J Mol Evol (1993) 37:457–463

Journal of Molecular Evolution

© Springer-Verlag New York Inc. 1993

A New Member of the Balbiani Ring Multigene Family in the Dipteran *Chironomus tentans* Consists of a Single-Copy Version of a Unit Repeated in Other Gene Family Members

Joakim Galli, Lars Wieslander

Department of Cell and Molecular Biology, Medical Nobel Institut Karolinska Institutet, S-104 01 Stockholm, Sweden

Received: 5 February 1993 / Accepted: 13 April 1993

Abstract. The known Balbiani ring (BR) multigene family members in the dipteran *Chironomus tentans* encode salivary gland secretory proteins in the size range between 38 and 1,000 kDa. The proteins interact to form protein fibers used by the aquatic larvae to spin feeding and protective larval tubes or pupation tubes. Here, we describe a new BR multigene family member, the sp17 gene, which codes for an 89-amino-acid-long protein with a relative mobility of 17k. The gene has a high content of charged amino acid residues and consists of two structurally different halves. Five regularly spaced cysteine codons are present in the 5′ half while the 3′ half contains five proline codons. These two different halves exhibit similarities to the C and SR regions, respectively, which form the tandemly repeated units in the about 40-kb-long BR genes and which also, in different versions, are the building blocks of all genes in the BR multigene family.

In this multigene family, encoding interacting structural proteins, the long BR genes with their 125–150 tandemly arranged repeat units as well as the short sp17 gene with its single-copy version of such a repeat unit, have therefore evolved from a common ancestor.

Correspondence to: L. Wieslander

Key Words: Gene family — Repetitive sequences — Secretory proteins — Balbiani ring genes

Introduction

Multicomponent biological structures have often evolved by duplication, rearrangement, and divergence of specific genetic elements (e.g., Jones and Kafatos 1982; Mishina et al. 1985; Weatherall and Clegg 1979). The larval tube made by the dipteran species *Chironomus tentans* is one such example (Case and Wieslander 1992). The salivary gland cells in the *C. tentans* aquatic larvae produce approximately 15 different secretory proteins, ranging in size between 12 and 1,000 kDa. The secretory proteins are synthesized during the entire larval period. As they are excreted through the gland duct they form water-insoluble protein fibers, which are spun into a larval protective and feeding tube (Grossbach 1977; Case and Wieslander 1992). During the later stages of larval development, the protein composition changes and pupation tubes are formed. The largest proteins are coded for by the four closely related Balbiani ring (BR) genes BR1, BR2.1, BR2.2, and BR6 (Edström et al. 1980; Rydlander et al. 1980; Kao and Case 1985; Case 1986; Botella et al. 1988). Each gene is composed of 125–150 repeat units arranged in one uninterrupted array (Wieslander and Paulsson 1992, Paulsson et al. 1992b). A repeat unit has two structurally different

parts, the C (constant) and the SR (subrepeat) region (Pustell et al. 1984; Wieslander et al. 1984; Grond et al. 1987).

Seven additional genes coding for intermediate and small-sized secretory proteins have so far been described. (For references see Case and Wieslander 1992; Galli and Wieslander 1993.[1]) At least five of the genes are built from elements structurally related to the C and/or the SR regions of the large BR genes and three of the genes are internally repetitious. Almost all of the tissue-specifically expressed genes therefore belong to the same multigene family, the Balbiani ring (BR) gene family, in which the individual genes have diverged both in regulatory and coding regions.

Our aim is to identify all the secretory protein encoding genes in order to be able to investigate possible evolutionary relationships between the different genes and identify functionally important changes that have resulted in the set of cooperating genes responsible for the secretory function of the salivary gland cells.

Here we report the characterization of a new salivary-gland-specific secretory protein gene. This sp17 gene is transcribed into a 0.7-kb-long mRNA and codes for a protein with a relative molecular weight of 17k. The gene is composed of structural element similar to those present in the large BR genes, having both one cysteine-codon-containing C-like region and one proline-codon-containing SR-like region. The sp17 gene resembles a version of a single BR gene repeat unit and therefore belongs to the BR multigene family. The evolution and function of this BR repeat until monomer gene are discussed.

