encodes the major secreted protein synthesized in the female salivary glands. The stage- and sex-limited expression of the D7 gene, and the secretion of its product, indicate that the product is most likely involved with the blood feeding capabilities of the female mosquito.

Key words: Salivary gland; Sex-specific gene; Vector biology

Introduction

The possibility of using genetically engineered vector insects in disease control programs has prompted studies of insect organs and tissues where significant interactions with pathogens take

place. One objective is to use promoters of genes active in those specific tissues to control the expression of exogenous gene products selected to interfere with pathogen development. Our efforts are focused on the salivary glands of adult female mosquitoes because prior to transmission most pathogens must reside for some period of time in these organs. In addition, adult female mosquitoes are highly adapted ectoparasites of vertebrates, and as such, their salivary glands produce and secrete proteins that enable them to blood feed efficiently [1].

We began our analysis by isolating and characterizing the major genes expressed specifically in the salivary glands. These genes are principally those that encode the secreted salivary proteins. Our characterization of each gene includes the determination of the primary structure, and its expression pattern, and where possible, the function of the gene product. Here we report the molecular and developmental properties of the major salivary gland-specific gene expressed abundantly and preferentially in *Aedes aegypti* females. The

Correspondence address: A.A. James, Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92717, U.S.A.

*Present addresses: ¹Tufts Medical School, Tufts University, Boston MA, 02118, U.S.A.; ²Departamento de Parasitologia, ICB-Universidade de São Paulo C.P. 4365, 05580, São Paulo, S.P. Brazil; ³Division of Science, College of Basic Studies, Boston University, 871 Commonwealth Ave., Boston, MA 02215, U.S.A.

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBankTM database with the accession numbers M33156 and M33157.

Abbreviations: AMV, avian myeloblastosis virus; BSA, bovine serum albumin; cDNA, complementary DNA; poly(A)⁺, polyadenylated; rRNA, ribosomal RNA; SDS, sodium dodecyl sulfate; SDS-PAGE, SDS-polyacrylamide electrophoresis; SSC, saline sodium citrate; TBS, Tris-buffered saline.

abundance of the D7 gene product and its sex- and stage-specificity suggest a role in blood feeding and/or digestion.

Materials and Methods

Mosquito strains. The mosquitoes used in this study are the Rockefeller strain of *A. aegypti*. Mosquitoes were reared using standard protocols [2] and the insects used in the following procedures were usually 3–5 days old after adult emergence.

Nucleic acid isolation and labeling protocols. RNA was prepared from whole adults and isolated tissues using a protocol from Chirgwin et al. [3]. Polyadenylated (poly(A)⁺) RNA was prepared following the protocol of Maniatis et al. [4].

DNA was prepared from adult mosquitoes using the protocol of Bender et al. [5]. Filter hybridization analyses of genomic and cloned DNAs were done according to Southern [6] using the high-stringency buffer from Maniatis et al. [4]. DNA was labeled using a random primer kit supplied by Boehringer Mannheim, or by a standard nick translation reaction [7]. RNA was end-labeled with [³²P] gamma-labeled ATP [8]. Riboprobes were synthesized from coding sequences cloned into the plasmid, pGEM1, using ³H- and ³⁵S-substituted ribonucleotides in reaction conditions supplied by the manufacturer (Promega Biotec).

Developmental and spatial analyses of RNA expression. Stage-, tissue- and sex-specific RNAs were examined by Northern analysis [9]. Exposures of Northern blots to X-ray film were typically overnight. Hybridizations in situ to tissue sections were performed as described [10]. Exposures of sections of female glands were from 1–3 days, while male glands were exposed for 45 days.

Sequence determination and analysis. Subcloned genomic and cDNA fragments were sequenced in both directions (Fig. 1) by the method of Sanger et al. [11] using a kit supplied by USBiochemicals. Oligonucleotide primers were either obtained commercially or synthesized with a Biossearch DNA synthesizer. RNA sequencing experiments were essentially identical to the DNA

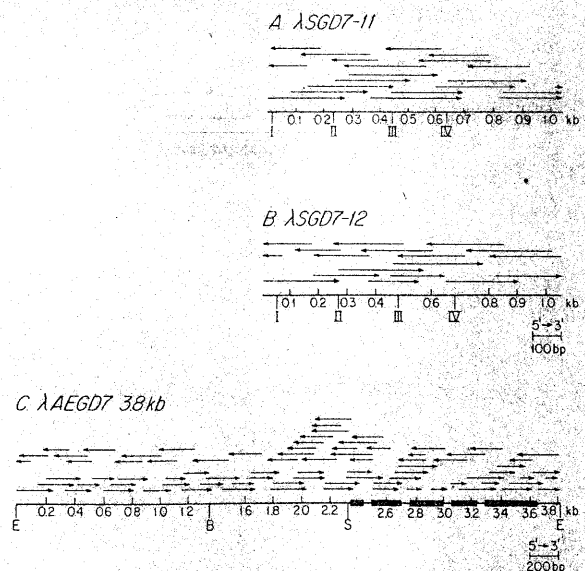


Fig. 1. Molecular maps of two cDNAs and a genomic clone homologous to the D7 gene. The thin horizontal lines in (A), (B) and (C) represent the sequenced DNA. The thick horizontal line in C represents the coding region of the D7 gene. Intron locations are marked by the Roman numerals I–IV in (A) and (B), and by the spaces in the thick horizontal line in (C). The arrows indicate the length and direction of the different sequencing gel readings. In (C), various restriction endonuclease sites are indicated; B, *Bam*HI; E, *Eco*RI; S, *Sma*I.

