

Molecular organization of the maternal effect region of the *Shaker* complex of *Drosophila*: characterization of an I_A channel transcript with homology to vertebrate Na^+ channel

A. Baumann, I. Krah-Jentgens, R. Müller, F. Müller-Holtkamp, R. Seidel, N. Kecskemethy, J. Casal¹, A. Ferrus¹ and O. Pongs

Ruhr-Universität Bochum, Lehrstuhl für Biochemie, 4630 Bochum, FRG, and ¹Instituto Cajal, Consejo Superior de Investigaciones Científicas, Madrid, España

Communicated by H. Thoenen

We have cloned 215-kb DNA containing the maternal effect region (ME) of the *Shaker* gene complex (ShC) at 16F of the *Drosophila* X chromosome. Five translocation and deletion breakpoints have been mapped on the cloned DNA allowing a correlation of the genetic map to transcription units. The ME region spans ~100 kb. The genetic behavior of this region correlates with the occurrence of maternal RNAs in this part of the ShC. Two transcripts have been identified in the vicinity of chromosomal rearrangements which cause a *Sh* phenotype. These are a 4.5-kb transcript interrupted by T(X;2)B27 and a 2-kb transcript interrupted by T(X;3)Sh^{LC} and T(X;Y)W32. The latter transcript is derived from a primary transcript which spans >65 kb genomic DNA. The cDNA-sequencing data show that this *Shaker* (I_A channel) gene can encode a protein of ~35 kd with three α -helical membrane-spanning sequences near its carboxyl terminus. These have a striking homology with membrane-spanning sequences of the vertebrate Na^+ channel.

Key words: *Drosophila/Shaker* locus/ K^+ channel/ Na^+ channel

Introduction

The molecular analysis of genes whose mutations cause abnormal membrane physiology provides an important mean to approach the study of ion channels. Electrophysiological as well as pharmacological studies sustain the hypothesis that the *Shaker* (*Sh*) (X;57.6) locus of *Drosophila* might encode components of a channel that mediates a fast, voltage-dependent K^+ current (I_A). *Sh* mutants exhibit action potentials with delayed repolarization (Jan *et al.*, 1977; Tanouye *et al.*, 1981; Tanouye and Ferrus, 1985) in an identified neuron (Koto *et al.*, 1981) as well as altered I_A in voltage-clamped muscle cells (Salkoff and Wyman, 1981; Wu *et al.*, 1983; Wu and Haugland, 1986). The I_A defects include altered closing kinetics in some alleles (Salkoff and Wyman, 1981). The molecular components of K^+ channels have not yet been identified in any organism. The heterogeneity of K^+ currents (Adams *et al.*, 1980) and the absence of ligands with the required specificity and binding constants render the standard biochemical approaches (Popot and Changeux, 1984) unproductive. The available data indicate that the *Sh* locus controls the properties of the I_A current and thereby action potential duration.

The genetic analysis indicates that the *Sh* locus is a gene complex. Several lethal complementation groups can be included within the *Sh* complex (*ShC*) (Tanouye *et al.*, 1986; A. Ferrus *et al.*, in preparation) together with the viable alleles where the

physiological studies were performed. Although the physiological properties of the lethal mutants are not yet known, they seem to be related to the viable mutants on the basis of genetic behavior, nature of the mutations and other phenotypic traits (A. Ferrus *et al.*, 1987). Several chromosomal rearrangements with cytogenetically separable breakpoints cause typical *Sh* physiology. The *ShC* is limited, at present, by the most proximal [Df(1)S4010] and the most distal [T(X;3)JC153] breakpoints that do not show *Sh* phenotypes. The complex can be subdivided into a haplolethal and a maternal effect region. The haplolethal region is limited by the breakpoints of T(X;Y)V7 and T(X;Y)W32 while the maternal effect region is limited by those of T(X;Y)B55 and T(X;3)JC153. Here we describe the chromosomal walk along 215 kb covering entirely the maternal effect region. Four breakpoints from chromosomal rearrangement showing *Sh* phenotypes have been mapped. This alignment allows the transcripts encoded in the cloned *ShC* DNA to be correlated with genetic components of the *ShC*. The partial structure of a 2-kb transcript interrupted by the breakpoints of T(X;3)Sh^{LC} and T(X;Y)W32 has been determined. This transcript apparently encodes an I_A channel protein. The membrane-spanning sequences of this protein have a striking homology with membrane-spanning sequences of the vertebrate Na^+ channel (Noda *et al.*, 1986).

Results

A cDNA clone (adm 135H4), which hybridized *in situ* at 16E/F, had been obtained from D. Hogness' laboratory (Wolfner, 1979). This location was near the sites of X-chromosomal breakpoints [such as T(X;Y)B55 and T(X;Y)W32] causing *Sh* mutant phenotypes (Tanouye *et al.*, 1981). Therefore, this cDNA clone was used to screen genomic DNA libraries to isolate a continuous set of overlapping clones. The chromosomal walk, as summarized in Figure 1, represents a continuous region of 215 kb extending 94 kb distal to the adm 135H4 site and 121 kb proximal of it. Both λ -phage and cosmid libraries were used (see Materials and methods). The *Drosophila* DNA contained in these libraries originated either from OregonR (E clones and cos 401) or CantonS (M clones) fly stocks. We have only detected a few restriction site polymorphisms between DNA of different fly strains, most notably in the -10, +15, +60 and +100 regions (data not shown). Clones isolated from the -60 to -65 region differed by an insertion of 2 kb (as indicated in Figure 1), which was detected in clones of CantonS origin (M7a, M7i), but not of OregonR origin (E44). This insertion was included in the restriction map shown in Figure 1. In this paper, we describe a set of clones which include the maternal effect region of the *ShC*. This region is defined by the breakpoints of T(X;3)JC153 and T(X;Y)W32. These and three other chromosomal breakpoints of genetically mapped rearrangements were located on the physical map from the *ShC* (Figures 2–4).

Mapping the breakpoints of chromosomal rearrangements

The distal limit of the maternal effect region of the *ShC* is defined by the distal breakpoint of the insertional T(X;3)JC153. In

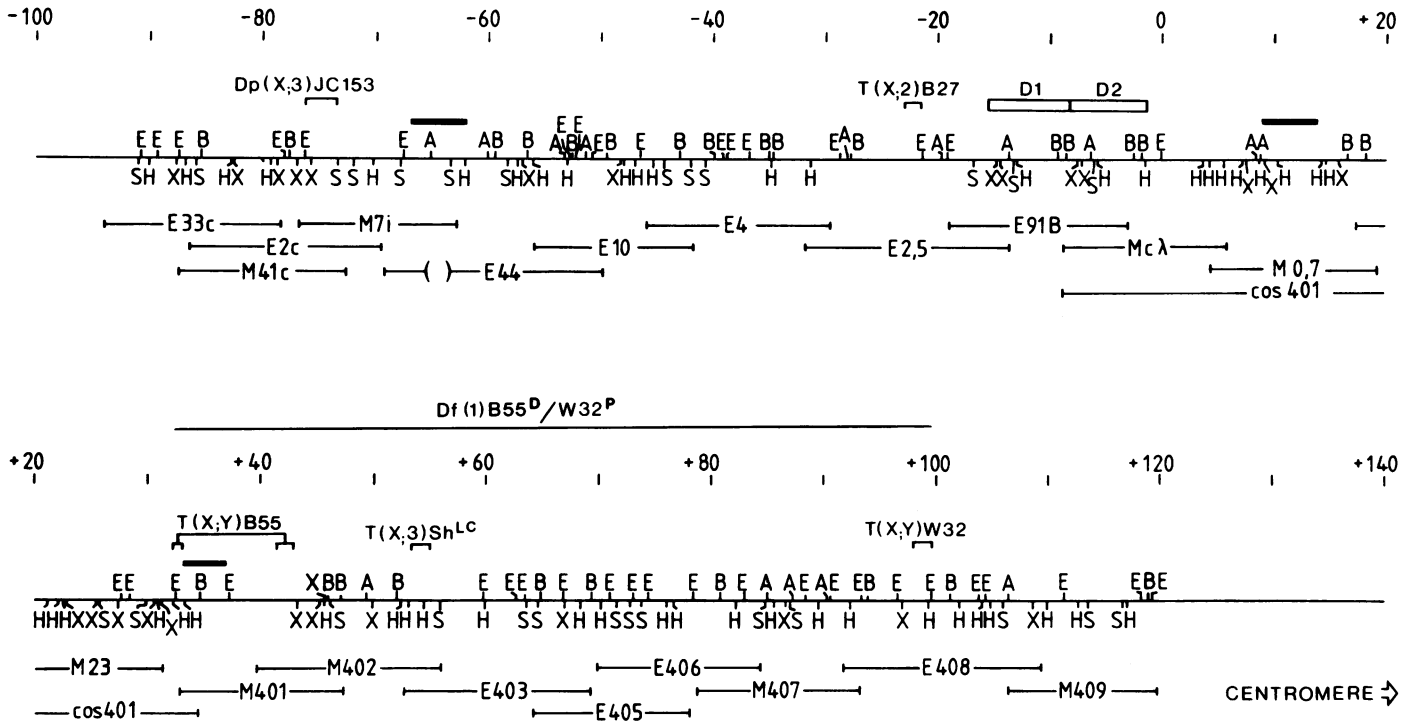


Fig. 1. Restriction map of the cloned *Shaker* gene complex. The distance is measured in kb from the start site of the chromosomal walk. The positive direction is towards the centromere, negative towards the telomere. Only the most representative clones for each interval are indicated below the map. The restriction maps were checked in all cases using restriction fragments to hybridize back to Southern blots. *XbaI* sites were not mapped in the region -40 to -20. Breakpoints of duplications and translocations are indicated above the map by brackets. DNA, which is deleted in *Df(1)B55^D/W32^P*, is indicated above the map by a solid line. Thick bars above the map indicate repetitive DNA in the cloned region. Part of the repetitive DNA was absent from clone E44 as indicated by the parenthesis. Boxed D1 and D2 indicate a tandem duplication. *EcoRI* = E, *BamHI* = B, *HindIII* = H, *SstI* = S, *ApaI* = A, *XbaI* = X.