Materials and Methods

Extraction of DNA, RNA and Proteins. High-molecular-weight DNA was extracted from cultured *C. tentans* epithelial cells (Wyss 1982) as described (Gross-Bellard et al. 1973). For extraction of RNA and proteins, salivary glands were dissected manually from fourth-instar larvae, fixed in 70% ethanol at 4°C, and stored in glycerol:ethanol (1:1) at −20°C. RNA was extracted as described (Eström et al. 1982). RNA to be used as template for cDNA primer extension was pelleted through a cushion of 5.7 M CsCl in 0.1 M EDTA pH 7.5 in an SW 50 rotor at 35,000 rpm for 12 h. For extraction of the gland lumen proteins, all cells were removed from the fixed glands with dissection needles. Proteins were dissolved in 62.5 mM Tris-HCl pH 6.8 containing 3% SDS and reduced in 0.7 M mercaptoethanol.

Construction and Screening of cDNA and Genomic Libraries. The λ ZapII cDNA library was constructed, as described (Galli and Wieslander 1993), from size-selected cDNAs in the range of 0.3–1 kb. For construction of the λ Gem-11 genomic library see Galli and Wieslander (1993). The genomic library was screened with the ct.k37 cDNA as probe, labeled by random priming (Feinberg and Vogelstein 1983).

Plaque hybridizations were performed according to Benton and Davis (1977). Phage DNA was prepared from liquid cultures (Sambrook et al. 1989).

In Situ Hybridization. The ct.k37 cDNA clone was random labeled using biotinylated dATP and hybridized to salivary-gland squash preparations. The hybridization and the washing of the squash preparations were performed as described (Galli and Wieslander 1993).

Northern, Southern, and Western Blots. Total salivary gland RNA was denatured and electrophoresed as previously described (Paulsson et al. 1992a). As size markers an RNA ladder (BRL) was used. For quantitation of mRNA, the same amount of total RNA from the various larval stages (Ineichen et al. 1983) was loaded onto the agarose gels. After blotting, the filters were hybridized with the ct.k37 cDNA, ^{32}P-labeled by random priming.

DNA and RNA were transferred to nylon fibers (Amersham) by vacuum blotting.

For immunological detection of size-separated proteins, the reduced secretory proteins were electrophoresed in SDS-containing 15% polyacrylamide gels (Laemmli 1970). The proteins were transferred to nitrocellulose filters as described (Burnette 1981).

The filters were incubated with affinity-purified fusion protein antibodies. (See below.) The bound antibodies were detected with antibody alkaline phosphatase conjugates (Promega).

Sequence Determination of DNA. The cDNA inserts and the genomic fragments were sequenced using Bluescript vector (Stratagene)-specific primers and insert-sequence-specific primers (Innovagen). PCR fragments were recovered from agarose gels (Magic PCR Preps, Promega) before sequence determination. The dideoxy sequencing method was used in combination with ^{35}S-dATP and the modified T7 DNA polymerase (Sequenase, USB). The DNA sequence was partly determined by using both the dye-terminator kit (ABI) and the dye-primer kit (ABI) in combination with the ABI model 373A automatic sequencer.

Sequence analyses were performed with the programs of Devereaux et al. (1984).

Expression of Fusion Protein and Immunological Techniques. A 293-bp-long PCR fragment (positions 323–615 in Fig. 1) was subcloned in the pQE-13 expression vector (QIAGEN) and expressed as fusion protein with a derivative of mouse dihydrofolate reductase (DHFRS). The fusion protein was affinity purified in 8 M urea on QIAexpress NI-NTA columns as described by the manufacturer. The fusion protein was mixed with Freund's adjuvans and injected intradermally into rabbits. Specific antibodies were affinity purified as previously described (Galli and Wieslander 1993).

cDNA Primer Extension. An end-labeled oligodeoxynucleotide primer (positions 324–343 in Fig. 1) was hybridized to 20 μg of total salivary gland RNA in 0.4 M NaCl, 0.04 M PIPES buffer pH 6.5 for 3 h at 50°C. The RNA-primer complex was ethanol precipitated and reverse transcribed using Superscript RT-H (BRL). The obtained cDNA was ethanol precipitated, dissolved in formamide, and electrophoresed in 6% sequencing gels.