sequencing experiments, except that 3.25 μ g total salivary gland RNA was used as the template and 5 units of AMV reverse transcriptase as the polymerase. The primer extension protocols are described in James et al. [10] and 0.9 μ g total salivary gland RNA was used per reaction. Two 40-mer primers were used: D7R12, 5'-GGTTCGTTAATCCATTCTTTCCATTTCGCTACATTGGAA-3'; and RD7GEN12, 5'-CGGGATGAATGTTAGACAGGCCTAATGCGTTTTATATGCA-3'. D7R12 is complementary to nucleotides +223 through +262, and RD7GEN12 to nucleotides –33 through +7 of the genomic sequence (Fig. 4). Samples were run on 6% polyacrylamide sequencing gels for visualization.

Computer analysis of sequence information was performed using the Amersham Staden software package. An analysis of similarity was performed using the Lipman-Pearson rapid search algorithms and RDF analysis [12], resident in the Georgetown University PIR database programs.

Preparation of antibodies and immunodetection protocols. The full length λ SGD7-12 cDNA was cloned into the pT7 expression plasmid [13] generating pT7-D7. Recombinant D7 protein synthesis was induced by heat shock of transformed *Escherichia coli* containing both pT7-D7 and pGP1-2.

The recombinant D7 protein was purified by SDS-PAGE [14] and injected subcutaneously into a rabbit. A booster injection was given 40 days after the first immunization, and the rabbit bled 7 days later. The serum obtained was kept at -20° until used for immunoblotting procedures.

Material for immunoblotting was collected by dissecting salivary glands of 3–5-day-old males and females, or by collecting saliva as described [15]. Saliva and glands were stored at -20° in TBS (10 mM Tris, pH 7.5/150 mM NaCl). An equal volume of loading buffer containing SDS and 2-mercaptoethanol was added to each sample and the samples resolved by SDS-PAGE on a 10% gel. Proteins were electrophoretically transferred to nitrocellulose using a custom apparatus. The resulting filter was blocked for 3 h in 3% BSA in TBS. The membrane was then incubated overnight in a 1/500 dilution of rabbit polyclonal anti-D7 serum in 3% BSA/TBS. After washing in 0.05% Triton X-100/TBS, the filter was incubated in the presence of ^{125} I-labeled protein G (10^5 cpm ml^{-1}) for 1 h. The filter was washed as before, air-dried, and exposed to Kodak AR film with a Dupont Cronex intensifying screen at -70°C .

Results

Isolation of clones homologous to the D7 gene. A differential screening of an adult female cDNA library with end-labeled salivary gland RNA resulted in the isolation of clones homologous to RNAs expressed specifically in the adult salivary gland [10]. Six of these clones were apparently identical and were designated D7 after a representative clone, λ SGD7-1. Approximately 150 bp were sequenced at both ends of λ SGD7-1 using universal primers (New England Biolabs) and double-stranded template sequencing techniques. Using the information from these sequences, we synthesized two oligonucleotides and used them to select other clones from a salivary gland cDNA

library and genomic clones from an *A. aegypti* genomic library. Fig. 1 shows a molecular map of a genomic clone, λ AEGD7 that includes the D7 gene.

Temporal and spatial expression of the D7 gene. Northern analysis using ^{32}P -labeled λ SGD7-1 and total RNA from different developmental stages and isolated tissues demonstrates that the D7 cDNA is homologous to an approx. 1.2-kb polyadenylated RNA present in adult female salivary glands (Fig. 2). No expression is seen in embryos, larvae or pupae, and little is observed in adult males. The RNA can be detected up to 48 h following blood feeding (O. Marinotti and A.A. James, unpublished). The intensity of the signal in the female salivary gland lane after an exposure of 18 h indicates that this is an abundantly expressed gene.

