The role of localization of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo

Thomas Berleth, Maya Burri¹, Gudrun Thoma, Daniel Bopp¹, Sibyll Richstein, Gabriella Frigerio¹, Markus Noll¹ and Christiane Nüsslein-Volhard

Max-Planck-Institut für Entwicklungsbiologie, Abteilung Genetik, Tübingen, FRG and ¹Department of Cell Biology, Biozentrum der Universität, Basel, Switzerland

Communicated by C.Nüsslein-Volhard

The organization of the anterior pattern in the *Drosophila* embryo is mediated by the maternal effect gene *bicoid*. *bcd* has been identified in an 8.7-kb genomic fragment by germ line transformants that completely rescue the mutant phenotype. The major transcript of 2.6 kb includes a homeobox with low homology to previously known homeoboxes, a PRD-repeat and a M-repeat. *In situ* hybridizations reveal that *bcd* is transcribed in the nurse cells. The mRNA is localized at the anterior tip of oocyte and early embryo until the cellular blastoderm stage. The localization of the transcript requires the function of the maternal effect genes *exuperantia* and *swallow* while transcript stability is reduced by functions depending on posterior group genes.

Key words: Drosophila/mRNA localization/pattern formation/anteroposterior axis/homeobox

Introduction

Position along the anteroposterior axis of an insect egg is determined by the polar distribution of maternal factors (Sander, 1976; Frohnhöfer *et al.*, 1986). In *Drosophila*, systematic screens for maternal effect mutations have identified 13 loci that are involved in the establishment of two centers of organizing activity positioned at the egg poles (Nüsslein-Volhard *et al.*, 1987). Each of these centers harbors factors which can reorganize major portions of the anteroposterior pattern in transplantation experiments involving mutant embryos as donors or recipients (Lehmann, 1985; Lehmann and Nüsslein-Volhard, 1986).

The gene *bicoid* (*bcd*) is required for the development of the entire anterior half of the embryo: in embryos from flies homozygous for strong *bcd* alleles, all structures of head and thorax are missing and the blastodermal fate map of the remaining abdominal segments is shifted anteriorly. In addition, the anterior-terminal acron is transformed into a posterior telson in *bcd*⁻ embryos (Frohnhöfer and Nüsslein-Volhard, 1986).

Transplantation experiments revealed concentrated bcd^+ activity in the anterior 15% of the wild-type embryo. Cytoplasm taken from this region can induce anterior development (ordered head and thoracic structures) either by normalizing the phenotype of bcd^- embryos or by

inducing the development of anterior structures at ectopic sites (Frohnhöfer and Nüsslein-Volhard, 1986). In the latter case the degree of anteriorness reached at ectopic sites declines with proximity to the posterior center, presumably due to an inhibitory effect of the posterior activity on the anterior bcd^+ activity (Frohnhöfer *et al.*, 1986; Nüsslein-Volhard *et al.*, 1987; Lehmann, 1985; Frohnhöfer and Nüsslein-Volhard, 1986).

In addition to posterior factors, two other genes, *swallow* and *exuperantia*, influence bcd^+ activity. Mutations in these maternal genes result in distortion of pattern elements along the anteroposterior axis (Schüpbach and Wieschaus, 1986; Frohnhöfer and Nüsslein-Volhard, 1987; Stephenson and Mahowald, 1987). The anteriormost structures are lacking in *exu* and *swa* mutant embryos, while the subterminal regions of posterior head segments and thorax are enlarged and spread anteriorly. Pricking and transplantation experiments suggest that the *exu* and *swa* phenotypes are caused by a shallow distribution of bcd^+ activity throughout the egg (Frohnhöfer and Nüsslein-Volhard, 1987).

To investigate the interactions between the various maternal functions at the molecular level, we have cloned and sequenced the *bcd* locus and analyzed its transcripts. We demonstrate that localization is affected by mutations in other maternal genes: *bcd* mRNA is strictly localized at the anterior end of wild-type oocytes and embryos, but is almost evenly distributed in *exu* and *swa* embryos. A negative effect of posterior activity on *bcd* mRNA stability is observed.