[1] Galli J, Wieslander L (1993) Structure of the smallest salivary gland secretory protein gene in *Chironomus tentans*. Submitted for publication

```
AAGTTGCAATAACATTCTTTGTCTTTACAATTTCACTGTACAGTTGATAATTATTAAAAT  60
GATAATTAAATTTTACTAACATTATCTTTATCTTTTTAATAATACTTAAATGTATTGATA 120
AGCATCAAGAGCTTTTTCCCCTTCCCTTAATACTCATAAATTTTATATGTTAGCTTTCGA 180
AGCTTTCCTTTAAACTTTTAAAGCATATTTTTATATAAAAGGATGGATCACATCATCATC 240

   ▼
GGCACGTTAGTTACCGTTTAGTAATCATTAACCAGGTACGATGAGATTCACTTTAATTTT 300
                                          M  R  F  T  L  I  F

CCTTATAGTGCTCGCTTGCATCGCTTTCACACTTGCATGTGATAAGAAATGTGCAGCAAA 360
 L  I  V  L  A  C  I  A  F  T  L  A (C) D  K  K (C) A  A  K

AAAAGCAAAAAGAGCGGAAGAAAAGGCATTGAAAAAAGCATGCAAGGAATTGAATGGAGG 420
 K  A  K  R  A  E  E  K  A  L  K  K  A (C) K  E  L  N  G  G

CACCAAGATCAAGTTGGAAATTGACTTAGATGGAGCATGTAGAGGAAAATGCAAAAACAA 480
 T  K  I  K  L  E  I  D  L  D  G  A (C) R  G  K (C) K  N  K

GAAAAAGAAGGCTAAGAAGGCTGGTGAACCAAAGAAGATTACTAAGAAGATTGAAGGAAA 540
 K  K  K  A  K  K  A  G  E  P  K  K  I  T  K  K  I  E  G  K

ACCAAAACCAGGGCCATCAGCAGGGGGACCAGTAGCAGCAGCAGCACAAGCTAAAAATTG 600
 P  K  P  G  P  S  A  G  G  P  V  A  A  A  A  Q  A  K  N  *

AAGAGTTATTAAGGGATAGCAAACAATAGATATTAAAGTGTTTTTCGCTCAAATTTCATG 660
TATTTTGGATGTTTAACAACATACATCAACTAGACTTTCATGTTTATTTTATTCTCAATG 720
TAAAATAAAATTTTTAGAATCAAAA 745
```

Fig. 1. Nucleotide and corresponding amino acid sequence of the sp17 gene. The transcription start site is indicated by an *arrow-head*. The putative TATA box is *underlined by double lines* and the poly(A) addition sequence by a *single line*. The cysteine residues are *encircled*.

Dideoxy-sequencing reactions obtained by oligodeoxynucleotide priming on a subcloned fragment served as size marker.

Results

Structure, Chromosomal, Location, and Expression of the sp17 Gene

The aim was to isolate genes encoding the 15–17-kDa components of the salivary gland secretion. We therefore used a *C. tentans* λ Zap salivary gland cDNA library, constructed from cDNA in the size range of 0.3-1 kb, which should be enriched for sequences encoding proteins of the proper size. To detect gene sequences abundantly and specifically expressed in the salivary glands, the library was screened with ^{32}P-labeled cDNA made from the salivary gland poly(A)$^+$ RNA. In parallel, replica filters were screened with cDNA made from poly(A)$^+$ RNA present in all larval tissues but the salivary glands and with a mixture of probes representing sequences from the previously identified secretory protein genes (Case and Wieslander 1992). Clones were selected which were positive with the salivary gland cDNA probe and negative both with the cDNA probe from the other tissues and with the probes for the known genes. The cDNA inserts from the selected clones were individually hybridized to Northern blots of total salivary gland RNA. One cDNA clone, ct.k37, hybridized to an abundant mRNA with a size of 0.7 kb. (See below.) This 486-bp-long cDNA sequence, primed at the poly(A) tail, covered almost the entire transcript of the corresponding gene except for the first 16 bp of the 5′ untranslated region.

The structure of this gene, which according to convention was called the sp17 gene, is shown in Fig. 1. The structure was determined by a combination of direct sequencing of genomic fragments amplified by the polymerase chain reaction (PCR) and sequencing of a fragment isolated from a λ Gem-11 *C. tentans* genomic library. No introns are present in the sp17 gene, which was seen when the cDNA sequence and the genomic sequence were compared.