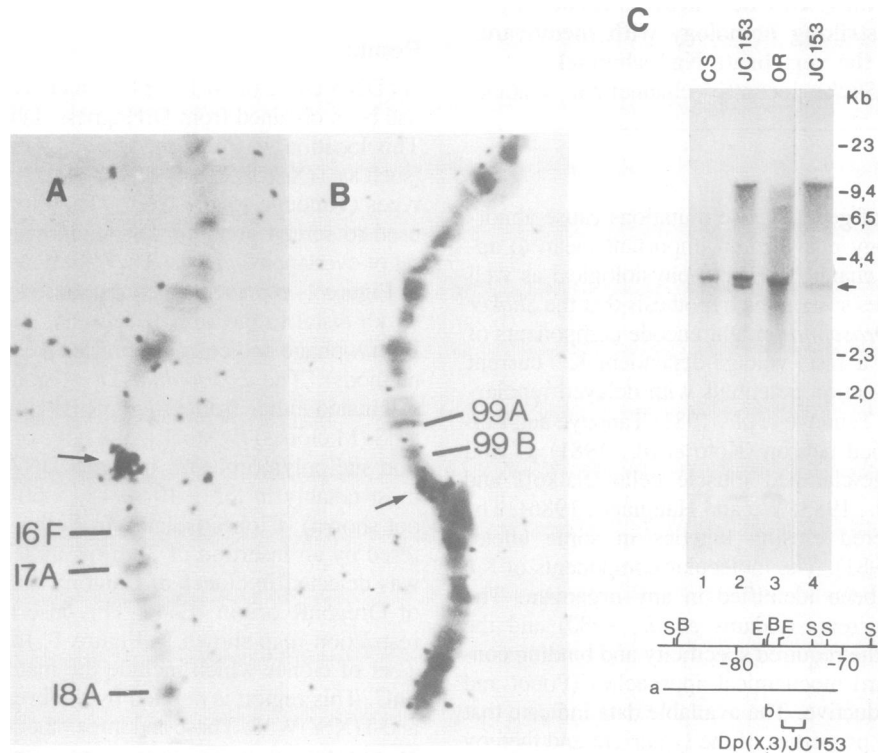


Fig. 2. Mapping the distal limit of *ShC*. (A) and (B) *in situ* hybridization of ³H-labeled M41c DNA to polytene chromosomes of *Dp(X;3)JC153*. Only the X chromosome (A) and the third chromosome (B) are shown. Arrowheads indicate hybridization sites. (C) Genomic Southern blots of DNA from male OregonR (OR), CantonS (CS), and *Dp(X;3)JC153* (JC153) flies digested with *SstI* and hybridized with labeled DNA probes. Probes a (lanes 1-3) and b (lane 4) indicated by corresponding lines below the restriction map were derived from clone E2c. *EcoRI* = E, *BamHI* = B, *SstI* = S. The arrow points to the altered *SstI* fragment of *Dp(X;3)JC153* in lanes 2 and 4. Bracket at bottom indicates the limits within which the chromosomal rearrangement of *Dp(X;3)JC153* was mapped. Mol. wt standards in kb are indicated at right.

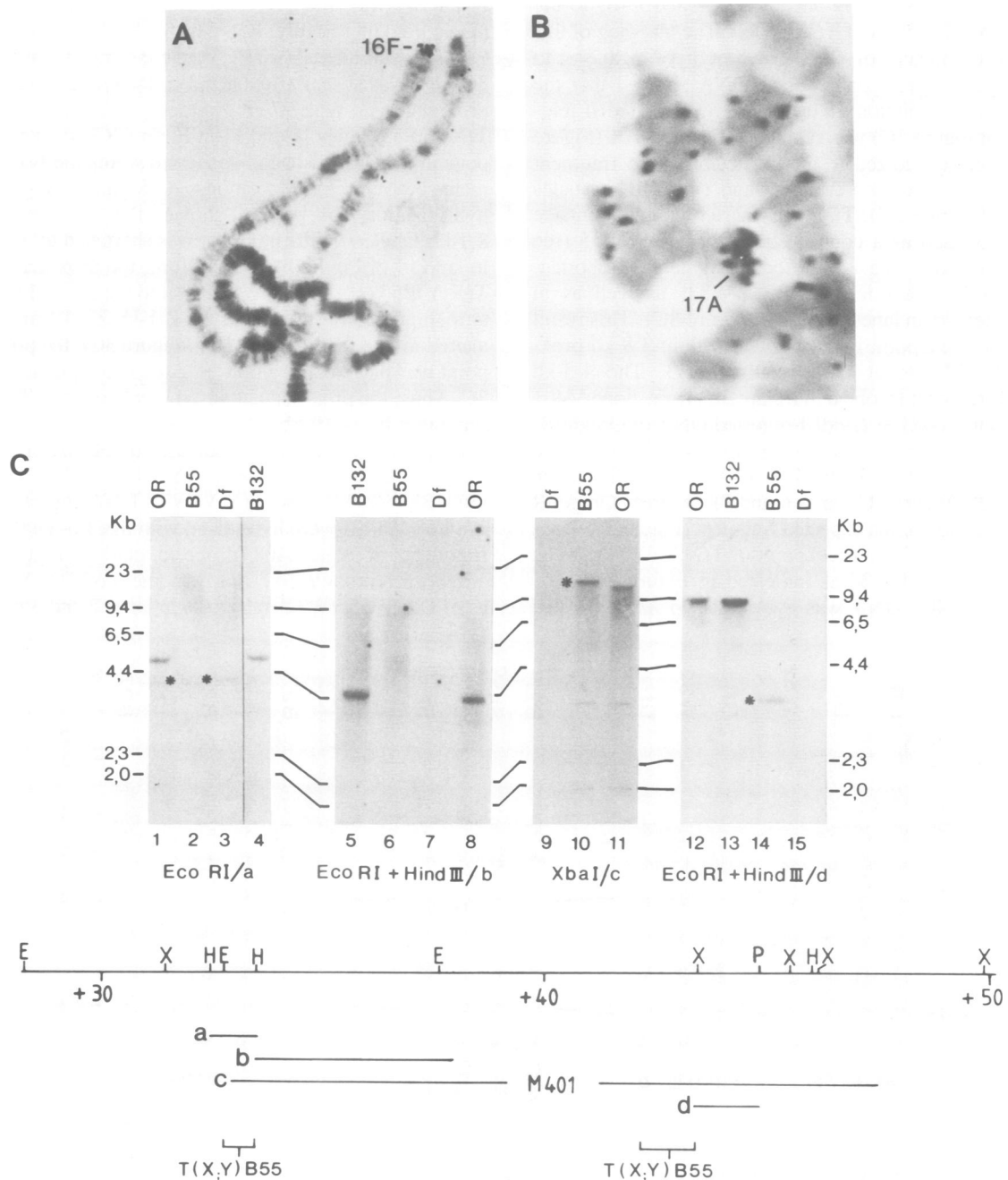


Fig. 3. Mapping the breakpoint of T(X;Y)B55. (A) *In situ* hybridization of ^3H -nick-translated clone M0.7 to the distal element of T(X;Y)B55. (B) *In situ* hybridization of clone E403 to the proximal element of T(X;Y)B55. Arrowheads indicate hybridization sites. (C) Genomic Southern blots of DNA of OregonR (OR), T(X;Y)B55 (B55), T(X;Y)B132 (B132) and Df(1)B55^D/W32^P (Df) males digested with *Eco*RI (lanes 1–4), with *Eco*RI + *Hind*III (lanes 5–8, 12–15) or with *Xba*I (lanes 9–11) were hybridized with labeled DNA probes derived from clone cos401 (probe a) or from clone M401 (probes b, c and d). Probes are indicated below the restriction map by corresponding lines. Lanes 1–4 were hybridized with probe a, lanes 5–8 with probe b, lanes 9–11 with probe c, and lanes 12–15 with probe d. Brackets below the map indicate the limits within which the breakpoints of T(X;Y)B55 were mapped. Asterisks in lanes 2, 3, 10 and 14 mark restriction fragments which were altered by the rearrangements T(X;Y)B55 or Df(1)B55^D/W32^P. Mol. wt standards in kb are indicated at left and at right. *Eco*RI = E, *Xba*I = X, *Hind*III = H, *Pst*I = P. Not all the *Pst*I sites are shown.

the polytene chromosomes the deficiency element of the X chromosome can be distinguished from the duplication element located in the third chromosome. *In situ* hybridization located the distal breakpoint of Dp(X;3)JC153 within clone M41c as shown in Figure 2A,B. The hybridization with E2c DNA (probe a in Figure 2C) showed that *Sst*I-digested CantonS and OregonR DNA had the expected fragments. Dp(X;3)JC153 had an additional 3.8-kb *Sst*I fragment (lane 2 in Figure 2C), which was

also detected with a 2.8-kb *Eco*RI/*Sst*I fragment of E2c DNA (probe b, lane 4 in Figure 2C). Therefore, the distal limit of the ShC is located at –76 to –73 on the map in Figure 1.