Hybridizations in situ to sections of adult salivary glands were used to reveal those cells containing D7 transcripts. Antisense RNA was synthesized from a 461-bp, 5'-end *Eco*RI-*Hinc*II frag-

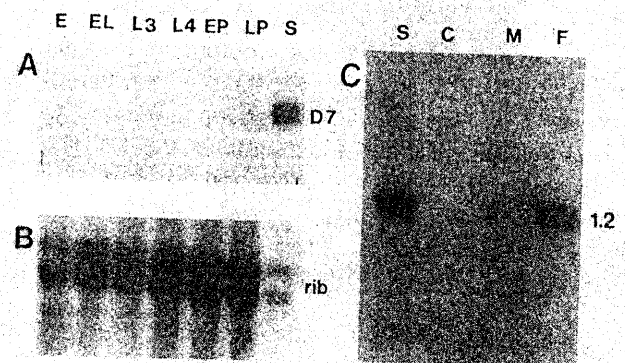


Fig. 2. Developmental expression of the D7 gene. RNA was prepared from various stages of development and different adult tissues, resolved on 0.9% agarose-formaldehyde denaturing gels, and transferred to nitrocellulose. The resulting Northern blot was probed with ^{32}P -labeled D7 cDNA (A and C). After exposure, the probe was boiled off and the filter was reprobed with a ^{32}P -labeled clone λ Aeg1 [10] with sequences homologous to the rRNA (B). (A) 5 μg each of embryonic (E), early larval (EL), third instar larval (L3), fourth instar larval (L4), early pupal (EP), late pupal (LP) total RNA, and 1 μg of adult female salivary gland (S) total RNA, probed with D7. (B) same as (A), but probed with the ribosomal clone. (C) 2 μg each of adult female salivary gland (S) and carcass (whole body with salivary glands removed) (C) total RNA, and 1 μg each of male (M) poly(A)⁺ RNA and female (F) poly(A)⁺ RNA probed with SGD7-1.

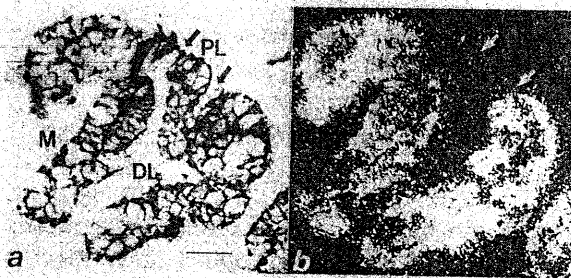


Fig. 3. Spatial distribution of the D7 gene expression in adult female salivary glands. Adult female salivary glands were dissected and prepared for in situ hybridization as in James et al. [10]. Sections were probed with antisense RNA probes made to the coding region of D7: (a) Giemsa-stained section of the female gland showing the medial (M), proximal-lateral (PL) and distal-lateral (DL) lobes. (b) Dark-field image of the same tissue section as in (a) showing the grains corresponding to hybridization. Note that the proximal-lateral lobes (arrows) are the only regions not expressing the D7 RNA.

ment of the λ SGD7-12 cDNA and hybridized as in James et al. [10]. A total of 52 antisense-probed and 7 sense-probed female salivary glands were examined. Positive hybridization was seen only with the antisense probe and was confined to the distal-lateral and medial lobes of the female salivary glands (Fig. 3). These regions of the gland are highly differentiated in females as compared to males, and this sexual dimorphism coincides with the female-enriched expression of the D7 gene. Repeated efforts to detect localized hybridization in male salivary glands were unproductive; after 6-week exposures, a few grains were scattered over all regions of the male salivary glands.

Characterization of the D7 gene. A genomic clone, λ AECD7, was selected by homology to the cDNA clone λ SGD7-12. From this genomic clone a single *Eco*RI restriction endonuclease fragment 3819 bp in length was sequenced (Fig. 4). This fragment was determined to contain the entire

coding region of the D7 gene by comparison to sequenced cDNAs. Four introns, I-IV, interrupt the coding region. The observed splice junctions conform to known consensus acceptor and donor sequences. Other features consistent with known eukaryotic genes, a TATAAA sequence and a consensus polyadenylation sequence, AATAAA, were observed. The probable start of transcription was defined by comparing primer extension products and dideoxy sequencing of RNA, and thus the nucleotides C and A were designated as +1 and +2, respectively (Fig. 5). Although primer extension analysis and RNA sequencing can be confounded by secondary structure of the RNA, the designated transcription start site falls between a consensus TATAAA sequence and the start of the translated portion of the gene.

The coding region of this genomic clone displayed nucleotide sequence polymorphisms with the sequenced cDNAs, λ SGD7-11 and SGD7-12 (Fig. 6). In addition, cDNAs λ SGD7-11 and λ SGD7-12 also show sequence polymorphisms between themselves. Thus we have sequence information on three variants of the D7 gene from the Rockefeller strain. As discussed below, whether this variation results from multiple alleles or multiple copies of the D7 gene is still to be determined.

Characterization of the D7 cDNAs and putative translation products. Eleven different cDNAs with homology to the original clone, λ SGD7-1 were isolated from a salivary gland cDNA library and the largest were selected for further characterization. As mentioned previously, the sequenced cDNAs, λ SGD7-11, λ SGD7-12, displayed nucleotide sequence polymorphisms (Fig. 6). The sequence of an additional cDNA clone, λ SGD7-8, was identical to that of λ SGD7-12. In general, the 5' ends of the cDNAs are more conserved than

Fig. 4. Primary nucleotide sequence of a genomic DNA fragment containing the D7 gene. A 3.8-kb *Eco*RI fragment containing the complete coding region of the D7 gene was subcloned from λ AECD7 and sequenced as described in Materials and Methods and Fig. 1. Various features of the sequence are listed in bold. A consensus eukaryotic TATAAA and polyadenylation consensus sequence (AATAAA) are present. The most likely translation initiation (ATG) and termination (TAG) codons are indicated. The dinucleotide CA were identified as the transcription initiation sites by primer extension analysis and direct dideoxy sequencing of RNA and the C has been denoted +1 in the sequence. Introns I-IV are in lower-case and underlined. In addition, the putative intron in the 3' end non-translated region is underlined. The restriction endonuclease cleavage sites for *Stu*I and *Bam*HI used for subcloning portions of the putative promoter are indicated.