The transcription start site is located 37 bp upstream of the AUG translation start codon as determined by primer extension (Fig. 2). A putative TATA box is located 32 bp upstream of the transcription start site. In the 144-bp-long 3′ untranslated region a poly(A) addition sequence is present 23 bp upstream of the poly(A) tail. The sp17 gene is most likely present in the genome in a single copy as indicated by Southern blots of genomic *C. tentans* DNA digested with four different restriction enzymes (*Eco*RI, *Hin*dIII, *Pst*I, and *Xho*I). In each case only one band was detected (data not shown). The genomic location of the gene was determined by in situ hybridization of the ct.k37 cDNA to polytene salivary gland chromosomes. The sp17 gene is located on chromosome I, region 5C (Fig. 3).

In addition the steady-state level of the sp17 gene mRNA was analyzed. Salivary-gland RNA was prepared from different larval stages and analyzed in Northern blots (Fig. 4). The mRNA level is approximately the same throughout the fourth larval instar.

We conclude that a new *C. tentans* gene has been identified. This sp17 gene is expressed abundantly and specifically in the salivary gland cells. No significant homology to other sequences was found in the EMBL or the Swissport databases.

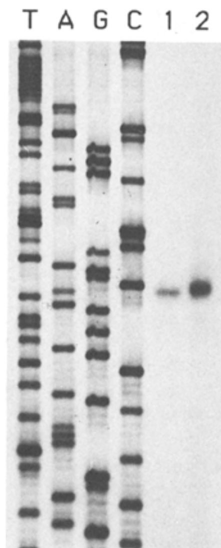

Fig. 2. Mapping of the transcription start site by cDNA primer extension. T, A, G, C and lanes are sequencing reactions serving as size markers. Lanes 1 and 2 show the extension product. Different amounts of material were loaded in the two lanes.

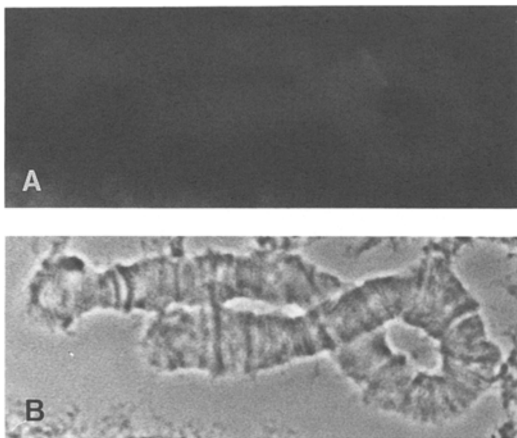

Fig. 3. Chromosomal location of the sp17 gene on chromosome I. A shows the immunofluorescence in region 5C of chromosome I after in situ hybridization with the ct.k37 probe. B shows the same chromosome in phase contrast. The chromosome I is about 200 μm in length.

Identification of the sp17 Gene-Encoded Protein and its Amino Acid Sequence

To identify the sp17 gene-encoded protein, a 293-bp-long coding fragment was expressed as a fusion protein which subsequently was used to immunize rabbits. The resulting polyclonal serum was affinity purified and used in Western blots of total salivary gland secretion. The antibodies specifically detected a protein with a relative mobility of 17k (Fig. 5). This protein is clearly visible after Coomassie staining and represents a quantitatively major component of the salivary-gland secretion. The molecular weight of the protein, calculated from the nucleotide sequence of the gene, is 9,312 Da. The difference between the calculated and the observed relative molecular weights is likely due to the basic character of the protein (see below) and/or to protein modifications. A similar discrepancy has been observed for three other secretory proteins in *C. tentans* (e.g., Galli and Wieslander 1993).

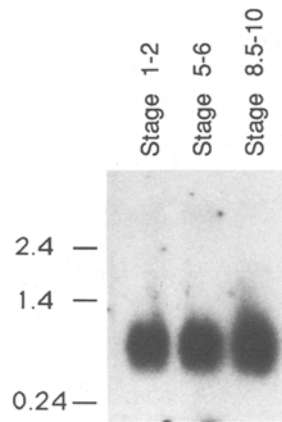

Fig. 4. Steady-state levels of sp17 gene mRNA at different stages in the fourth larval instar. The same amount of RNA was loaded in each lane and after blotting, the filter was hybridized with the ct.k37 cDNA probe.