Df(1)B55^D/W32^P is an aneuploid generated by the distal element of T(X;Y)B55 and the proximal one of T(X;Y)W32. Male flies carrying this deletion are viable and have an extreme *Sh* phenotype which places a prominent part of the ShC at 16F3–4 (Tanouye *et al.*, 1981). Figure 3A,B illustrates *in situ* hybridiza-

tions of ^3H -labelled DNA from phages MO.7 and E403 to the distal and proximal elements of T(X;Y)B55. The mapping of this rearrangement by means of Southern blot hybridizations to restriction enzyme digests of genomic DNA showed that the breakpoint includes a deletion. In detail: the 0.9-kb *HindIII* fragment (probe a in Figure 3C) was subcloned from cos401 (Figure 1). This subclone hybridized to the expected 4.8-kb fragment of *EcoRI*-digested DNA from OregonR and T(X;Y)B132 males (lanes 1 and 4 of Figure 3C). T(X;Y)B132, a breakpoint outside of the ShC, was used as a control strain for T(X;Y)B55 (see Materials and methods). The 4.8-kb *EcoRI* fragment was shortened to a 3.8-kb fragment in Df(1)B55^D/W32^P as well as in T(X;Y)B55 (asterisks in lanes 2 and 3 in Figure 3C). This result placed the distal breakpoint in T(X;Y)B55 within 3.8 kb proximal to the *EcoRI* site at map position +33. The 4.3-kb *HindIII/EcoRI* fragment (probe b in Figure 3C) from clone M401 was hybridized to *EcoRI* + *HindIII*-digested DNA of OregonR, T(X;Y)B132, T(X;Y)B55, and Df(1)B55^D/W32^P males. No hybridization was detected in T(X;Y)B55 as well as in Df(1)B55^D/W32^P (Figure 3C, lanes 6 and 7), whereas OregonR and T(X;Y)B132 showed the expected fragment (Figure 3C, lanes 5, 8). This located the distal breakpoint of T(X;Y)B55 between map positions +33 and +34 as indicated in Figure 3C.

Subsequently, M401 DNA was hybridized to *XbaI*-digested

DNA of OregonR, T(X;Y)B55 and Df(1)B55^D/W32^P males (lanes 9–11 in Figure 3C). The 12-kb *XbaI* fragment in wild-type flies was altered in T(X;Y)B55 (asterisk in lane 11 of Figure 3C) and was absent as expected in Df(1)B55^D/W32^P. Since the proximal wild-type *XbaI* fragment of 2.1 kb was present in T(X;Y)B55 (lanes 10 and 11 in Figure 3C), the proximal breakpoint in T(X;Y)B55 had to be located within the *EcoRI/XbaI* fragment between map positions +38 and +43. Finally, the hybridization with probe d showed that the relevant wild-type 8.1-kb *EcoRI/HindIII* fragment was shortened to 3.7 kb (asterisk in lane 14 of Figure 3C). Therefore, the proximal breakpoint of T(X;Y)B55 is at most 3.7 kb distal to the *HindIII* restriction site at map position +46. Since the 2.1-kb *XbaI* fragment was not altered in T(X;Y)B55 (lane 10 in Figure 3C), the proximal breakpoint was placed at position +42 to +43 as indicated in Figure 3C. Thus, the proximal and distal breakpoints of T(X;Y)B55 are separated by ~10 kb.

In situ hybridization had indicated that DNA of E408 covered the T(X;Y)W32 breakpoint. Southern blots of OregonR, Df(1)B55^D/W32^P, and T(X;Y)W32 DNA digested either with *XbaI* or *HindIII* were hybridized with the 3.0- and 3.2-kb *EcoRI* fragments of clone E408 (probes a and b in Figure 4A). Probe a hybridized to a 9.2-kb *XbaI* fragment of T(X;Y)W32 as well as of Df(1)B55^D/W32^P (asterisks in lanes 2 and 3 of Figure 4A).

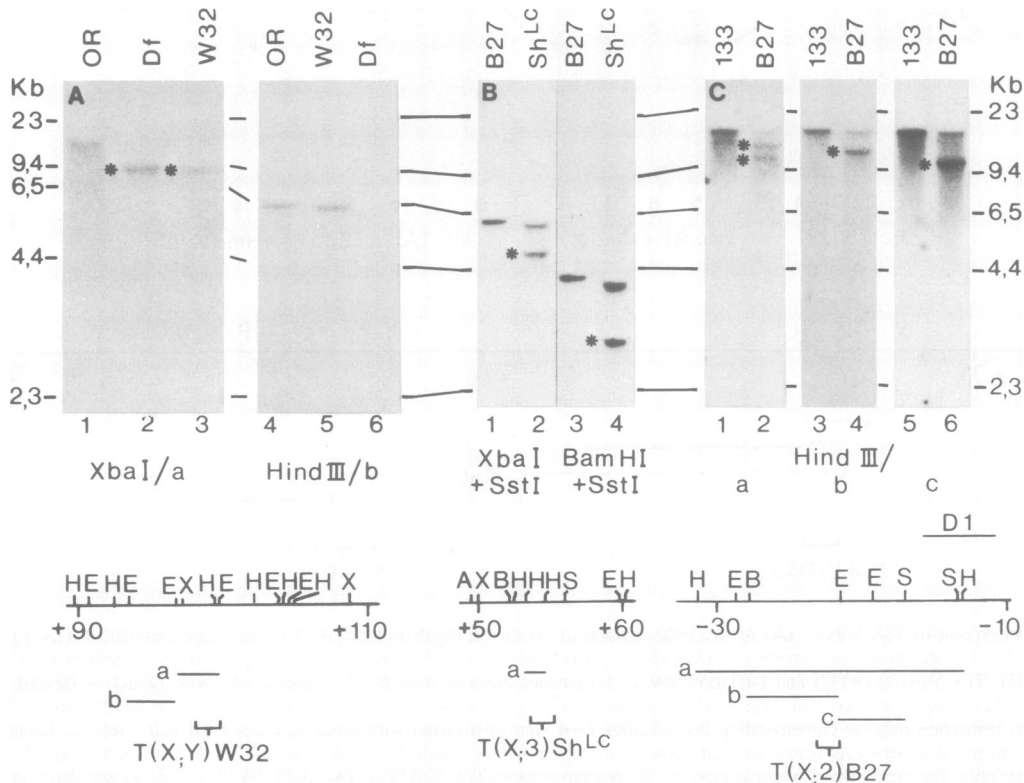


Fig. 4. Mapping the breakpoints of T(X;Y)W32, T(X;3)Sh^{LC}, T(X;2)B27. (A) Genomic Southern blots of DNA of male OregonR (OR), T(X;Y)W32 (W32) and Df(1)B55^D/W32^P (Df) flies digested with *XbaI* (lanes 1–3) or with *HindIII* (lanes 4–6) and hybridized with labeled DNA probes derived from clone E408. Probes a and b are indicated by corresponding lines below the restriction map. Lanes 4–6 were hybridized with probe a, lanes 1–3 with probe b. Asterisks mark altered restriction fragments in T(X;Y)W32 and in Df(1)B55^D/W32^P. Mol. wt standards in kb are indicated at left. *EcoRI* = E; *HindIII* = H; *XbaI* = X. (B) Genomic Southern blots of DNA of T(X;2)B27/FM6 (B27) and T(X;3)Sh^{LC}/FM6 (Sh^{LC}) flies digested with *XbaI* + *SstI* (lanes 1 and 2) or with *BamHI* + *SstI* (lanes 3 and 4). Hybridization was with the DNA probe indicated by a corresponding line below the restriction map. *XbaI* = X, *EcoRI* = E, *BamHI* = B, *HindIII* = H, *SstI* = S. Asterisks indicate altered restriction fragments in T(X;3)Sh^{LC}/FM6. Bracket at bottom indicates the limits within which the chromosomal rearrangement of T(X;3)Sh^{LC} was mapped. (C) Genomic Southern blots of DNA of *Sh*^{KS133} (133) and T(X;2)B27/FM6 (B27) flies digested with *HindIII* and hybridized with labelled DNA probes indicated by corresponding lines below the restriction map. Probes a (lanes 1 and 2), b (lanes 3 and 4), and c (lanes 5 and 6) were derived from clone E 2.5. Mol. wt standards in kb are at right. *EcoRI* = E, *BamHI* = B, *SstI* = S, *HindIII* = H. Asterisks in lanes 2, 4 and 6 indicate altered restriction fragments in T(X;2)B27/FM6. Bracket at bottom indicates the limits within which the chromosomal rearrangement of T(X;2)B27 was mapped.

By contrast, this fragment had a size of 12 kb in wild-type flies (lane 1 in Figure 4A). Since probe d did not hybridize to Df(1)B55^D/W32^P (lane 6 in Figure 4A), the T(X;Y)W32 breakpoint was located between map position +98 and the *EcoRI* site of probe a at +99.5 as shown in Figure 4A.