Results

Identification of the bcd locus

bicoid has been mapped genetically to the interval between the loci proboscipedia (pb) and Deformed (Dfd) in the cytological region 84A on the right arm of the third chromosome (Frohnhöfer and Nüsslein-Volhard, 1987). In this region a molecular walk initiated from the site of an insertion into Dfd was available (Scott et al., 1983). We mapped the proximal breakpoint of the deficiency Df(3R)LIN which uncovers bcd—within that walk and screened the sequences encompassing ~60 kb proximal of Dfd for hybridization to maternal transcripts using ³²P-labelled cDNA made from poly(A)⁺ RNA of staged embryos (Figure 1a).

The 8.7-kb *Eco*RI fragment at walk positon -42 to -33 kb hybridized to cDNA from mRNA of cleavage stage embryos as well as egg follicles. Probes from that region detect a 2.6-kb RNA which is present in oocytes and early embryos (Figure 1c). This same restriction fragment has also been isolated by homology to a sequence (PRD-repeat) in the segmentation gene *paired* (Frigerio *et al.*, 1986; Kilchherr *et al.*, 1986), and the corresponding cDNAs have been shown to detect an RNA localized at the anterior tip of oocytes (Frigerio *et al.*, 1986). The localization of this

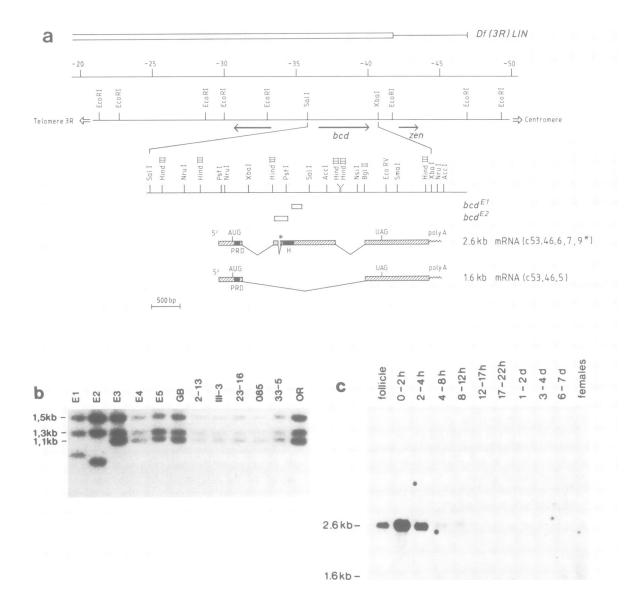


Fig. 1. (a). Map of the chromosomal region and transcriptional organisation of the bicoid gene. The chromosomal region has been isolated by Scott et al. (1983) in a chromosomal walk to which the scale (second line) and EcoRI restriction map (line three) refers (Scott et al., 1983). The orientation and approximate extent of the bcd and adjacent transcription units are indicated by the arrows. The proximal gene has been identified with the zerknüllt gene (Wakimoto et al., 1984; Doyle et al., 1986), the distal gene [producing a 1.5-kb poly(A)⁺ RNA during the first half of the embryogenesis) has not been identified with a genetically characterized gene. The splicing patterns of the bcd transcription unit result in a major bcd mRNA of 2.6 kb and a minor bcd mRNA of 1.6 kb. They have been derived from DNA sequences of several bcd cDNAs. The acceptor site of the second intron in c53.46.9 is 15 nt upstream from that found in c53.46.6 and 7 is marked by an asterisk (for DNA sequence see Figure 3). PRD = PRD-repeat H = homeo domain. (b) Southern blot analysis after digestion with HindIII of genomic DNA of flies hemizygous for 11 EMSinduced bcd alleles revealing small deletions in the alleles bcd^{E1} and bcd^{E2} . The size reduction of the HindIII fragment that corresponds to a large portion of the third exon (see 1a) reflects deletions of ~ 180 bp (bcd^{E1}) and 260 bp (bcd^{E2}). The position of the deletions was determined by SI nuclease mapping using a probe labelled at the position of the Sall site in a DNA fragment derived from cDNA clone pc53.46.6 but corrected to the genomic sequence at two points of divergence (see legend to Figure 3). The protected fragment corresponds to the entire cDNA portion of the probe (894bp) in wild-type, while it is of 120 bp (bcd^{E1}) and 390 bp (bcd^{E2}) length in the two mutants (not shown). The deduced location of the deletions within the bcd coding region is indicated in a. (c) Developmental profile of bicoid transcripts. Poly(A)⁺ RNA of the developmental stages indicated above each lane were analysed by Northern blot analysis using c53.46.6 DNA as probe. For size calibration, end-labelled fragments of a partial RsaI digest of pBR322 (Bopp et al., 1986) and of a HindIII digest of lambda phage DNA were used (data not shown).

transcript at the anterior egg pole of cleavage stage embryos suggests identity with the bcd^+ activity detected in transplantation experiments. To test whether the 8.7-kb fragment indeed included the *bcd* locus we used it for P-elementmediated transformation. As a selective system, a recipient strain which was mutant for the alcohol dehydrogenease gene (Adh^{n2}) as well as for bcd^{E1} was transformed with the 8.7-kb fragment cloned into the P^{PA} vector (Goldberg *et al.*, 1983). Three independent transformant lines were obtained in which both the ethanol sensitivity and the *bcd*⁺ function were completely rescued by the insertion of the cloned DNA fragments into the genomic sites 12E, 19A and 61D respectively. These results demonstrate that the complete functional unit of the *bcd* gene is contained in this

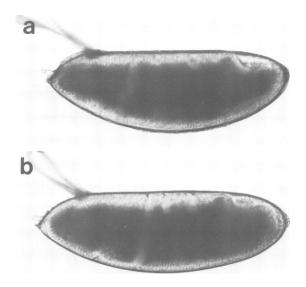


Fig. 2. The position of the head fold in embryos from transformed flies (a) Embryos derived from a female containing a mutation (bcd^{E1}) at both endogenous copies, but diploid for a bcd^+ gene copy inserted in the X chromosome at position 19A (bcd^{+8}) . (b) The maternal genotype of this embryo is two normal endogenous copies and two extra copies caused by a bcd^+ insertion at 12E (bcd^{+5}) . The position of the head fold is indicated by an arrow.

8.7-kb fragment. Only the 2.6-kb transcription unit could be detected in this DNA fragment, while the transcription units to the right and left of *bcd* (Figure 1a) are strictly zygotically expressed.

Extra copies of bcd⁺ cause oversized head regions

A sensitive assay for bcd^+ activity in the embryo is the position of the headfold which normally appears during early gastrulation at 65% egglength, approximately at the border between head and thorax. The position of the headfold is strictly dependent on the number of bcd^+ copies in the female producing the embryo. It is shifted anteriorly in embryos from hemizygous females, posteriorly in embryos from females with three wild-type copies of the bcd^+ gene (Frohnhöfer and Nüsslein-Volhard, 1986). Figure 2a shows an embryo from a female in which both bcd^+ copies are derived from the transformed gene. The headfold is at the normal position of 65% egglength. The transformed gene copies in the genome allow the construction of females with additional copies of the bcd^+ gene. Figure 2b shows that in embryos from females with four bcd^+ copies (two of which are the endogenous copies) the head region is very large. Surprisingly, these embryos can develop into apparently normal adults. This indicates powerful size regulation capacities in the Drosophila embryo, which have also been noted for other genotypes causing oversized embryonic thoracic (exu and swa, Schüpbach and Wieschaus, 1986) or abdominal (bicaudal, Mohler and Wieschaus, 1985; Nüsslein-Volhard, 1977) anlagen.

Transcriptional organization and DNA sequence of the bicoid gene

Two classes of *bcd* mRNA, 2.6 and 1.6 kb in size, can be seen in early embryos (Figure 1c). The 2.6-kb transcript is predominant in oocytes and cleavage stage embryo. It dis-

appears after gastrulation (>4 h). A much weaker transcript of 1.6 kb persists throughout all developmental stages.