The sp17 gene-encoded protein is 106 amino acids long, of which the first 17 constitute a typical signal peptide sequence. (See Fig. 1.) The protein excluding the signal peptide sequence has a highly biased amino acid composition—29% lysine, 18% alanine, and 10% glycine residues—giving an overall basic protein with an isoelectric point of 10.8. Five regularly spaced cysteine residues are present in the first half of the protein. The first two and the last two cysteine residues are grouped in pairs with a spacing of three residues in between the cysteine in each pair. The fifth cysteine residue is positioned between the two pairs, approximately at the same distance from each pair. In the C-terminal part of the protein five proline residues are present in a 33-residue-long region. In this region, lysines are very abundant, making up 13 of the 33 residues. Following this region is a short C-terminal part of nine residues, five of which are alanine. At the nucleotide level, four consecutive alanine residues are encoded by four tandemly arranged GCA codons.

Discussion

We have identified a previously unknown secretory protein gene in the midge *Chironomus tentans*. This sp17 gene is transcribed into a 0.7-kb-long mRNA and codes for a quantitatively major component of the salivary gland secretion with a relative molecular weight of 17k.

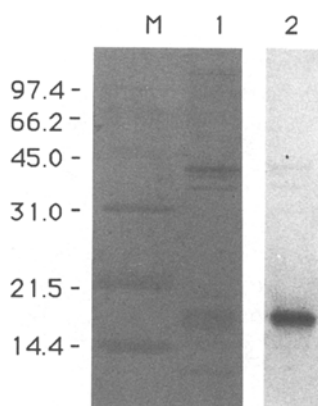

Fig. 5. Identification of the sp17 gene protein product. Salivary gland lumen proteins were separated in a 15% polyacrylamide gel, blotted onto a nitrocellulose filter, and probed with sp17 protein-specific antibodies. *M* stands for size markers, given in kDa. Lane 1 shows the stained secretory proteins and lane 2 shows the antibody reaction.

The sp17 Gene Belongs to the BR Multigene Family

Eleven *C. tentans* salivary-gland-specific secretory protein genes have so far been identified. At least nine of these share a common ancestry and form the BR multigene family. The proteins encoded by this multigene family can be divided into large, about 1,000-kDa proteins; intermediate, 100–200-kDa proteins; and small, less than 100-kDa proteins. A characteristic feature of the genes in the BR multigene family is that they are built from cysteine-codon-containing regions and/or regions containing short repeats with codons for proline and charged amino acids, often with typical −proline+ tripeptide motives (Pustell et al. 1984; Wieslander et al. 1984; Grond et al. 1987). These two kinds of regions are most pronounced in the four closely related BR genes, encoding the four large secretory proteins, where they are called the C (constant) and SR (subrepeat) regions. In most of the genes the C-like and SR-like regions are repeated; for example, in each BR gene a C and an SR region form a repeat unit which is present in 125–150 copies (Paulsson et al. 1992b), and the sp115,140 gene consists of about 65 copies of an SR-like region (Dignam et al. 1989; Galli et al. 1990). The sp17 gene encodes one of the smallest proteins in the salivary-gland secretion. Based on the following structural properties we suggest that the sp17 gene is related to the C and SR regions of the BR multigene family. First, the sp17 gene has two almost equally long, but structurally different, halves. In the 5′ part, five cysteine codons have a regular spacing and proline codons are lacking. In the 3′ part, five proline codons are present but no cysteine codons. In this latter part there is an accumulation of charged residues; in a 33-codon-long stretch, 13 codons are lysine codons (39%). Second, the 3′ part contains one copy of the SR region characteristic −proline+ tripeptide motif and one copy of the sequence glycine-lysine-proline-lysine-proline-glycine-proline, which is similar to subrepeats in the SR regions of the BR genes. Third, the overall length of the sp17 gene is almost the same as the length of a BR gene repeat unit and the 5′ and 3′ halves are very similar in length to the C and SR regions, respectively.

We conclude that the sp17 gene appears to be part of the BR gene multigene family and that it represents a monomer version of the repeat units building the large BR genes, with one C-like and one simple SR-like region.