Two other breakpoints within the maternal effect region cause *Sh* phenotype: T(X;3)*Sh*^{LC} and T(X;2)B27 (Tanouye *et al.*, 1981). The breakpoint of T(X;3)*Sh*^{LC} appeared to be located ~10 kb proximal to that of T(X;Y)B55, because *in situ* hybridization to polytene chromosomes showed that M401 DNA hybridized distal and E403 DNA mainly proximal to the breakpoint in T(X;3)*Sh*^{LC} (data not shown). T(X;3)*Sh*^{LC}/FM6 DNA was

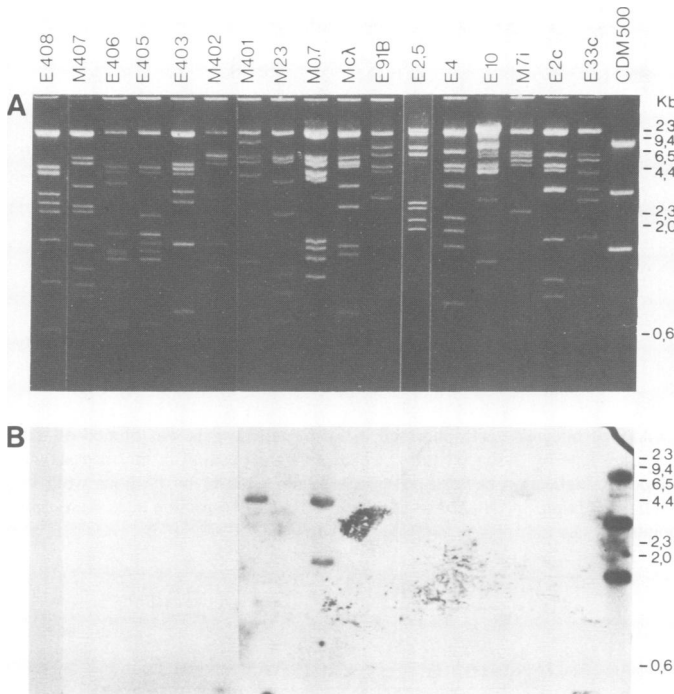


Fig. 5. Repetitive sequences in the ShC. Labeled *Drosophila* DNA hybridized to clones from the ShC region. The upper panel (A) shows the gel stained with ethidium bromide, the lower panel (B) shows the autoradiograph of the Southern blot. Mol. wt standards in kb are at right. All clones were cut with *EcoRI* + *HindIII* except clone E 2.5, which was cut with *EcoRI* + *SstI*. Clone CDM 500 contains histone gene DNA. It was cut with *HindIII* + *SstI*.

digested either with *XbaI* + *SstI* (lane 2 in Figure 4B) or with *BamHI* + *SstI* (lane 4 in Figure 4B). T(X;2)B27/FM6 DNA was used for control in the Southern blot hybridizations (lanes 1 and 3 in Figure 4B). The 1.4-kb *HindIII* fragment of E403 DNA (probe a in Figure 4B) hybridized to the expected 5.9-kb *XbaI/SstI* and 3.9-kb *BamHI/SstI* fragments in accordance with our restriction map. The T(X;3)*Sh*^{LC} breakpoint additionally generated *XbaI/SstI* and *BamHI/SstI* fragments, both 1.1 kb shorter than the wild-type fragments. Therefore, the breakpoint of T(X;3)*Sh*^{LC} was located within the *BamHI/SstI* fragment at map position +53 to +55 as shown in Figure 4B.

The rearrangement T(X;2)B27 was induced in a *Sh*^{KSI33}-bearing chromosome and was detected because of the modified *Sh* phenotype (Tanouye *et al.*, 1981). The breakpoint appeared to reside near the distal limit of the maternal effect region based on cytogenetic observations. Since E4 DNA hybridized *in situ* very close to the breakpoint of T(X;2)B27, labelled E 2.5 DNA was hybridized to genomic Southern blots of *HindIII*-digested DNA of *Sh*^{KSI33} (used as control strain) and T(X;2)B27/FM6 females (Figure 4C, lanes 1 and 2). Within the 18-kb *HindIII* fragment of E 2.5 DNA (lane 2 in Figure 4C) the T(X;2)B27 breakpoint generates two new *HindIII* fragments (9.5 and 11.5 kb) which were not present in *Sh*^{KSI33}. The 6.5-kb *BamHI/EcoRI* fragment of E 2.5 DNA (probe b in Figure 4C) hybridized to the 11.5-kb *HindIII* fragment and probe c to the 9.5-kb *HindIII* fragment of T(X;2)B27/FM6 (lanes 4 and 7 in Figure 4C). Therefore, the T(X;2)B27 breakpoint was placed at position -22 to -21 (Figure 4C).

Repetitive sequences

Three repetitive sequences and one duplication were encountered during the chromosomal walk. We used hybridization with genomic OregonR DNA to Southern blots of cloned ShC DNA as a diagnostic test to identify repetitive sequences (Pirota *et al.*, 1983). Cloned *Drosophila* histone genes (CDM 500, Goldberg, 1979) were used as reference DNA in order to have an estimate for sequences which occur ~100 times in the *Drosophila* genome. Several restriction fragments gave a signal corresponding to ~20–50 copies per genome (Figure 5). These middle repetitive sequences were located at map positions -67 to -62 (clone M71) +9 to +14 (clone M0.7), and at +33 to +37 (clone M401). The repetitive sequences at +33 to +37 are located within the deletion of T(X;Y)B55 (see Figure 1). Sequences within map positions -2 to -17 were exactly duplicated. The repetitive element 297 (Potter *et al.*, 1979) was inserted into the duplica-

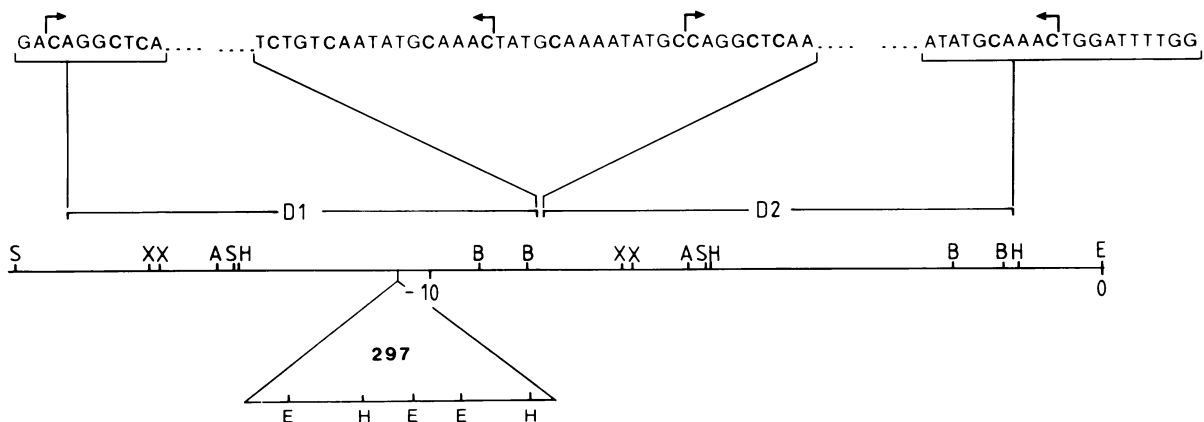


Fig. 6. Arrangement of a duplication (D1, D2) within the ShC. *EcoRI* = E, *HindIII* = H, *BamHI* = B, *ApaI* = A, *SstI* = S, *XbaI* = X. The limits of the duplication are indicated by arrows on top of the respective DNA sequence. The triangle indicates the position of the insertion of the 297 repetitive element into CantonS DNA.

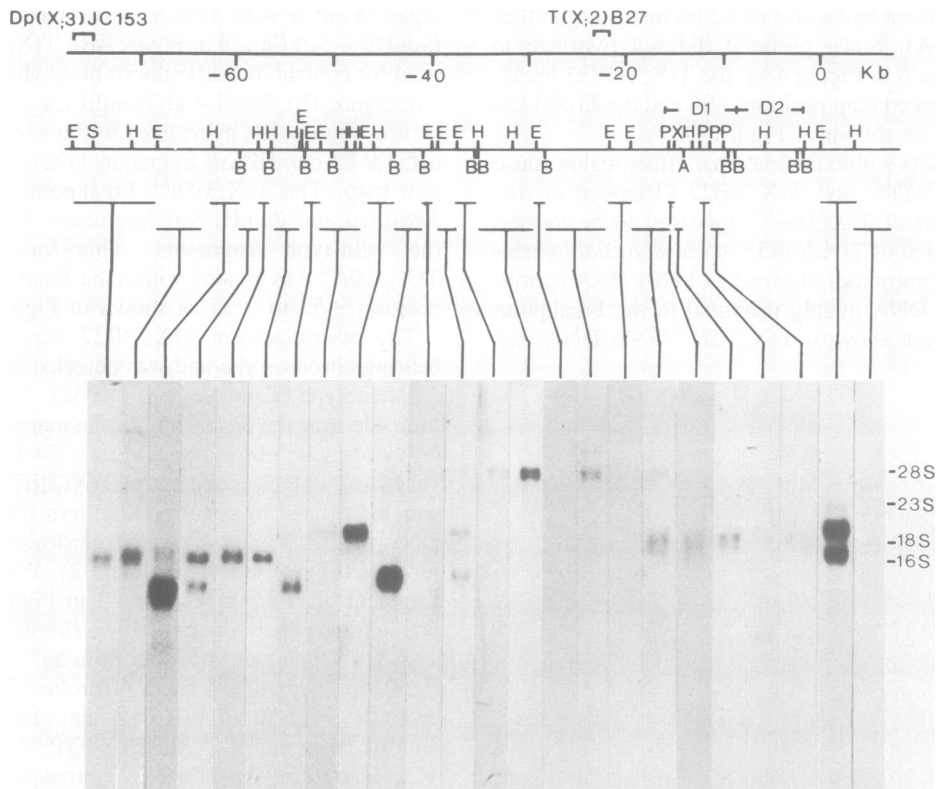


Fig. 7. Transcripts from the maternal effect region of the ShC. Poly(A)⁺ RNA of wild-type oocytes electrophoresed in formaldehyde gels was blotted on nylon filters. Hybridizations were done with labelled subclones as indicated on top of each lane by T-shaped lines, which give the position in the map shown at the top. rRNAs used as size markers are indicated at right. D1/D2 indicates the tandem duplication between position -2 and -17. Hybridizations were done only with labeled subclones of D1. *EcoRI* = E, *HindIII* = H, *SstI* = S, *BamHI* = B, *ApaI* = A, *XbaI* = X, *PstI* = P. *PstI* restriction sites were only mapped for the D1 region. Brackets indicate Dp(X;3)JC153 and T(X;2)B27 breakpoints.

tion of CantonS DNA as illustrated in Figure 6. The duplication was absent in the wild-type strain Berlin (data not shown). The limits of this duplication were determined by sequencing the corresponding genomic DNA (Figure 6). The duplicated sequences end with the decamer ATATGCAAAC. Fragments from this sequence are directly repeated between the duplication in the form of the octamer TATGCAAA, followed by the hexamer ATATGC.

Transcripts from the cloned Shaker gene complex

The maternal effect region of the ShC contains a number of homozygous lethal complementation groups, which have been genetically mapped between the distal breakpoints of T(X;3)JC153 and T(X;Y)B55 (A.Ferrus *et al.*, in preparation). Therefore, the maternal region is located between map positions -76 and +32 (Figure 1). The dearth of repetitive sequences in this region permitted us to investigate the transcriptional activity of this region by systematically hybridizing subcloned ShC-DNA to Northern blots of oocyte poly(A)⁺ RNA as shown in Figure 7. Nearly every subclone in the -76 to +5 region hybridized to oocyte RNA. Subclones in the +5 to +34 region did not hybridize to oocyte RNA or RNA of any other developmental stage. There is a good correlation between the occurrence of oocyte transcripts within defined region of the physical map of the ShC and the genetically mapped maternal effect region. The rearrangement T(X;2)B27 apparently occurred within a transcription unit coding for a ~4.5-kb poly(A)⁺ RNA.

Subclones from that part of ShC where we have mapped T(X;Y)B55, T(X;3)Sh^{LC} and T(X;Y)W32 breakpoints (map positions +32 to +110 in Figure 1), did not hybridize to Northern blots of oocyte RNA, but instead to poly(A)⁺ RNA deriv-

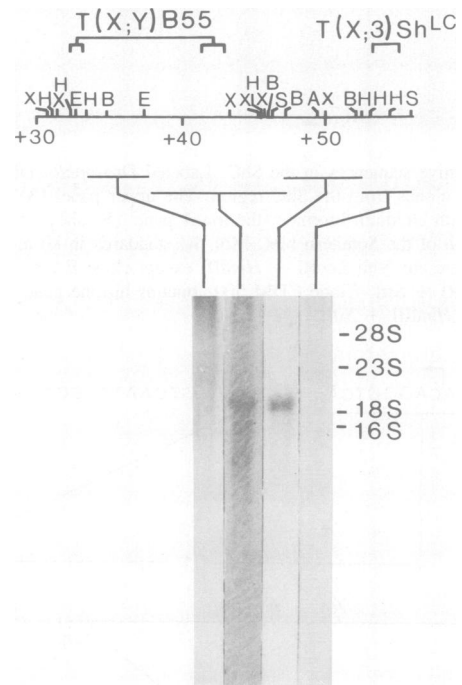


Fig. 8. Location of transcript between T(X;Y)B55 and T(X;3)Sh^{LC}. Poly(A)⁺ RNA of wild-type embryos electrophoresed in formaldehyde gels was blotted on nylon filters. Hybridizations were done with labeled DNA probes as indicated by T-shaped lines on top of each lane. Map positions of these probes are shown at top. rRNAs were used as size markers. *EcoRI* = E, *HindIII* = H, *BamHI* = B, *XbaI* = X, *SstI* = S, *ApaI* = A. The breakpoints for T(X;Y)B55 and T(X;3)Sh^{LC} are indicated by brackets.

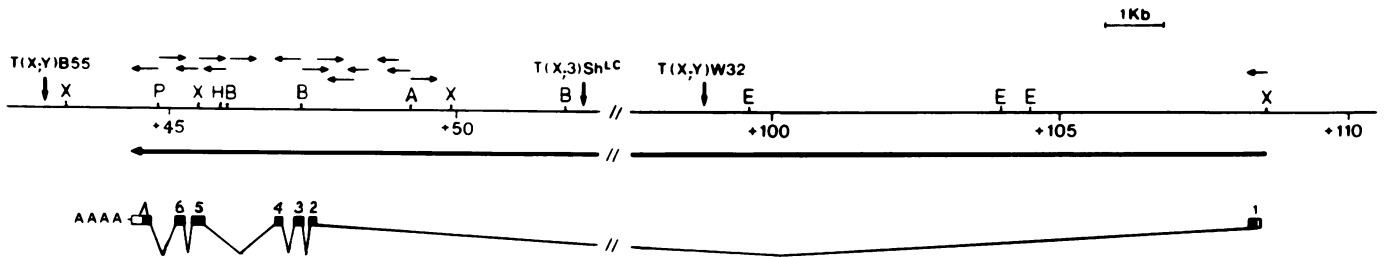


Fig. 9. Structure of the *Sh* transcription unit. The restriction map is as in Figure 1. Translocation breakpoints T(X;Y)B55, T(X;3)*Sh*^{LC} and T(X;Y)W32 are indicated by vertical arrows. The centromere is to the right. Calibration is in kb. Thin horizontal arrows on top of the restriction map indicate the genomic regions which were sequenced. Thick horizontal arrow indicates direction and extent of the primary *Sh* transcript. Intervening sequences are indicated by thin lines underneath, exons by numbered boxes. Open boxes are untranslated regions, filled boxes correspond to the 912-nucleotide open reading frame. Lengths are to scale.

ed from embryos (Figure 8). Between map positions +40 and +50 we have detected an ~2-kb-long transcript located immediately proximal to the breakpoint of T(X;Y)B55 and distal to that of T(X;3)*Sh*^{LC}. This transcript has been analyzed in detail. Further proximal (map positions +50 to +110) we have not found other transcription units so far (but see below).

With the 5-kb *Bam*HI restriction fragment of M402 DNA we have isolated a cDNA clone (gt141) (Figure 9) which hybridized to a similarly sized transcript in preparations of poly(A)⁺ embryonic RNA as in Figure 8. Both strands of the cDNA clone were sequenced, as well as 5 kb of the genomic DNA from the region where the transcript is encoded (Figure 9). The gt141 clone contained 1200 bp cDNA with a poly(dA)·(dT) tail at one end. Northern analysis with strand-specific probes from gt141 cDNA revealed that the poly(dA) tail was at the 3' end of this *Sh* transcript (data not shown). Comparison of the cDNA sequence with the corresponding genomic region only revealed one nucleotide difference outside the cDNA's open reading frame. Six points of non-colinearity were revealed, indicative of six intervening sequences in the primary transcript (Figure 9). The splice junctions for these intervening sequences are located at positions 240, 364, 533, 630, 826 and 948 (Figure 10). The 3' end of the *Sh* transcript (exon 7) was located within the *Eco*RI/*Xba*I fragment of M401 DNA in the immediate vicinity or at the proximal breakpoint of T(X;Y)B55. The 5' end of the *Sh* transcript (exon 1) was not derived from M401 or M402 DNA. Instead, the first exon was found to be encoded in an *Eco*RI/*Xba*I fragment of M409 DNA, i.e. 10 kb proximal to the breakpoint of T(X;Y)W32 (Figure 9). Therefore, the gt141 cDNA encompasses the translocation breakpoints T(X;3)*Sh*^{LC} and T(X;Y)W32. The corresponding chromosomal rearrangements interrupt joining of exon 1 to exon 2 which are separated by an intervening sequence of 60 kb.

The most extensive open reading frame in the gt141 cDNA has 912 bp (nucleotides 194–1105). Translation of the 912-bp open reading frame from the position of its first AUG codon (194) would produce a protein with 304 amino acids and a molecular mass of 34 949 daltons (Figure 10). There is one additional in-frame AUG codon located very near (994) the 3' end. Sequences characteristic of integral membrane proteins have been found near the 3' end of the deduced protein sequence (Figure 11). The hydrophobicity profile (Figure 11A) indicates three potentially bilayer-spanning regions, which have the general properties expected for intramembraneous sequences of proteins which traverse the bilayer several times (Brandl and Deber, 1986) as outlined by a model for the structure of the protein (Figure 11B). This observation is supported by the homologies between the *Drosophila* membrane-spanning helices and the corresponding se-

quences of membrane-spanning helices of the vertebrate voltage-dependent Na⁺ channel (Figure 12). The homologies are 43–59% at the nucleotide and 53–70% at the amino acid level. The first membrane-spanning sequence (736–796) is homologous to the S1 segments of *E. electricus* (Noda *et al.*, 1984) and rat brain Na⁺ channel (Noda *et al.*, 1986) (Figure 12A). Accordingly, the amino acid sequences can be aligned to give a 53–62% homology. This is similar to the amino acid homology amongst the various S1 segments of the Na⁺ channels. The second membrane-spanning sequence (899–958) of the *Sh* gene product has an even higher homology to the second membrane-spanning sequence of the vertebrate Na⁺ channel. Twelve to 14 out of 20 amino acids of the S2 segments can be aligned with the *Sh* protein sequence (Figure 12B). The third membrane-spanning sequence (988–1054) is homologous to the S6 segment of the Na⁺ channel as outlined in Figure 12C. The homologies are 52–53% at the nucleotide and 55–59% at the amino acid level. These homologies suggest that the *Sh* gene product is a structural component of a voltage-dependent K⁺ channel in accordance with the *Sh* phenotype.

Discussion

The ShC is organized into two major regions: the maternal effect (ME) and the haplolethal (HL) regions (Tanouye *et al.*, 1981). The names of these regions describe the genetic behavior in aneuploids. Males deficient for the ME region are viable only if their mothers were at least diploid for this region. Females heterozygous for a ME deletion never give rise to viable ME null males. On the other hand, females heterozygous for a HL deletion are not viable irrespective of the genetic constitution of their progenitors. The ME and HL regions are separated by ~65 kb DNA located between T(X;Y)B55 and T(X;Y)W32. Its deletion, Df(1)B55^D/W32^P, is viable in males irrespective of the genotype of their progenitors. Therefore, it is distinguishable from the ME and HL regions by its genetic behavior.

By comparison to other cloned regions of *Drosophila* genome, the ME region does not show extensive polymorphism nor abundance of repetitive sequences. The breakpoints do not locate preferentially in these repetitive or polymorphic sites. The rearrangements show rather simple cuts except T(X;Y)B55. This rearrangement includes a 10-kb deletion with respect to wild-type DNA. The cuts have been delimited in all cases within a margin of 1–3 kb.

The ME region encodes several transcripts which were detected in oocyte poly(A)⁺ RNA. The minimum estimate from Figure 7 is five transcriptional units, each one encoding multiple RNAs. The genetic behavior of the ME region correlates with the occurrence of maternal RNAs in this part of the ShC. Most of them

```

t atcggctgatctatatacagacctatagaatcttaagtgtttcccgacttgc      52
atagccgactagatatatgctggatatcttagaattcacaaagggctgaac

cagcgaataaaaaaggaaaatccagacagatttggaaaagcgattaaaatcgcataaaaa 112
gtcgcttatttttcttttaggtctgtctaaaccttttcgctaattttatgcgattttt

gaaaaacaaaATCCCAGTGGAGATCAAAGTGCACGTGCAGAGAGAGAAAAAGTGGAGTAG 172
ctttttgtt

GTCGATGAATGGATTGCCTGTATGGCACACATCACGACGACGCACGGCAGCTTAAGCCAA 232
MetAlaHisIleThrThrThrHisGlySerLeuSerGln
      1                                10
GCGACAAGGTCTTTGCCCAAATTGAGCAGTCAAGACGAAGAAGGGGGGGCTGGTCATGGC 292
AlaThrArgSerLeuProLysLeuSerSerGlnAspGluGluGlyGlyAlaGlyHisGly
      20                                30
TTTGGTGGCGGACCGCAACACTTTGAACCCATTCTCACGATCATGATTTCTGCGAAAGA 352
PheGlyGlyGlyProGlnHisPheGluProIleProHisAspHisAspPheCysGluArg
      40                                50
GTCGTTATAAATGTAAGCGGATTAAGGTTTGGAGACACAACACTACGTACGTTAAATCAATC 412
ValValIleAsnValSerGlyLeuArgPheGluThrGlnLeuArgThrLeuAsnGlnPhe
      60                                70
CCGGACACGCTGCTTGGGGATCCAGCTCGGAGATTACGGTACTTTGACCCGCTTAGAAAT 472
ProAspThrLeuLeuGlyAspProAlaArgArgLeuArgTyrPheAspProLeuArgAsn
      80                                90
GAATATTTTTTTGACCGTAGTCGACCGAGCTTCGATGCGATTTTATACTATTATCAGAGT 532
GluTyrPhePheAspArgSerArgProSerPheAspAlaIleLeuTyrTyrGlnSer
      100                               110
GGTGGCCGACTACGGAGACCGGTCAATGTCCCTTTAGACGTATTTAGTGAAGAAATAAAA 592
GlyGlyArgLeuArgArgProValAsnValProLeuAspValPheSerGluGluIleLys
      120                               130
TTTTATGAATTAGGTGATCAAGCAATTAATAAATTCAGAGAGGATGAAGGCTTTATTA 652
PheTyrGluLeuGlyAspGlnAlaIleAsnLysPheArgGluAspGluGlyPheIleLys
      140                               150
GAGGAAGAAAGACCATTACCGGATAATGAGAAAACAGAGAAAAGTCTGGCTGCTCTTCGAG 712
GluGluGluArgProLeuProAspAsnGluLysGlnArgLysValTrpLeuLeuPheGlu
      160                               170
TATCCAGAAAGTTCGCAAGCCGACAGAGTTGTAGCCATAATTAGTGTATTTGTTATATTG 772
TyrProGluSerSerGlnAlaAlaArgValValAlaIleIleSerValPheValIleLeu
      180                               190
CTATCAATTGTTATATTTTGTCTAGAAACATTACCCGAATTTAAGCATTACAAGGTGTT 832
LeuSerIleValIlePheCysLeuGluThrLeuProGluPheLysHisTyrLysValPhe
      200                               210
AATACAACAACAAATGGCACAAAAATCGAGGAAGACGAGGTGCCTGACATCACAGATCCT 892
AsnThrThrThrAsnGlyThrLysIleGluGluAspGluValProAspIleThrAspPro
      220                               230
TTCTTCCTTATAGAAACGTTATGTATTATTTGGTTTACATTTGAACTAACTGTCAGGTTC 952
PhePheLeuIleGluThrLeuCysIleIleTrpPheThrPheGluLeuThrValArgPhe
      240                               250
CTCGCATGTCCGAACAAATTAATTTCTGCAGGATGTCATGAATGTTATCGACATAATC 1012
LeuAlaCysProAsnLysLeuAsnPheCysArgAspValMetAsnValIleAspIleIle
      260                               270
GCCATCATTCCGTACTTTATAACACTAGCGACTGTCGTTGCCGAAGAGGAGGATACGTTA 1072
AlaIleIleProTyrPheIleThrLeuAlaThrValValAlaGluGluGluAspThrLeu
      280                               290
AATCTTCCAAAAGCGCCAGTCAGTCCACAGGTATGAGATTTCTGTTTGCCCAATGCCAAC 1132
AsnLeuProLysAlaProValSerProGlnVal***
      300                               304

TACACCAAAAAATATCAATATTAATCACCCACACACGACCACACACACACATGCATATA 1192
CAGTTGAACCGATTAAGTAATAAATGTTGCCATCATCATTTATGAGTTGCCTGACATGA 1252
TAAAGAAGGGTTGCTGCCCGCCCGAAACCAAAAAAAAAAAAAA 1312

```

Fig. 10. Sequence and translation of *Sh* cDNA. Capital letters are the sequence of cDNA clone gt141. Both strands were sequenced as well as the corresponding genomic DNA (Figure 9). The 5' end of the first exon has not been mapped. Small letters indicate how the cDNA sequence continues in the sequenced genomic DNA. The splice junctions of the six introns are indicated by arrows. Polyadenylation signals, AATAAA and GATAAA are found at nucleotides 1212–1217 and at 1251–1256. The translation shown is of the 912-nucleotide open reading frame starting with the ATG at nucleotide 194. Amino acids are numbered starting with this ATG. Transmembrane spanning helices are underlined. They correspond to sequences of amino acids having positive hydrophathy indices (Figure 11A).

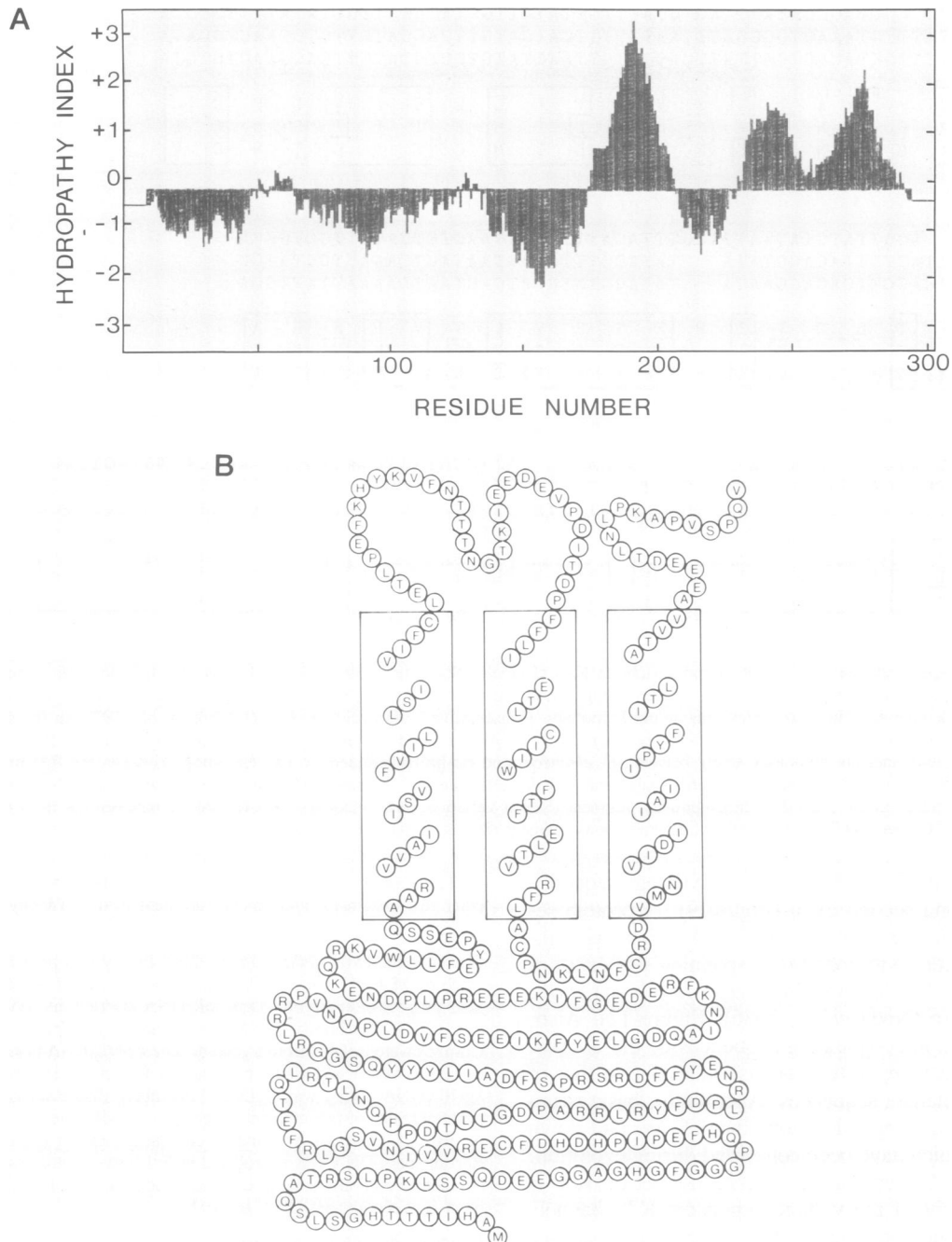


Fig. 11. A model for the structure of the *Sh* gene product. (A) Hydropathy plot of the *Sh* protein. Hydropathy indices are plotted versus amino acid residues. The plot was generated by the program of Kyte and Doolittle (1982) with a window size of 19 amino acids. Positive index values indicate hydrophobic groups, negative values hydrophilic groups. The residue numbers correspond to those given in Figure 10. (B) The transmembrane segments in the model correspond to the sequences of amino acids having positive hydropathy indices and are drawn as α -helices. One-letter symbols are used to represent the amino acid sequence as given in Figure 10.

are also expressed at other developmental stages. The T(X;2)B27 breakpoint is located within the ME region. This rearrangement was induced in a *Sh*^{KS133}-bearing chromosome and modifies the phenotype of *Sh*^{KS133}. The break locates within a region coding for a 4.5-kb transcript (Figure 7). We are presently studying this transcript in more detail.

The T(X;Y)B55–T(X;Y)W32 region encodes a very large transcription unit (65 kb at least) reminiscent of *Antp* (Schneuwly *et al.*, 1986) or *Ubx* transcription units (Hogness *et al.*, 1985). The cDNA clone gt141 contained 65–70% of the sequence of the mature transcript. The 5' end of this transcript has not been

mapped yet. The identification of this site might extend the already large primary transcript even further. The cDNA sequence (Figure 10) contains an entire open reading frame. More importantly, the first and second exon in this sequence are separated by an ~ 60 -kb intervening sequence. The intervening sequence encompasses the T(X;3)*Sh*^{LC} and T(X;Y)W32 breakpoints. T(X;3)*Sh*^{LC} as well as T(X;Y)W32 flies do not have a detectable I_A current (Tanouye and Ferrus, 1981; Wu and Haugland, 1986). Therefore, this transcription unit is a prime candidate for being the molecular basis of the *Sh* phenotype of T(X;3)*Sh*^{LC} and of T(X;Y)W32 flies. The X chromosome is

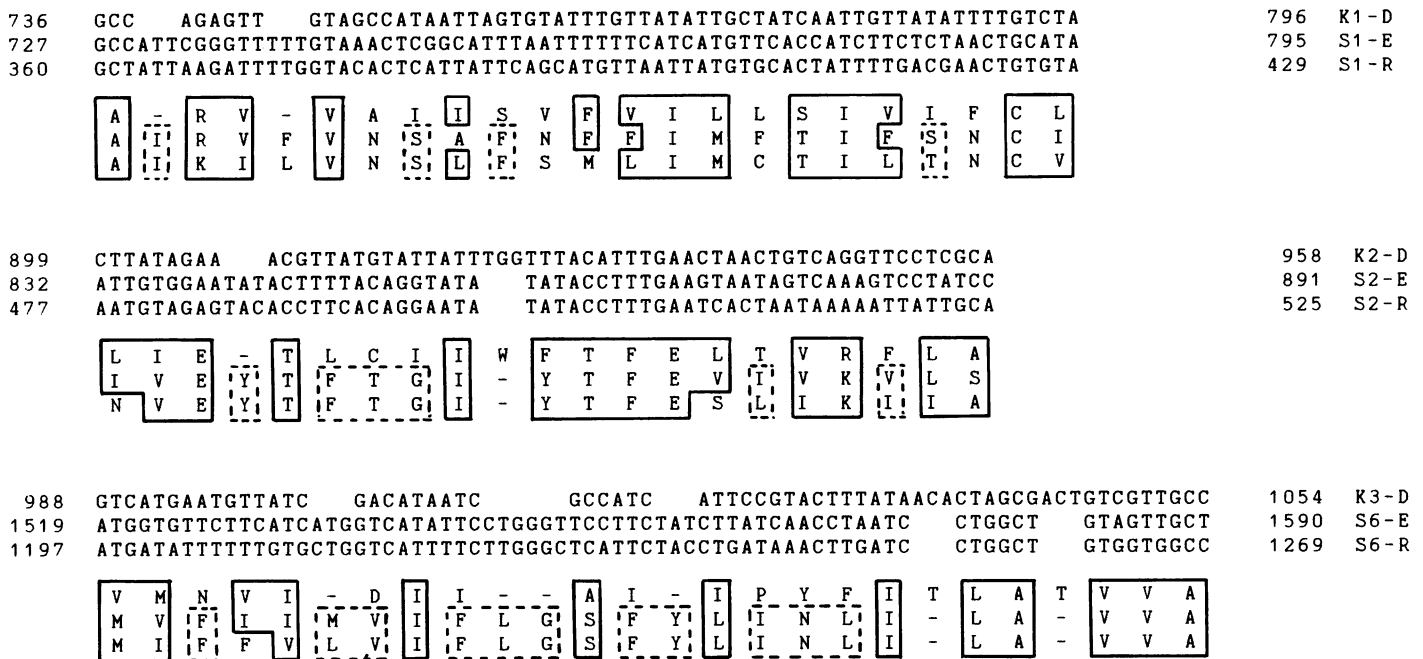


Fig. 12. Sequence homologies between the *Sh* protein and vertebrate Na⁺ channels. Nucleotide sequences of cDNAs encoding *Sh* protein (first row), *E. electricus* Na⁺ channel (second row) and rat brain Na⁺ channel I (third row) are numbered as in Figure 10 and in Noda *et al.* (1986) respectively. Gaps have been inserted into the nucleotide sequences to achieve maximum homology. Conservative amino acid residues (Dayhoff *et al.*, 1978) in the deduced amino acid sequences are boxed by solid lines to indicate homology between the *Drosophila* protein (D) and the *E. electricus* (E) and/or rat (R) protein. Amino acid residues boxed by broken lines indicate homology solely between *E. electricus* and rat protein sequences. (A) Homology between the first membrane-spanning sequence (K1) of the *Sh* protein to the first intramembranous segment (S1) of vertebrate Na⁺ channels. (B) Homology between the second membrane-spanning sequence (K2) of the *Sh* protein to the second intramembranous segment (S2) of vertebrate Na⁺ channels. (C) Homology between the third membrane-spanning sequence (K3) of the *Sh* protein to the sixth intramembranous segment (S6) of vertebrate Na⁺ channels.

broken in these mutants within the intervening sequence separating the first and second exon (Figure 9). A protein sequence is derived from the open reading frame which has striking structural similarities with membrane-spanning sequences of the voltage-dependent Na⁺ channel in *E. electricus* Noda *et al.*, 1984) and in rat brain (Noda *et al.*, 1986) (Figure 12). Also, the order of the membrane-spanning sequences of the *Sh* protein, each encoded in a separate exon, is analogous to that of the respective Na⁺ channel sequences. Apparently, basic structural features of voltage-dependent ion channels are laid down in these sequences which have been conserved during evolution. This implies that this *Sh* gene product, albeit not yet proven, is a structural component of the voltage-dependent K⁺ channel.

