

Sequence of a cDNA from the mosquito *Anopheles gambiae* encoding a homologue of human ribosomal protein S7

Cristina E. Salazar¹, Diane Mills-Hamm², Vasantha Kumar³ and Frank H. Collins^{1,2,*}

¹Department of Biology, Emory University, Atlanta, GA 30322, ²Malaria Branch, Division of Parasitic Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30333 and ³Department of Cellular and Developmental Biology, Biological Sciences Laboratory, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA

Received July 15, 1993; Accepted July 21, 1993

GenBank accession no. L20837

Studies of ribosomal proteins have given insight to the function and structure of the ribosome. During ribosome assembly, individual ribosomal proteins that initially recognize rRNA must bind to the assembly particle in a certain order and work cooperatively, rather than independently (1). Two of these important ribosomal proteins are S7 and S8. Detailed study of these and other ribosomal proteins is vital to the understanding of protein synthesis.

Here we report the structure, chromosome location and complete sequence of a cDNA in the mosquito *Anopheles gambiae* encoding a homologue of the human ribosomal S7 protein. The clone was isolated from a cDNA library constructed using poly(A⁺) RNA of adult female and male mosquitoes from the G3 strain of *A. gambiae* (2). The complete 817 nucleotide-long cDNA contains a continuous open reading frame of 193 codons which potentially encodes a peptide with a molecular weight of 22206.90 Daltons and an isoelectric point of 9.945. Furthermore, this peptide shows a high level of basic amino acids (37), consistent with characterizations of S7 and S8-like proteins in other organisms (3). A presumed polyadenylation signal, AATAAA is located at bases 771–776.

At the amino acid level, the open reading frame shows 71.7% identity with human ribosomal protein S7; 70.6% identity with *Xenopus laevis* ribosomal protein S8, and 61.7% identity with rat ribosomal protein S8 (Figure 1). Nucleotide comparison of the *A. gambiae* ribosomal protein cDNA coding sequence reveals a 66.7% identity with human ribosomal protein S7, but there are no evident similarities among the 5' or 3' non-coding sequences.

In situ hybridization to the ovarian nurse cell polyrene chromosomes of *A. gambiae* revealed its location at division 38B in the *A. gambiae* genome.

ACKNOWLEDGEMENT

We thank Brian Holloway and the staff of the Center for Infectious Diseases Biotechnology Core Facility for synthesizing the oligonucleotides used in sequencing.

REFERENCES

- Moore, P.B. (1988) *Nature* **331**, 223–227.
- Salazar, C.E., Mills-Hamm, D., Wesson, D.M., Beard, C.B., Kumar, V.K. and Collins, F.H. *Insect Mol. Biol.*, submitted.
- Mariottini, P., Bagni, C., Annesi, F. and Amaldi, F. (1988) *Gene* **67**, 69–74.
- Bilofsky, H.S., Burks, C., Fickett, J.W., Goad, W.B., Lewitter, F.I., Rindone, W.P., Swindell, L.D. and Tung, C.S. (1986) *Nucleic Acids Res.* **14**, 1–14.

Ag	MVFGSKVIKAGNGEPDAFETQIGQAILLEENNSDLKQLRDLVITRAREVEF.NNKKAI	
Hs	SSSAKIVKPNKPKPDEFESGISOALLEENNSDLKQLRDLVITRAREVEF.EVGGGRKAI	
Xl	MFSTSAKIVKPNKPKPDEFESGISOALLEENNSDLKQLRDLVITRAREVEF.EVGGGRKAI	
Rn	MSSSAKIVKPNKPKPDEFESGISOALLEENNSDLKQLRDLVITRAREVEF.EVGGGRKAI	60
Ag	IIIVVVPKQKAFQKQVQTR.LVRELEKPKFSGKHVVVIAERRILLPKPNRGRDPPNKQKRP	
Hs	IIIVVVPQLKSPQKIQVR.LVRELEKPKFSGKHVVVIAERRILLPKPTRSRKTKKQKRP	
Xl	IIIVVVPQLKSPQKIQVR.LVRELEKPKFSGKHVVVIAERRILLPKPTRSRKTKKQKRP	
Rn	IIIVVVPQLKSPQKMAKSGKVRLEKPKFSGKHVVVIAERRILLPKPTRSRKTKKQKQKP	120
An	PNVTAVYDAILEDLVFPAEVGKRIRVKLDGSQLIKVHLDKQQTTEHKVDTFASVYKK	
Hs	RTLTAVHDAILEDLVFPAEVGKRIRVKLDGSQLIKVHLDKQQTTEHKVDTFASVYKK	
Xl	RTLTAVHDAILEDLVFPAEVGKRIRVKLDGSQLIKVHLDKQQTTEHKVDTFASVYKK	
Rn	TLTSSADRRILEDLVFPAEVGKRIRVKLVGSQLIKVHLDKQQTTEHKVDTFASVYKK	180
An	LTGRDVTFFPPENYLX	193
Hs	LTGKDVNFFPFPQLX	192
Xl	LTGKDVNFFPFPQLX	194
Rn	FLGKDVNFFPFPVSVX	194

Figure 1. Amino acids alignments of S7 and S8-like proteins from mosquito, human, frog and rat. Comparison of the *A. gambiae* ribosomal protein with GenBank data base revealed identities of 71.7% with human ribosomal protein S7 (HsrpS7A: accession no. M77233), 70.6% with *Xenopus laevis* ribosomal protein S8 (MrpS8B: accession no. M21486), and 61.7% with rat ribosomal protein S8 (RnrpS8MR: accession no. X56846). Underlined amino acids indicate long stretches of high homology among all four organisms. Hs sequence lacking Met start codon obtained from GenBank data base, accession number listed above (4).

*To whom correspondence should be addressed at: Malaria Branch, Mail stop F-12, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA