Identification of a different-type homeobox gene, BarH1, possibly causing Bar (B) and Om(1D) mutations in Drosophila

(compound eye/homeodomain/morphogenesis/malformation/M-repeat)

TETSUYA KOJIMA, SATOSHI ISHIHARU, SHIN-ICHI HIGASHIJIMA, EUI TAKAYAMA, HIROSHI AKIMARU, MASAKI SONE, YASUFUMI EMORI, AND KAORU SAIGO*
Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, 7-3-1 Jongo, Bunkyo-ku, Tokyo 113, Japan

Communicated by Melvin M. Green, January 25, 1991

ABSTRACT The Bar mutation B of Drosophila melanogaster and optic morphology mutation Om(1D) of Drosophila ananassae result in suppression of ommatidium differentiation at the anterior portion of the eye. Examination was made to determine the genes responsible for these mutations. Both loci were found to share in common a different type of homeobox gene, which we call “BarH1.” Polypeptides encoded by D. melanogaster and D. ananassae BarH1 gene consist of 543 and 604 amino acids, respectively, with homeodomains identical in sequence except for one amino acid substitution. A unique feature of these homeodomains is that the phenylalanine residue in helix 3, conserved in all metazoan homeodomains so far examined, is replaced by a tyrosine residue. By Northern blotting, considerably more BarH1 RNA was detected in the Bar mutant than in wild type. P element-mediated transformation showed Bar-like eye malformation to be induced by transient overexpression of the BarH1 gene in the late third-instar larvae. Somatic recombination analysis indicated normal gene functions of the Bar region, including the BarH1 gene, to be required for normal eye morphogenesis.

The Drosophila compound eye consists of an array of some 800 ommatidia, each a precise, stereotyped assembly of about 20 cells including 8 photoreceptor cells (R cells) (1). Based on the sequential order of neural differentiation, R cells may be classified as R8, R2/R5, R3/R4, R1/R6, or R7 (2). Classical anatomical and genetic studies support the notion that there are no strict cell lineage relationships between R cells. The fates of cells that join a developing ommatidium later appear to be determined by cells that already have differentiated within a unit (2). Several genes involved in these determination processes have been cloned, and their possible roles in ommatidium formation have been indicated. sevenless (sev), whose expression in R7 cells is required for R7 determination, encodes a transmembrane receptor that possesses a putative tyrosine kinase domain (3) and possibly interacts with bride of sevenless (boss)-dependent signal molecules produced by R8 cells (4). The rough (ro) gene product, having a homeodomain, is essential for sending signals by the R2/R5 pair to their neighbors (5). The seven up (svp) gene product, mandatory for determining the cell identities of R3/R4 and R1/R6, encodes a presumptive steroid hormone receptor (6), whereas the glass (gl) gene product is a putative DNA-binding protein with zinc fingers and apparently is involved in photoreceptor determination (7).

In contrast to the fate-determination processes at the level of individual photoreceptors, little is known of the molecular events at very early stages of concerted ommatidium formation, although a recent finding by Mlodzik et al. (8) may suggest an important role of the scabrous gene in establishing the initial periodicity. A class of mutations including Bar (B) may also be intimately related to early events (9). For clarification of this point, examination was first made of possible determinant genes for the Bar mutation B of Drosophila melanogaster (10) and optic morphology mutation Om(1D) of Drosophila ananassae (11), in both of which ommatidium differentiation is suppressed in the anterior portion of the eye. Both loci were found to share in common a different type of homeobox gene, called BarH1, which encodes a polypeptide of Mᵦ = 60,000. RNA (Northern) blotting indicated considerably more BarH1 RNA in the Bar mutant than in wild type. P element-mediated transformation showed Bar-like eye malformation to result from the transient overexpression of the BarH1 homeobox gene in the late third-instar larvae. Normal gene functions of the Bar region have also been shown to be necessary for the normal eye morphogenesis.

MATERIALS AND METHODS

Fly and Phage Stocks. D. melanogaster strains carrying B and w mutations were obtained from Y. N. Tobari and Y. Hotta, respectively. Strain Df(1)B63-20/In(1)sc7 AM, sc7car, from which strain Df(1)B63-20 w/FM7 is derived, was obtained from Mid-America Drosophila Stock Center, Bowling Green State University. According to Sutton (12), this chromosome 1 Bar deficiency (lethal) includes almost all of the entire 16A section [16A1-6 (or 7)] of chromosome location 1:57.0. MS59 is a viable enhancer trap fly line and carries an “enhancer trap” P element marked by w⁺ and lacZ that is inserted near the Bar locus (M.S., unpublished data). Other fly strains were from our fly stocks. A phage clones AOM1(1D3) and AOM1(1D5), each including a part of the Oml(1D) locus, were kindly supplied by S. Tanda and C. H. Langley (13).

Somatic Recombination Techniques. Virgin females heterozygous for Df(1)B63-20 w/FM7 were crossed with MS59 males. Mosaic eyes were found in 2-3% of female progeny heterozygous for Df(1)B63-20 w/FM7 w/F(w⁺) w. X-ray irradiation (1000-2000 rads) of second-instar larvae was used to induce somatic recombination. Fixation and sectioning were performed as described by Tomlinson and Ready (2).

Molecular Analyses. Biotinylated DNA probes were prepared by DNA labeling and use of a nonradioactive kit (Boehringer Mannheim). Synthetic oligonucleotides 3'-ATITTTTCTIATITITCTTT (I = inosine) and 3'-GA(C or I)GG(I or C)GG(I or C)GG(I or C)CCTC(T or C)ACCTACCG(I or C)GICTT, corresponding to Tyr-Lys-Asn-Asp-Tyr-Tyr-Arg-lys-Arg and Leu-Pro-Val-Arg-Trp-Met-Ala-Pro-Glu, respectively, were used to screen tyrosine kinase

Abbreviations: R cells, photoreceptor cells; hsp, heat shock protein. *To whom reprint requests should be addressed.
†The sequences reported in this paper have been deposited in the GenBank data base (accession nos. MS5962 and MS5963 for D. ananassae BarH1 and MS5965 for D. melanogaster).
genes, while $S'$-$(CAG)_{10}$ was used to screen the M-repeat sequence (14). cDNA libraries were made from Canton-S embryos and pupae. Two other cDNA libraries made by L. Kauvar (26) were also used. All other procedures were essentially as described by Sambrook et al. (15).

RESULTS

Molecular Cloning of the Bar Locus. The Omt(ID) mutation in D. ananassae results from the insertion of a retrotранспон sor tom (11). The DNA fragment including the Omt(ID) locus has been cloned by tom-tagging (13). By using a part of this DNA as a probe, the D. melanogaster counterpart was isolated (Fig. 1a). In situ hybridization to the polyteny chromosomes of D. melanogaster indicated the wild-type X chromosome to have a single hybridization band at 16A corresponding to the Bar locus, whereas two adjacent hybridization bands could be seen on the X chromosome with the Bar mutation, B (Fig. 2a, Inset 2), associated with a duplication of 16A1–16A7 (18). Recently, a partial physical map of the Bar region was shown by Tsubota et al. (19). Since there appears to be no overlapping in restriction maps between their clone and ours, the latter may likely represent an internal Bar sequence situated more than 50 kilobases (kb).

![Fig. 1. (a) Structures of two BarH1 homeobox genes: D. ananassae BarH1 (Upper) and D. melanogaster BarH1 (Lower). A thick horizontal line labeled Y shows the location of a D. ananassae EcoRI fragment hybridized with the M-repeat probe described in Materials and Methods. The coding regions are shown by boxes, and the thin horizontal lines associated with them show exon size and location. The thin line in the middle shows the region covered by four cDNA clones so far obtained. The filled boxes indicate homeodomains in the protein products and M shows the location of an M-repeat (14). Large vertical arrows (D9, D73, or D88), either over or under the physical map of the D. ananassae genomic DNA, show the locations of the insertion sites of tom elements (13). The arrow, labeled BREAK POINT, shows the location of one breakpoint in Omt(ID) (a), a revertant of Omt(ID) (b). B, BamHI; E, EcoRI; H, HindIII; P, Pst I; S, Sal I; X, Xho I. (b) Structure of pBH1. A 1875-base-­long BarH1 (BH1) cDNA fragment, which includes the entire coding sequence (1623 bp) and a part of the third untranslated region (246 bp), was ligated to the BamHI linkers and inserted into the BamHI site of pHST14, a pUC plasmid having a partial P' sequence along with the 5' and 3' portions of the gene encoding a 70-kDa heat shock protein (hp70; hp70, ref. 16). After insertion of a rosy+ marker, the resultant plasmid, pBH1, was used for injection. Bg, Bgl II. Other restriction sites are the same as in A. Since the hp70 promoter includes the region encoding the first 338 amino acids of the hp70 protein (16), the resultant hp-BH1 fusion protein is expected to consist of 885 amino acids. (c) A scheme of somatic recombination for the generation of w clones lacking Bar functions in a wild-type background.]


![Fig. 2. Chromosomal location (a) and expression (b) of D. melanogaster BarH1 homeobox gene. (a) In situ hybridization was carried out with fragment 1 of Fig. 1a as a probe. Under phase-contrast microscopy, a photograph was taken of the wild-type X chromosome showing a single hybridization band at 16A, the site of the Bar locus (12). The hybridized region indicated by an arrowhead appears as a very bright band in contrast to the picture taken without phase contrast (Inset 1). Inset 2 shows a part of the hybridization pattern of the X chromosome. (b) Northern blot analysis of D. melanogaster BarH1 RNA. Hybridization was carried out with a cDNA fragment, fragment 2 of Fig. 1a, as a probe. Odd- and even-number lanes, respectively, are for the wild-type and B mutant RNAs. Lanes: 1 and 2, embryos; 3 and 4, larvae; 5 and 6, pupae; 7 and 8, adults; 9 and 10, early third-instar larvae (collected 96–120 hr after oviposition at 25°C); 11 and 12, late third-instar larvae (120–168 hr); 13 and 14, early pupae (120–168 hr pupae); 15 and 16, middle stage of pupae (168–192 hr); 17 and 18, late pupae (192–240 hr). Molecular size in kb is shown on the right. Ac indicates patterns of rehybridization with an actin probe (pACT5C, ref. 17).]

from the distal B breakpoint. The phenotype of B is very similar, if not identical, to that of Omt(ID); thus, the above finding indicates the Bar locus in D. melanogaster to possibly be the counterpart of the Omt(ID) locus in D. ananassae. Accordingly, this would imply the aberration of an identical gene to cause both Omt(ID) and B mutations.

Identification of Similar Homeobox Genes in Omt(ID) and B Loci. The ectopic expression or deficiency of genes related to the switching mechanism for cell differentiation has been shown to cause malformation in various tissues (20). To determine whether a gene (or genes) belonging to the same category takes part in Omt(ID) and/or B mutations, a 50-kb-long cloned Omt(ID) locus (13) was examined by blot hybridization. Three sets of mixed oligonucleotides were used as probes, two to search for receptor genes encoding tyrosine kinase and one to find genes with M-repeat or opa present in certain neurogenic genes (14) and homeobox genes (21) in Drosophila. Although no positive signal could be detected with the receptor probes, an M-repeat-positive sequence was present in the vicinity of a tom insertion hot spot (see a thick line labeled Y in Fig. 1a).

D. melanogaster cDNA clones were isolated by using as a probe a DNA fragment (fragment 3 in Fig. 1a), corresponding to the M-repeat-positive region. The nucleotide sequences of the cDNA inserts were determined as were also those of the genomic counterparts of D. melanogaster and D. ananassae (Fig. 3). Conceptual translation shows both loci to share in common a gene in which three coding exons are present. The exon/intron structure of the D. ananassae gene was deduced by homology with that of D. melanogaster. A 33-base-pair (bp)-long M-repeat was found embedded in the putative first exon of the D. ananassae gene. However, the D. melanogaster counterpart lacks this repeat. Hereafter we refer to these two genes as the BarH1 homeobox gene collectively.
The BarH1 homebox gene products of *D. melanogaster* and *D. anna
daeae* consist of 604 and 543 amino acids, respectively (Fig. 3). If deletions/insertions are neglected, the amino acid sequence homology is 82%. Four cDNA clones so far ob
tained, as a whole, cover the entire coding sequence along with the 3' untranslated region and 69 nucleotides of the 5' untranslated region (bounded by horizontal L-shaped arrow in fifth line of Fig. 3). The 5' untranslated regions are similar in nucleotide sequence to each other (see underlines in Fig. 3) and include a TATA box-like sequence (TATATAAA) along with a 30-bp-long identical sequence of unknown function. Although the 3' untranslated regions are less homologous, a
polyadenylation signal, AATAAA, was seen in a 23-bp-long conserved sequence (Fig. 3). Two adjacent poly(A) sites were detected in the case of D. melanogaster.

The most striking feature of the BarH1 gene product is the presence of a homeodomain (Figs. 3 and 4). The amino acid sequence of the D. melanogaster homeodomain is identical to that of the D. ananassae counterpart except for one amino acid substitution. The BarH1 homeodomain contains eight of nine invariant amino acids deduced from 20 Drosophila homeodomains (5) and resembles most in sequence that of eve (22), the identity being 50%. A phenylalanine residue in presumptive helix 3 (see a box in Fig. 4), strictly conserved in all homeodomains of metazoa so far known (23), is replaced by a tyrosine residue in both BarH1 homeodomains, so that the BarH1 gene may be classed as a unique subclass of homeobox genes.

Another interesting feature of the putative BarH1 gene product is that the homeodomain is surrounded by many simple repeats of one or two amino acids. Some may be of less important biological function, since many deletions/insertions are present within the repeat-rich regions (Fig. 3).

Induction of B-like Eye Malformation by Transient BarH1 Homeobox Gene Expression. Poly(A)+ RNA extracted from embryos, larvae, pupae, and adults of D. melanogaster was examined by Northern blotting (Fig. 2b). Two distinct BarH1 transcripts, 3.1 kb and 0.8 kb long, were found. The 3.1-kb RNA, mainly expressed in both embryos and pupae, was found to encode the homeodomain (data not shown). At the pupal stage, considerably more 3.1-kb RNA was detected in the strain carrying the B mutation than in wild type (compare lanes 5 and 6). Although the total expression was much less, the same was noted in the late larval third instar (lanes 11 and 12), the time when ommatidium differentiation begins (1, 2).

To further clarify the relationship between abnormal BarH1 homeobox gene expression and the mutation, D. melanogaster transformants possessing a BarH1 coding sequence with the hsp70 promoter (Fig. 1b) were made by P element-mediated transformation (24). Two transformants were obtained, both of which exhibited similar morphological changes upon heat induction. The heat induction of the BarH1 homeobox gene in late third-instar larvae caused ommatidia to decrease in number in nearly all adult eyes (Fig. 5 a and b) as in the case of the B mutation (Fig. 5c). Additional scars were also detected in about 30% eyes. Except for minor defects occasionally found in the leg, there was no other morphological change. In contrast to the late larval third instar, no appreciable morphological change was found by heat treatment at the pupal stage. The transient overexpression of the BarH1 homeobox gene at the late larval third instar is thus concluded to cause eye malformation in a manner virtually the same as that noted in the Bar mutant.

Requirement of Normal Gene Functions of the Bar Region for Normal Eye Morphogenesis. Using x-ray-induced mitotic recombination, mosaic eyes possessing both clones of normal cells and those lacking Bar gene function were produced. Normal cells are marked with a transposed white gene, P[w+], whereas almost the entire B region [16A1–6 (or 7)] is deleted in the chromosomes of deficient cells without pigment granules (Fig. 1c). In all of the 20 eyes so far examined, the unpigmented region lacking the Bar gene function was found to be occupied by irregularly fused, bulging ommatidia greatly differing in the number of rhombomeres (0–12 per ommatidium) (Fig. 5 d–f). Fusion or ramification of rhombomeres was noted occasionally. Bristles were completely absent. The normal gene functions of the Bar gene region thus appear quite likely required for normal eye morphogenesis.

DISCUSSION

The Bar locus of D. melanogaster and the Om(1D) locus of D. ananassae share in common a different type of homeobox gene, which we call BarH1. As with other homeodomain-
containing proteins, the BarH1 gene product is likely a DNA-binding protein performing key roles in cell differentiation (20). Intimate relation between the BarH1 homeobox gene overexpression and the Bar mutation, B, was exhibited by Northern blotting and P element-mediated transformation. Reduced eye morphology in strains carrying the Om(1D) mutation also may reasonably be considered to be due to abnormal activation of the BarH1 gene by the retrotransposon *tom*, since normal eye morphology is restored on introducing a chromosomal breakage between the BarH1 homeobox gene and *tom* resulting in mutation (ref. 13; see Fig. 1a).

Besides *B*, many other mutations giving rise to *B*-like eye morphology have been mapped in the Bar locus (12). All of them are gain-of-function mutations associated with inversion and translocation, in which there is one breakpoint in the region 16A1–2 or its vicinity. By restriction mapping, Tsutbota et al. (19) showed the breakpoints of *B*¹⁰, *B*¹², and *B*²⁰ה to be within 30 kb from the distal *B* breakpoint. By our calculations, the BarH1 coding sequence was located more than 20 kb from the nearest breakpoint, that of *B*²⁰ה. Thus, should the BarH1 homeobox gene be the only gene responsible for all *B* mutations, chromosomal fragments translocated to the Bar region upon Bar gene rearrangement would surely contain an enhancer element that functions in a relatively long range.

Alternatively, the Bar gene region may be much more complicated and include plural genes capable of inducing similar narrow eye phenotypes on aberrant expression. A recent genetic dissection of the *Om(1D)* locus of *D. ananassae* (13) appears partially to support this view, since a new locus designated *Om(1H)* with a mutant phenotype similar to that of *Om(1D)* was found situated quite close to the *Om(1D)* locus. It appears pertinent to point out here that another homeobox gene, BarH2, similar in sequence to the BarH1 gene, was recently identified in the Bar gene region (T. Michiue, S.I., Y.E., and K.S., unpublished data).

Somatic recombination analysis indicated the Bar region to include genes responsible for normal eye morphogenesis. This region may contain 10 or so genes in addition to BarH1 and BarH2. The severely disrupted pattern formation in the region lacking Bar functions can be explained as being due to loss of control over cluster-spacing at the morphogenetic furrow, where ommatidium assembly begins, or to the defective recruitment of cells into different types. At present, the phenotype for the loss-of-function mutation of the BarH1 homeobox gene is not known because of lack of available mutants. However, in consideration of the general roles of homeodomain proteins (20), the absence of the BarH1 homeobox gene would certainly appear responsible in part for the morphological changes in the Bar gene function-defective region.

Recent analysis has shown the Antennapedia (Antp) homeodomain to contain four helices connected by short peptides and held together by a hydrophobic core (23). Five of 10 amino acids constituting the hydrophobic core are identical, and 2 are similar in the BarH1 homeodomain of *D. melanogaster* (see Fig. 4). Amino acid 9 of helix 3, possibly the determinant of binding specificity (25), is glutamine and is identical to that of Antp. This homeodomain may thus be similar in both tertiary structure and target sequence specificity to the Antp homeodomain. However, the most conserved phenylalanine residue in helix 3 (23) is replaced by a tyrosine residue in the BarH1 homeodomain. Since identical change was also noted in BarH2 (T. Michiue, S.I., Y.E., and K.S., unpublished data), the Bar gene region may be considered a unique region having a pair of homeobox genes belonging to a peculiar subclass.

We are grateful to Y. Sakoyama, L. Kauvar, and T. Uemura for providing some of the cDNA libraries used here. We thank S. Tanda and C. H. Langley for the Om(1D) phages and their comments. We also thank R. Ueda and T. Michiue for discussion. The initial stages of the in situ hybridization experiments were conducted in collaboration with Y. N. Tobari, to whom we are grateful for sharing unpublished data. This work was supported in part by grants from the Ministry of Education, Science and Culture of Japan to K.S.