In spite of the striking homology of the characterized sequence of the *Sh* gene to the vertebrate Na⁺ channel a note of caution seems to be appropriate. The available data from the formal genetic (Tanouye and Ferrus, 1985) and molecular analyses indicate that the phenotype cannot be ascribed to a single transcriptional unit. The B55^D element, by itself, causes dominant mutant phenotype (A. Ferrus *et al.*, in preparation). Furthermore, the *Sh*^{K5133} phenotype is modified by T(X;2)B27, the breakpoint of which is located ~60 kb distal to the transcription unit described in this paper. Finally the HL region is an essential part of the *Sh*C. Nevertheless, the transcription unit related to T(X;Y)B55, T(X;3)*Sh*^{LC} and T(X;Y)W32 will be the primary focus to study the molecular basis of the *Sh* phenotype in the near future.

Materials and methods

Fly strains

T(X;Y)B132 is an X-ray induced rearrangement with the breakpoint in 17.A9-12. It was induced in the same parental chromosome as T(X;Y)B55 and thus it is

used as a control strain. T(X;Y)B132 does not show any *Shaker* phenotype. All other rearrangements and mutants are described in Tanouye *et al.* (1981).

Libraries

One genomic *Drosophila* CantonS DNA library was prepared and kindly provided by Maniatis *et al.* (1978). The other library was constructed from partial *Sau*3a digest of embryonic OregonR DNA inserted into EMBL4 vector (Murray, 1983). This library was kindly provided by H. Jaeckle (Tübingen, FRG). Cosmids were isolated from a genomic *Drosophila* DNA library constructed with *cosp20* as vector and kindly provided by E. Mohier (Strasbourg, France). The λ t10 cDNA library was a gift of T. Kornberg (San Francisco, USA). Restriction fragments of recombinant phages or cosmids were subcloned into pAT153 or pUC 18/19 vector by standard cloning techniques (Maniatis *et al.*, 1982). Recombinant DNA was propagated in Er1 host-vector system under L1 containment conditions, as defined in the guidelines of the Federal German Government for recombinant DNA research.

Isolation and labeling of DNA

Recombinant DNA was isolated according to Maniatis *et al.* (1982). Cosmid and phage libraries were screened by the method of Benton and Davies (1977). Bacterial colonies were screened according to Grunstein and Hogness (1975). *In situ* hybridization was carried out with ³H-labeled, nick-translated probes according to Rigby *et al.* (1977). All other hybridizations were carried out with DNA radioactively labeled by nick-translation with [α -³²P]dATP (800 Ci/mmol), [α -³²P]-dCTP (800 Ci/mmol) or with both nucleotides simultaneously. Genomic DNA of wild-type and mutant fly stocks was isolated according to Brennan *et al.* (1984).

Isolation of RNA

RNA of the different developmental stages of *Drosophila* OregonR stocks was isolated by the guanidinium-thiocyanate method (Maniatis *et al.*, 1982). Poly(A)⁺ RNA was selected by oligo(dT)-cellulose chromatography and stored frozen in 1 mM EDTA, 0.05% SDS, 10 mM Tris (pH 7.4) buffer before further use.

Southern and Northern hybridization

Restriction enzyme-digested genomic DNA was separated on 0.7% agarose gels and transferred to Hybond-N (Amersham) or Biodyne (Pall) membranes. Pre-hybridization was in 50% deionized formamide, 5 × SSC (20 × SSC is 3 M sodium chloride, 0.3 M sodium citrate, pH 7.0) 100 μ g/ml denatured salmon sperm DNA, 5 × Denhardt's solution, 0.1% SDS for 15-60 min at 42°C.

Hybridization was in the same buffer containing $1-5 \times 10^6$ c.p.m./ml ^{32}P -labeled probes for 20 h at 42°C. Washing was in $1 \times \text{SSC}$, 0.1% SDS for 5 min at room temperature, then twice for 30 min in $0.2 \times \text{SSC}$, 0.1% SDS at 65°C. Hybridizations to Southern blots of restriction fragments of recombinant phages/cosmids were done without formamide at 65°C. Blots were washed as above.

Northern blots

Poly(A)⁺ RNA (3 µg) was separated on horizontal 1.2% agarose gels in 20 mM morpholinopropane sulfonic acid, 5 mM sodium acetate, 2.2 M formaldehyde. After electrophoresis, the RNA was electroblotted onto nylon filters (Biodyne, Pall or Hybond-N, Amersham). Northern blots were prehybridized for 24 h at 42°C in 50% formamide, $5 \times \text{SSC}$, 0.1% SDS, $5 \times \text{Denhardt's}$ solution, 20 mM potassium phosphate (pH 6.5) and 250 µg/ml denatured salmon sperm DNA (Thomas, 1980). Blots were hybridized under the same conditions with ^{32}P -labeled probes (5×10^6 c.p.m./ml) for 16–48 h at 42°C. Blots were washed several times in $1 \times \text{SSC}$, 0.1% SDS followed by $0.1 \times \text{SSC}$, 0.1% SDS at 65°C before autoradiography.

Restriction maps

Restriction maps were derived by a combination of complete, single and double digests followed by gel electrophoresis of the resulting fragments on 0.7% agarose gels (Maniatis *et al.*, 1982). The colinearity of the restriction map and the genomic DNA was verified by hybridization of the recombinant phage to whole genome Southern blots.

Sequencing of DNA subcloned into M13mp18/19 phages followed the Sanger protocol (Sanger *et al.*, 1977).

Acknowledgements

This work was supported by grant SFB 168/B4 to O.P. of the Deutsche Forschungsgemeinschaft and grants from CAYCIT and R.Pascual Fund to A.F. We are especially indebted to the Volkswagenstiftung for its support of this collaboration. A German–Spanish exchange programme 'acciones integradas' provided travel funds.

References

- Adams,D.J., Smith,S.J. and Thompson,S.H. (1980) *Annu. Rev. Neurosci.*, **3**, 131–167.
- Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.
- Brandl,C.J. and Deber,D.M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 917–921.
- Brennan,M.D., Rowan,R.W. and Dickinson,W.J. (1984) *Cell*, **34**, 147–151.
- Dayhoff,M.O., Schwartz,R.M. and Orcutt,B.C. (1978) *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation Silver Spring, MD, Vol. 5, Suppl. 3, pp. 345–352.
- Ferrus,A., Tanouye,M.A., Pongs,O. and Llamazeres,S. (1987) in press.
- Goldberg,M.L. (1979) Ph.D. Thesis, Stanford University.
- Grunstein,M. and Hogness,D.S. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961–3965.
- Hogness,D.S., Lipshitz,H.D., Beachy,P.A., Peattie,D.A., Saint,R.B., Goldschmidt-Clermont,M., Harte,P.J., Gavis,E.R. and Helfand,S.L. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 181–184.
- Jan,Y.N., Jan,L.Y. and Dennis,M.J. (1977) *Proc. R. Soc. London., Ser. B*, **198**, 87–108.
- Koto,M., Tanouye,M.A., Ferrus,A., Thomas,J.B. and Wyman,R.J. (1981) *Brain Res.*, **221**, 213–217.
- Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.
- Maniatis,T., Hardison,R.C., Lay,E., Lauer,J., O'Connell,C., Oun,D., Sim,G.K. and Efstratiadis,A. (1987) *Cell*, **15**, 687–701.
- Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York, Vol. 7.
- Murray,N.E. (1983) In Hendrix,R.W., Roberts,J.W., Stahl,F. and Weisberg,R.A. (eds), *Lambda II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Noda,M., Shimizu,S., Tanabe,T., Takai,T., Kayano,T., Ikeda,T., Takahashi,H., Nakayama,H., Kanaoka,Y., Minamino,N., Kangawa,K., Matsuo,H., Raftery,M.A., Hirose,T., Inayama,S., Hayashida,H., Miyata,T. and Numa,S. (1984) *Nature*, **312**, 121–127.
- Noda,M., Ikeda,T., Kayano,T., Suzuki,H., Takeshima,H., Kurasaki,M., Takahashi,H. and Numa,S. (1986) *Nature*, **320**, 188–192.
- Pirotta,V., Hadfield,C. and Pretorius,G.H.J. (1983) *EMBO J.*, **2**, 927–934.
- Popot,J.L. and Changeux,J.P. (1984) *Physiol. Rev.*, **64**, 1162–1239.
- Potter,S., Brorain,W.J., Dunsmuir,P. and Rublin,G.M. (1979) *Cell*, **17**, 415–427.
- Rigby,P.W.J., Dieckmann,M., Rhodes,C. and Berg,P. (1977) *J. Mol. Biol.*, **113**, 237–246.
- Salkoff,L. and Wyman,R. (1981) *Nature*, **293**, 228–230.

- Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Schneuwly,S., Kurviwa,A., Baumgartner,P. and Gehring,W. (1986) *EMBO J.*, **5**, 733–739.
- Tanouye,M.A., Ferrus,A. and Fujita,S.C. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6548–6552.
- Tanouye,M.A. and Ferrus,A. (1985) *J. Neurogenet.*, **2**, 253–271.
- Tanouye,M.A., Kamp,C.A., Iverson,L.E. and Salkoff,L. (1986) *Annu. Rev. Neurosci.*, **9**, 255–276.
- Thomas,P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
- Wolfner,M. (1979) Ph.D. thesis, Stanford University.
- Wu,C.F., Ganetzky,B., Haugland,F.M. and Liu,A.X. (1983) *Science*, **220**, 1076–1078.
- Wu,C.F. and Haugland,F. (1986) *J. Neurosci.*, **5**, 2626–2640.

Received on March 5, 1987; revised on July 1, 1987