GAATTCAGGAACCTTTAAATGAATCTCTTGAATATCCCTGGAACATTACTCGTGAAGTCTTGAGAAATTTTCAAGTATAAATCTGGGAAGAACTCTTGCATGAATCCCTGGGGTAAT
 -2345 -2230
 TTATGAAGAAATCCTTTGAGGTATTTTGAAGAAATCTTTGACGTTGAATGCATTTCCTTGGCTTGAATATTTTAAAGTTTGGTGAATATAGAAATCTTGAGAAATCTGTGATAA
 -2110
 TATTTTGAATATATCTTGAATATATTTTGAAGAGTACCAGCATTATTTGAAGTGAATCCCTGGCGAATTTTGTGAATATCTCAAGAAATTTCAAAATGTCTGTGAAAAACC
 -1990
 CTGAAGGAAATTAAGGATAATTTGACAGATTTTTTTTATTTATTTGACTCTTTATTTATGGAGAAATTCAGCCCTTAGCTAGTTCATCTCTTGACAGAAATTAAGGAAACGACTCTG
 -1870
 TTTTGTATCTACTAATTCAGAGTCCAACGCTCGGGTTTTGAATTTCCCATATTTCTACACAAATCCATGAAGAAATAGTCGAGAACTGTCAAAAGTTATCAAAATATTGA
 -1750
 AAACGGATCGAGACTAGCACTGTAAGAGTCACTATTGTCTATTTCAGTTACAGGGCCAAATCTTGACACAAACACAAGCTTCACCAACTCTCCAGAAATTTTAAAGCATCTACT
 -1630
 GATATTGGCTTAGATATTTATACGTTATCTCTAGATTGGCCAGTTGTAGTTACATGAAGGCCAAATGTTAAAGTTAAATATCCGAGAGTCCATTGTCAATCAGTGAAGTAT
 -1510
 TTTAAGCCAGATGAAGGGTATTATCAACATTTTTTCCAAACCCCTCCGGGTTCTATTTGCTGAGAAGGAAAAACCGAAGTGATTTTTTGGCTCTGCTGCTGCATATAACCTT
 -1390
 ACCGCTCTCAATTAACCTTTCCGTCAAAGTCTTCATAGTCCAGCAATAAAAAAAGATGAATTTCTTCAATAAGGTATCACCAAGACACATACCATACCGAGGACCAATATGCA
 -1270
 GTACTAGAGTGATACCAATCAGAGCCGAGTGGGACGGACCTGGTACTCGACATATTCACTGTTTACTGATGTCGTACAAACAGCTATTGCTGATTTCTTCTGCTGCCACTTC
 -1150
 AAATGAACCTCTGAGTGAACGGAGCAACATCCAACTTTCTTCAGCAATCTTGTCTACCCATCAGAAAGTGTATTTATGAGGACTCTGAGCGGTACTGCAATGCTCTCAT
 -1030
 ATGACGGCTGGAGGCACTGACGTTCACTTCCGATCTCTGGAGTGCTGCTGTTTCGACGCCATGTTGATTAAGTCTGAGTATTACTCTCAACAGAGCAGTAAGCTTCACCT
 -910
 TAAGGCTAAGTAACCCGTCATTGTTTTGCAACATGATGACTTTTCAGCTTGCAATTTCAAAATGATAAAGCAGCTCTGATAGTTTATCTGACTGAAAAAGTATCACTGTACGCA
 -790
 TTTACATGCATAAAGTATCTGATATTTTTGAGCTGTGCTGAGTGAAGCAACTGATTTCTTTGATTCGAATCATGAGTGAATAGCAACACCATCAAGCAGCGGTACAAATTT
 -670
 TAATGACGGCTACTTCCGCTTACAGCCTTGTGCTGATGACGATCTTGCAGTTGCGACAGGAATCTTCGCAAAATCTAGCTAGATTTCCGATCTGAGCAGTTTGTAAAAAAA
 -550
 ATATTTTCAGTTGAGAAATGATTTTCTCACCTCAGAGGTTACCGGATCAACCGTATTAAAGTTGTCAATCAGGATATCGCATCCAGTACCGATTTCTGGATAACGATGGCAAT
 -430
 CGATTAACCTTGTCTACGCAATTAATAGTTCTAGTTGTCTTAAGCAATCTATCAGGCTTTTTGGTTCTCGGAAATCTGATTAATTCATCGACATTACCGTCCGAGGCTCGAGC
 -310
 TTGATTTGCGCAACATATTATGATGATTTGTTGAATACATAAGCTTTTGAAGCCTTAATCCAATAATAGATCTCTTCAGAGACGAATGGAAGCACCATAGTACTTTATAATAT
 -190
 TGAACATTTGTAGTACTGTGACGCTGTGACACATTATGATACATATGGATGATTTGATCAGCTTTACTCAATCTTTCTTCACACTCTGCTGGTTGTTATATTGCGTAAATCTATT
 -70
 ATCTGTTATCAGTCTGACTGCTAATCACATTCACTTGCATATAAAACGCAATTAAGCCTGTCTAACATTTCATCCGATCATACTTACAACTAGATAAGGCTACTTCTATTACTCGCA
 +51
 ATTTTACAACTTTTCCGTGtaagatgttccatagtgcatccaactcttcttcttaacaaatactccaccttttaagGTTGCCCTCAATGGGACCATTTGATCCGAGGAGATGCTGT
 +171
 TCACTTTACGCGCTGCATGGAAGACAATTTGAAGATGGAGCGAATCGACTTCCAATGTTAGCGAAATGSAAGAATCGATTAAACGAACCGGTAGACAGCCCGCAACTCAGTGTTCG
 +291
 GCAATCGCTCTGGTAAGAACAGGTCTGTACGATCCGGTAGCCCAAAAGTTCGATgaagttggtggtgatatccgggtggaattggataatttcatcgctcaattgtttagGCGT
 +411
 CGGTGATCCAGGAGCAATTTAAGGCTTATCCGCTCTTGGGGAAAAGAGCAAGTTGAAGCATATGCTAACGCAATTAAACAGTTGCCCTCCACAATAACGACTGTGCCGCTGTTTCA
 +531
 AAGCGTACGATCCTGTTCAAGGCGCATAGGACACAGCAAGAATCTGTTCCATGSAACAAAGGATTTGACCAAGGGCCTCTATGAGAAGTTGtaagtggttccgtg!attcgatt
 +651
 gaaacaattgtgtgtatgaacata!taattgcttcagGGAAGAGCAATTCGCCAGAAAGCAATCTACTTCGAGTTTTCGAGAAAGTACTACCCAGCTGGGTCAGATAGCGTCA
 +771
 GCACTTTGTGATAGGCAATACACTGTCTTAGATGATGCGCTGTTCAAGGAGCAGCTGATTGCGTATGAAGGTTATTCGATACATTACGAAGGATAATCAACTGGATgaagaaa
 +891
 ataataatatcatgaacatcatagtaatcgacatttcccttcagTGAAGAGGTGAAGCGGACTTCAGCTAGTGAATAAGATACGAAGGCACGCAAAAGGTTTTGAATGACT
 +1011
 GTAAGTCAAGCAACCAAGCAACGCAAAAGAGAAATATGCGCACTACTACAAATGTTGGTGAATCTTCGTTAAGGATGATTCAAGAGGCTTTGACTACTGAGGAGTACCGTAC
 +1131
 AGATTTATGCTTCAATTTGCCAAGAACAGGCTTACAGCAACAGCAGTCAATCTCAAGTATGGAATCGACGTAACAGTGTCCCAATAGAAATAGTGCATTGAAAATGCTT
 +1251
 CCAATTCATGTTTTCATCGAAGTAGAAGTAAATAAGATATAAATCATCAGCAGAGGAACCTTGTTTAACTTATATGTTTATTTGGTAATGTTGATGTTTAAATTA
 +1371
 TTTAATTTGGGAAGCCGAAAAACCGTTAATAGGAATTAAGTAGTTTTGAATCATGTTATTATAAATATCGCATAGTTCTGACTGAAGAATTC
 +1470