Subfamilies Within the BR Gene Multigene Family

Within the BR multigene family, the large proteins are encoded by a subfamily of four BR genes which have arisen by a series of gene duplications (Paulsson et al., 1992b). The genes encoding the intermediate proteins are more heterogeneous, although it is possible that a late gene duplication produced two genes encoding very similar proteins in this size range (Dignam et al. 1989; Galli et al. 1990).

As for the small secretory proteins, three corresponding genes have been described apart from the sp17 gene. Two of the genes, the sp38-40.A and B genes, have arisen by a recent gene duplication event and are very similar to each other (Galli and Wieslander 1993). These genes have two C-like regions interspersed with three short SR-like regions and exhibit several similarities to the sp17 gene. The sp17 gene has a similar spacing of the cysteine codons compared to the first C-like region in the sp38-40.A and B genes. The genes are also similar in the SR-like regions. In the third SR-like region of the sp38-40.A and B genes, a 16-bp-long sequence coding for the sequence proline-lysine-proline-glycine-proline is found also in the sp17 gene. In the sp38-40.A and B genes, derivatives of this sequence are tandemly repeated from three to seven times. In addition, the sp38-40.A and B and sp17 genes have similar amino acid sequences preceding this proline-codon-containing region.

In view of these similarities it is plausible that the sp17 gene and the sp38-40.A and B genes represent a subfamily within the BR multigene family. Substantial rearrangements must, however, have occurred since the duplication of a common ancestor gene.

The sp38-40.A and B genes have been proposed to represent a prototype for the repeat units building the large BR genes (Case and Wieslander 1992;

Galli and Wieslander 1993). Sequence variations have been observed in alleles of the sp38-40.A and B genes which suggest how longer repeat-unit arrays, similar to the ones in the large BR genes, could have evolved from a short ancestor sequence (Galli and Wieslander 1993). The sp17 gene, with only one C-like region and one simple SR-like region, resembles one single repeat unit of the BR genes and is an even simpler version of a prototype repeat unit than the sp38-40.A and B genes.

Evolution of the Long, Repetitive BR Genes and the Short Nonrepetitive sp17 Gene Within the Same Multigene Family

The sp17 gene and the BR genes represent two extreme gene-structure versions within one gene family. From the same ancestor gene, sequence duplications and recombinations have in one chromosomal locus resulted in the BR genes consisting of about 150 almost identical tandem copies while in another locus, the sp17 gene has retained its single copy of a version of the same sequence. One important factor involved should be the functional demands on the respective gene product, and an understanding of these demands requires knowledge about the molecular interactions in the formation of the larval tube. The large BR gene-encoded proteins have been shown to be the backbone in the protein fiber excreted by the salivary glands (Wellman and Case 1989), and results from structural studies of synthetic peptides representing the C or SR region indicate that the C region is primarily α-helical, while the SR region appears to be a more extended poly(Gly)II-type helix (Wellman et al. 1992). The extended helical BR proteins are thought to interact in the assembly of the protein fiber by electrostatic interactions between SR regions and by covalent disulphide bridges between the C regions. A key feature is then the possibility for out-of-register aligning between separate proteins (Hamodrakas and Kafatos 1984; Lendahl and Wieslander 1984; Wieslander et al. 1984; Case and Wieslander 1992). The repetitive nature of the proteins encoded by the BR genes may be important for the molecular strategy adopted to make the salivary-gland silk fibers. In line with this view, the number of repeats within the BR genes has been found to be conserved (Paulsson et al. 1992a). Several other examples of internally repetitious genes encoding proteins which form protein fibers are also known (Tsujimoto and Suzuki 1979; Yamada et al. 1980; Xu and Lewis 1990). On the other hand, the relationship between a repetitive gene design and protein function is not immediately obvious. Protein fibers can be built from nonrepetitive proteins (Pollard and Cooper 1986), and there are several examples of internally repetitive genes coding for proteins which are later cleaved into the individual functional repeat units (Dworkin-Rastl et al. 1984; Rothnagel and Steinert 1990; Sorimachi et al. 1990; Poole et al. 1992).