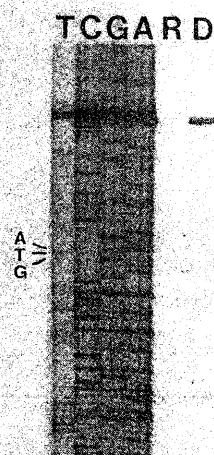


Fig. 5. Comparison of primer extension and dideoxy sequencing products of female salivary gland RNA. RNA was annealed to oligonucleotide primers D7R12 or RD7GEN12, and subjected to dideoxy sequencing reactions (lanes T, C, G, A) or primer extension analysis (lanes R and D), as described in Materials and Methods. Both methods produce strong stops at identical locations, 26 and 27 nucleotides upstream from the initiation codon ATG using the D7R12 primer. The RD7GEN12 primer overlaps the putative transcription initiation site and therefore produces no extension product. Lanes T, C, G, A are the T, C, G, and A, sequencing reactions, respectively (note that the RNA sequence in the figure has been presented as the complementary, sense strand, so that it may be read directly), and lanes R and D are the primer extension products generated by primers RD7GEN12 and D7R12, respectively.

the 3' ends. Of the 31 bases differing between the coding region of λ SGD7-11 and λ SGD7-12, only 5 cause changes in amino acids. λ SGD7-12 was the longest cDNA recovered, and by comparison with the genomic sequence was missing only 21 nucleotides from the 5' end. However, this cDNA contains a deletion in the 3'-end non-translated region proximal to the consensus polyadenylation signal sequence that is not observed in λ SGD7-11. Nucleotide variation in that region may have created a small intron that was excised before this particular cDNA was synthesized.