The function of the nonrepetitive sp17 gene is not known. The sp17 gene is expressed abundantly and tissue specifically in the salivary-gland cells at approximately the same level throughout the fourth larval instar. It is therefore likely to be used during the entire larval period, both in the formation of the larval feeding and housing tube and the modified pupation tube, made at the prepupal stage (Case and Wieslander 1992). We have preliminary immunohistochemical results indicating that the sp17 protein is present in the protein fiber of the larval tube. In view of the similarities to the repeat units in the BR-coded proteins—secondary-structure predictions also suggest mainly α-helical structures—it is possible that the sp17 protein is needed to cross-link the long BR gene-encoded proteins in the protein fiber network.

Whatever the function of the sp17 gene, the making of the protein fiber network utilizes different proteins which consist of versions of structurally similar domains either in a single copy or in a large number of tandem copies, and within the same gene family the corresponding drastically different gene structures have evolved from a common ancestor.

Acknowledgments. We thank Kerstin Bernholm for excellent technical assistance. The work was supported by the Swedish Natural Science Foundation, Torsten and Ragnar Söderbergs Stiftelser, Magnus Bergvalls Stiftelse; and the Karolinska Institutet.

References

Benton WD, Davis RW (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–182

Botella L, Grond C, Saiga H, Edström J-E (1988) Nuclear localization of a DNA-binding C-terminal domain from Balbiani ring-coded secretory protein. EMBO J 7:3881–3888

Burnette WN (1981) Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Analyt Biochem 112: 195–203

Case ST (1986) Correlated changes in steady-state levels of Balbiani ring mRNAs and secretory polypeptides in salivary glands of *Chironomus tentans*. Chromosoma 94:483–491

Case ST, Wieslander L (1992) Secretory proteins of *Chironomus tentans* salivary glands: structural motifs and assembly characteristics of a novel biopolymer. In: Case S T (ed) Results and problems in cell differentiation, vol 19. Springer-Verlag. Berlin, pp 187–236

Devereaux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Dignam SS, Young L, Lezzi M, Case ST (1989) Identification of a developmentally regulated gene for a 140-kDa secretory protein in salivary glands of *Chironomus tentans* larvae. J Biol Chem 264:9444–9452

Dignam SS, Case ST (1990) Balbiani ring 3 in *Chironomus tentans* encodes a 185-kDa secretory protein which is synthesized throughout the fourth larval instar. Gene 88:133–140

Dreesen TD, Bower JR, Case ST (1985) A second gene in a Balbiani ring: *Chironomus* salivary glands contain a 6.6 kb poly$(A)^+$ RNA that is transcribed from a hierarchy of tandem repeated sequences in Balbiani ring 1. J Biol Chem 260: 11824–11830

Dworkin-Rastl E, Shrutkowski A, Dworkin MB (1984) Multiple ubiquitin mRNAs during *Xenopus laevis* development contain tandem repeats of the 76 amino acid coding sequence. Cell 39:321–325

Edström J-E, Rydlander L, Francke C (1980) Concomitant induction of a Balbiani ring and a giant secretory protein in *Chironomus* salivary gland. Chromosoma 81:115–124

Edström J-E, Sierakowska H, Burvall K (1982) Dependence of Balbiani ring induction in phosphate in *Chironomus* salivary glands on inorganic phosphate. Dev Biol 91:131–137

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Galli J, Lendahl U, Paulsson G, Ericsson C, Bergman T, Carlquist M, Wieslander L (1990) A new member of secretory protein gene family in the dipteran *Chironomus tentans* has a variant repeat structure. J Mol Evol 31:40–50

Galli J, Wieslander L (1993) A repetitive secretory protein gene of a novel type in *Chironomus tentans* is specifically expressed in the salivary glands and exhibits extensive length polymorphism. J Biol Chem 268:11888–11893

Gross-Bellard M, Oudet P, Chambon P (1973) Isolation of high molecular-weight DNA from mammalian cells. Eur J Biochem 36:32–38

Hamodrakas SJ, Kafatos FC (1984) Structural implications of primary sequences from a family of Balbiani ring-encoded proteins in *Chironomus*. J Mol Evol 20:296–303

Ineichen H, Meyer B, Lezzi M (1983) Determination of the developmental stage of living fourth instar larvae of *Chironomus tentans*. Dev Biol 98:278–286

Jones CW, Kafatos FC (1982) Accepted mutations in a gene family: evolutionary diversification of duplicated DNA. J Mol Evol 19:87–103