We have generated a putative protein sequence based on the major open reading frame present in the sequenced cDNAs (Fig. 6). The putative open-reading frame of λ SGD7-12 was cloned and expressed in the phage T7 system [13], generating a protein with the predicted molecular weight, con-

firmed the reading frame of the D7 mRNA (O. Marinotti and A.A. James, unpublished). The deduced protein is 321 amino acids in length with a predicted minimum size of 36 920 Da. The protein is lysine-rich (12% of amino acid residues), has 10 cysteine residues, and a hydrophobic amino terminal region characteristic of a secretory signal peptide [16,17]. Several of the salivary gland secretory proteins are glycosylated [10,15], however, the deduced D7 protein lacks consensus sites for asparagine-linked glycosylation. A search of the PIR database for similarities of the D7 deduced protein with other reported proteins yielded no definitive functional similarity with other proteins.

SDS-PAGE-isolated recombinant protein was injected into a rabbit to produce antibodies to the D7 product. Fig. 7 shows that the serum recognizes a 37-kDa protein present in female salivary glands and saliva but not in males. Additionally, the antibodies recognize recombinant and native D7 protein following both native and denaturing electrophoretic separation and transfer protocols (C. Ghosn, O. Marinotti, and A.A. James, unpublished). These data indicate that we have been able to generate antibodies to the D7 protein, that the protein is expressed only in females, and that the D7 protein is a secreted protein as predicted from the signal peptide.

Discussion

The D7 cDNAs and homologous gene were isolated using standard differential screening techniques with labelled salivary gland-specific RNA. The D7 gene encodes an mRNA that is preferentially expressed in the distal-lateral and medial lobes of the adult female salivary glands. The high number of D7 representative clones in the salivary gland library, the short time of exposures of the in situ hybridizations, and the strong signal on the Northern blots indicate that the D7 gene is abundantly expressed.

The D7 gene encodes a major secreted salivary gland protein of 37 kDa. The putative signal peptide in the D7 protein indicates that it is a secreted salivary gland product, confirmed by the antibody detection of the D7 protein in saliva. The abundance and molecular weight of the D7 deduced gene product indicate that it is most likely the