Kao WY, Case ST (1985) A novel giant secretory polypeptide in *Chironomus* salivary glands: implications for another Balbiani ring gene. J Cell Biol 101:1044–1051

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature 27:680–685

Lendahl U, Wieslander L (1984) Balbiani ring 6 gene in *Chironomus tentans:* a diverged member of the Balbiani ring gene family. Cell 36:1027–1034

Mishina M, Takahashi T, Takai T, Kurasaki M, Fukuda K, Numa S (1985) Role of acetylcholine receptor subunits in gating of the channel. Nature 318:538–543

Paulsson G, Lendahl U, Galli J, Ericsson C, Wieslander L (1990) The Balbiani ring 3 gene in *Chironomus tentans* has a diverged repetitive structure split by many introns. J Mol Biol 211:331–349

Paulsson G, Höög, Bernholm K, Wieslander L (1992a) Balbiani ring 1 gene in *Chironomus tentans:* sequence organization and dynamics of a coding minisatellite. J Mol Biol 225:349–361

Paulsson G, Bernholm K, Wieslander L (1992b) Conserved and variable repeat structures in the Balbiani ring gene family in *Chironomus tentans*. J Mol Evol 35:205–216

Pollard TD, Cooper JA (1986) Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. Ann Rev Biochem 55:987–1035

Poole CB, Grandea III AG, Maina CV, Jenkins RE, Selkirk ME, McReynolds LA (1992) Cloning of a cutical antigen that contains multiple tandem repeats from the filarial parasite *Dirofilaria immitis*. Proc Natl Acad Sci USA 89:5986–5990

Pustell J, Kafatos FC, Wobus U, Bäumlein H (1984) Balbiani ring DNA: Sequence comparisons and evolutionary history of a family of hierarchically repetitive protein-coding genes. J Mol Evol 20:281–295

Rothnagel JA, Steinert PM (1990) The structure of the gene for mouse fillagrin and a comparison of the repeating unit. J Biol Chem 265:1862–1865

Rydlander L, Pigon A, Edström J-E (1980) Sequences translated by Balbiani ring 75 S RNA in vitro are present in giant secretory protein from *Chironomus tentans*. Chromosoma 81:101–113

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sorimachi H, Emori Y, Kawasaki H, Suzuki K, Inoe Y (1990) Organization and primary sequence of multiple genes coding for the Apopolysialoglycoproteins of rainbow trout. J Mol Biol 211:35–48

Tsujimoto Y, Suzuki Y (1979) Structural analysis of the fibroin gene at the 5′ end and its surrounding region. Cell 16:425–436

Weatherall DJ, Clegg JB (1979) Recent developments in the molecular genetics of human hemoglobin. Cell 16:467–479

Wellman SE, Case ST (1989) Disassembly and reassembly in vitro of complexes of secretory proteins from *Chironomus tentans* salivary glands. J Biol Chem 264:10878–10883

Wellman SE, Hamodrakas SJ, Kamitsos EI, Case ST (1992) Secondary structure of synthetic peptides derived from the repeating unit of a giant secretory protein from *Chironomus tentans*. Biochim Biophys Acta 1121:279–285

Wieslander L, Höög C, Höög J-O, Jörnvall H, Lendahl U, Daneholt B (1984) Conserved and nonconserved structures in the secretory proteins encoded in the Balbiani ring genes of *Chironomus tentans*. J Mol Evol 20:304–312

Wieslander L, Paulsson G (1992) Sequence organization of the Balbiani ring 2.1 gene in *Chironomus tentans*. Proc Natl Acad Sci USA 89:4578–4585

Wyss C (1982) Chironomus tentans epithelial cell lines sensitive to ecdysteroids, juvenile hormone, insuline and heat shock. Exp Cell Res 319:297–307

Xu M, Lewis RV (1990) Structure of a protein superfiber: spider dragline silk. Proc Natl Acad Sci USA 87:7120–7124

Yamada Y, Avvedimento VE, Mudryj M, Ohkubo H, Vogeli G, Irani M, Pastan I, de Crombrugghe B (1980) The collagen gene: evidence for its evolutionary assembly by amplification of a DNA segment containing an exon of 54 bp. Cell 22:887–892