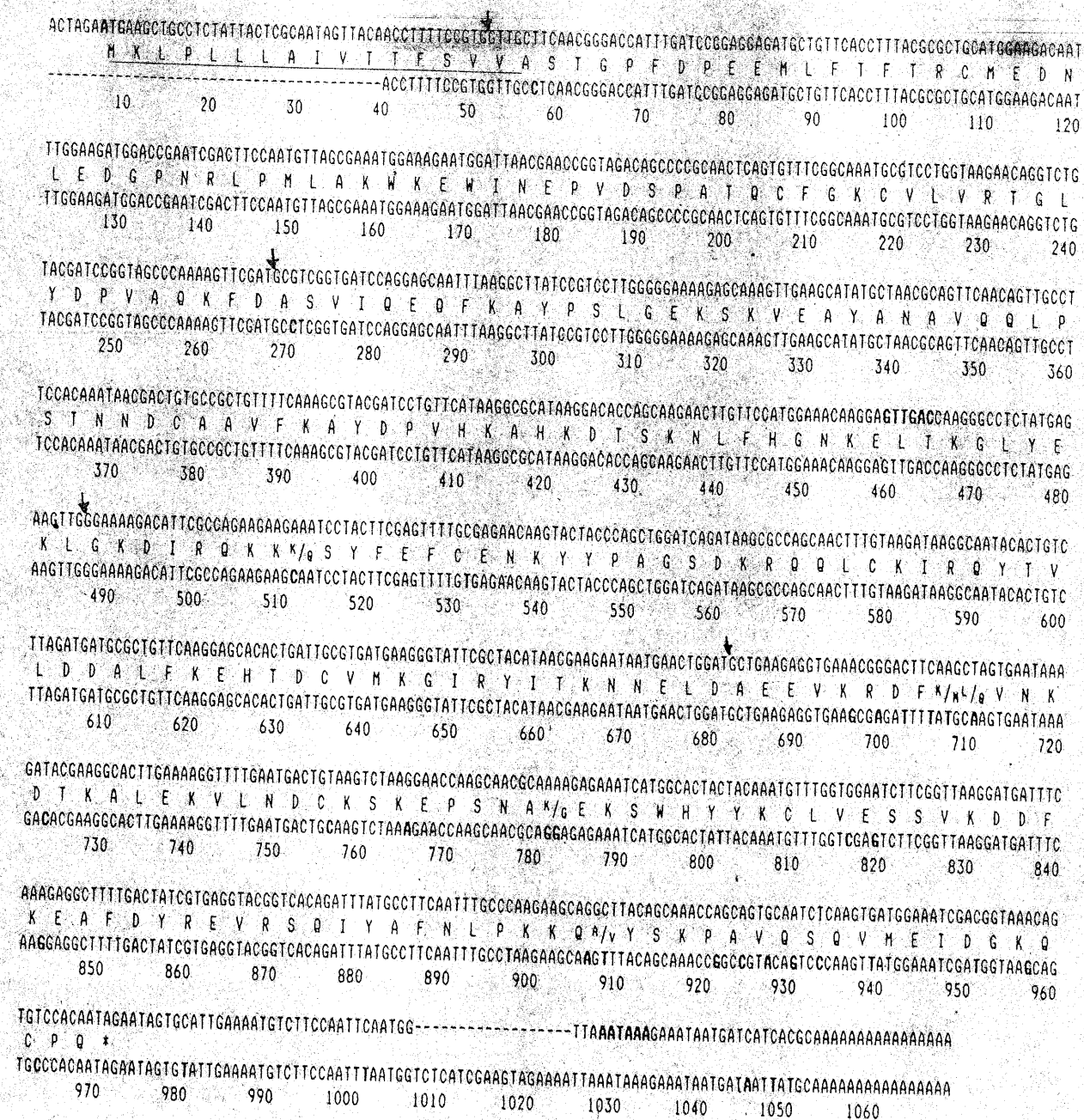


Fig. 6. Primary nucleotide sequence of the D7 cDNAs and their deduced amino acid sequence. The cDNA clones, λ SG7-12 (top) and λ SGD7-11 (bottom) were sequenced as described in Materials and Methods and in Fig. 1, and aligned to show sequence homology. The deduced amino acid sequence is listed in the middle with the putative secretory signal peptide underlined. Numbering refers to the nucleotide sequences and dashed regions represent sequences not found in the respective clones. Nucleotide sequence polymorphisms are highlighted in bold in λ SGD7-11. Amino acids differences resulting from sequence polymorphisms are indicated as doublets with the first amino acid derived from λ SG7-12 and the second amino acid from λ SGD7-11. A translation initiation codon (ATG) and consensus polyadenylation signal sequence, and the restriction endonuclease *HincII* site (GTTGAC) are listed in bold in the λ SG7-12 sequence. The positions of introns I-IV are indicated by the vertical arrows.

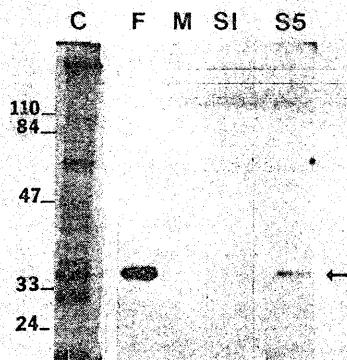


Fig. 7. Antibody detection of D7 protein in female salivary glands and saliva. C, Coomassie Blue-stained proteins from 6 pairs of whole female salivary glands separated by SDS-PAGE. F, antibody detection of D7 from 6 pairs of female salivary glands. M, antibody detection of D7 from 20 pairs of male salivary glands. S1, S5, antibody detection of D7 in saliva collected from 26 females, 1- and 5-day exposures, respectively. Proteins from saliva and salivary glands were resolved by SDS-PAGE, transferred to nitrocellulose filters, and specific recognition determined as described in Materials and Methods. Numbers refer to molecular weight markers in kDa.

major secreted salivary gland protein, designated polypeptide 7, by Racioppi and Spielman [15].

We take the signal evident in Northern analysis of male poly(A)⁺ RNA and the lack of a discernible pattern of expression the *in situ* hybridization to tissue sections of male thorax as an indication that all of the cells in the male glands synthesize the D7 RNA at a low level. However, it is possible that some other tissue in males is expressing the D7 RNA. If D7 expression is confined to the salivary glands, then sex-specific regulation of the D7 gene product appears to occur both transcriptionally and translationally. The low level of RNA expressed in males does not yield a protein detectable by polyclonal antibodies.

Efforts to resolve the origin of the nucleotide sequence differences between the genomic and cDNA clones have been so far unsuccessful. We do not yet know if these differences result from allelic variation of a single locus or from multiple genes. Southern analyses have revealed the presence of a copy of a dispersed, highly repetitive DNA sequence in a region adjacent to the D7 gene (C. Ghosn and A.A. James, unpublished), and variation in this region may account for the multiple bands often seen in these analyses and

complicate the standard approaches to evaluating gene copy number. We have isolated additional D7 genomic clones and the analysis of these should provide an answer.

Mosquitoes are holometabolous, having distinct larval, pupal, and adult stages. Larvae feed on bacteria and other particulate organic matter in their aqueous environment, adult males feed on nectar [18], and adult females feed on nectar and blood. The restricted expression of D7 in adult female salivary glands indicates that it is involved in one of the two adult feeding modes. The female salivary glands are morphologically distinct from the male salivary glands, possessing large distal-lateral and medial lobes not found in males. The fact that only females feed on blood is coincident with this structural dimorphism and it has been proposed that the female-specific lobes synthesize and secrete the salivary components involved in blood feeding [19]. For example, an apyrase (ATP-diphosphohydrolase) has been described in the female salivary glands. The apyrase acts as an anti-platelet aggregating factor and helps the female ingest blood [1]. The apyrase activity has been localized to the medial lobe and distal-lateral lobes of the female salivary gland [20]. Conversely, two salivary gland activities found in both males and females, an α -glucosidase [21] and a lysozyme-like activity [22], are restricted to the proximal-lateral lobes of the female salivary gland. These observations combined with the stage, sex, and spatial restricted expression of the D7 gene make it likely that the D7 gene product is involved in blood feeding.

The lack of reported proteins with sequence similarities to the D7 gene product is not surprising. Efforts to investigate the biochemistry and molecular biology of the secreted salivary gland proteins of hematophagous arthropods have only recently been initiated and there is a potential that many new families of proteins will be discovered.

In addition to the characterization of novel pharmacologically active products, the analysis of genes encoding the salivary gland proteins provides information about the DNA sequences that are good candidates for genetic transformation of vectors. The promoters of these abundantly expressed genes could be used to direct the expression of exogenous DNA sequences, which en-

code antipathogen activities, in the female salivary gland. This organ is a good target, as mentioned before, because it is a site where most pathogens must enter before transmission to a vertebrate host.

The functional analysis of vector insect promoters and their application to directing the expression of exogenous genes is now dependent on the development of techniques that enable the integration and stabilization of exogenous DNA sequences in the target insect genome. Our efforts in characterizing the 5'-end controlling DNA sequences containing the putative promoters of the abundantly expressed salivary gland genes anticipate the development of protocols for genetically transforming mosquitoes.

Acknowledgements

This work was supported by funds made available by the John D. and Catherine T. MacArthur Foundation, by a grant from the National Institutes of Health, NIAID, AI29746, and in part by BRSG S07 RR07008 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, NIH.

References

- 1 Ribeiro, J.M.C. (1987) Role of saliva in blood feeding by arthropods. *Annu. Rev. Entomol.* 32, 463-478.
- 2 Rossignol, P.A. and Spielman, A. (1982) Fluid transport across the ducts of the salivary glands of the mosquito. *J. Insect. Physiol.* 28, 574-581.
- 3 Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Isolation of biological active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18, 5294-5299.
- 4 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- 5 Bender, W., Spierer, P. and Hogness, D.S. (1983) Chromosomal walking and jumping to isolate DNA from the Ace and Rosy loci and the Bithorax Complex in *Drosophila melanogaster*. *J. Mol. Biol.* 168, 17-33.
- 6 Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503-517.
- 7 Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* 113, 237-251.
- 8 Maizels, N. (1976) Dictyostelium 17S, 25S, and 5S rDNAs lie within a 38,000 base pair repeated unit. *Cell* 9, 431-438.
- 9 Colot, H.V. and Rosbash, M. (1982) Behavior of individual maternal pA⁺ RNAs during embryogenesis of *Xenopus laevis*. *Dev. Biol.* 94, 79-86.
- 10 James, A.A., Blackmer, K. and Racioppi, J.V. (1989) A salivary gland-specific, maltase-like gene of the vector mosquito, *Aedes aegypti*. *Gene* 75, 73-83.
- 11 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- 12 Lipman, D.J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches. *Science* 227, 1435-1441.
- 13 Tabor, S. and Richardson, C.C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* 82, 1074-1078.
- 14 Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* 227, 680-685.
- 15 Racioppi, J. and Spielman, A. (1987) Secretory proteins from the salivary glands of the adult *Aedes aegypti* mosquitoes. *Insect Biochem.* 17, 503-511.
- 16 Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105-132.
- 17 Von Heijne, G. (1985) Signal sequences-the limits of variation. *J. Mol. Biol.* 184, 99-105.
- 18 Christophers, S.R. (1960) *Aedes aegypti*, the Yellow Fever mosquito. Cambridge University Press, Cambridge.
- 19 Marinotti, O., James, A.A. and Ribeiro, J.M.C. (In Press) Diet and salivation in *Aedes aegypti* female mosquitoes. *J. Insect Physiol.*
- 20 Ribeiro, J.M.C., Sarkis, J.J.F., Rossignol, P.A. and Spielman, A. (1984) Salivary apyrase of *Aedes aegypti*: characterization and secretory fate. *Comp. Biochem. Physiol.* 79B, 81-86.
- 21 Marinotti, O. and James, A.A. (In Press) An alpha-glucosidase in the salivary glands of the vector mosquito, *Aedes aegypti*. *Insect Biochem.*
- 22 Rossignol, P.A. and Leuders, A.M. (1986) Bacteriolytic factor in the salivary glands of the *Aedes aegypti*. *Comp. Biochem. Physiol.* 83B, 819-822.