Analysis of the genomic as well as several cDNA sequences showed that both the large and small mRNA are derived from the same primary transcript by differential splicing as indicated in Figure 1a. The minor 1.6-kb transcript corresponds to cDNA clone c53.46.5, in which an alternative splice pattern has deleted the second and third exon. As expected, this transcript is not detected in Northern blots hybridized with second and third exon probes (not shown).

Transcripts of the 2.6 kb size class contain all four exons (Figure 1a). However, the sequence of cDNA clones revealed heterogeneity with respect to the splice acceptor site of the second intron. Thus, in c53.46.9 the third exon may encode five additional amino acids (Asp-Val-Phe-Pro-Ser) close to the amino-terminal end of the homeodomain (see below) as compared to c53.46.6 and c53.46.7. Hence, these products of differential splicing produce three long open reading frames (ORFs) coding for proteins of 489 or 494 amino acids in the major 2.6-kb mRNA and for a protein of 149 amino acids in the minor 1.6-kb species (Figure 3).

All ORFs start with the ATG at nucleotides 1416-1418from the *Sal*I site and are terminated by an amber codon at nucleotides 4010-4012 (Figure 3). The termination signal is followed by 817 (possibily up to 820) bases of untranslated 3'-non-coding sequence with the likely poly(A) addition sequence AATATA, 21 (up to 24) nt from its end (Figure 3). It is likely that the 5' end of the *bcd* transcript is close to the 5' end of c53.46.6 (Figures 1 and 3) as the 5' ends of the two longest cDNAs are only separated by 15 nucleotides and their lengths of 2453 and 2438 bp [without poly(A)⁺] are close to the length of the 2.6-kb mRNA as determined by Northern analysis (Figure 1c). In addition, possible TATA boxes are located ~ 100 bp upstream of the 5' end of c53.46.6.

The first exon contains a PRD-repeat, consisting essentially of alternating histidines and prolines also found at the carboxy-terminal region of the *paired* protein and within a number of genes expressed during early development (Frigerio *et al.*, 1986). The third exon encodes in its aminoterminal portion a new type of homeodomain with no more than ~40% homology on the protein level to any of the known homeobox sequences (Frigerio *et al.*, 1986). In the second half of the third exon there is a region of repetitive glutamines called M- (McGinnis *et al.*, 1984) or opa-repeats (Wharton *et al.*, 1985). A remarkable feature of the entire *bcd* coding region is its high content of prolines (10%). Apart from the homeobox, M-repeat and PRD-repeat, no significant homologies to other sequenced genes have been detected (W.Driever, personal communication).

Analysis of bcd alleles

In several mutagenesis experiments, a total of 12 ethylmethane sulphonate (EMS)-induced alleles have been isolated, one of which has been shown to be a deletion including several adjacent complementation groups [Df (3R) LIN, Figure 1a] (Frohnhöfer and Nüsslein-Volhard, 1986). While nine of the alleles are apparent point mutations, in the two strong alleles bcd^{E1} and bcd^{E2} EMS mutagenesis has led to small deletions that could be detected by both S1 nuclease protection analysis and restriction site mapping. S1 nuclease mapping places the 3' end of the sequence

	Sall			
1	GTCGACTGGAGTGTCTGTGAATTGACTTTTGTTGCCAGTTGGCAGCGGCAGAAGCAGCAAAGCCCGGCCAACAAGCAACAA	90		
	Stul BATECEAAAABEAAACACEAEAATTATTTEGEAAATEGTEATTAAAAATATTTEGECTTAAAGECETTEGEACAETTGETTAAAGETEAACT	180	oserprothrproserthrprothralahismetthroluwistyrseroluserprothashalatyrtyrashtyrasholyölymisäs gtcacccacgcccctcaacgcccactgcaccacgagcactacggagcattacggagctattcaacgcagtcattcaacgacgccactacagcagc sait	2700
	HINDIII Geotogtigggtgtgttttaaaatgttaagottgggccaatgcactgagcaacttaatgcttgtagatatttacacaatattcttcaac	270	Ala nhisalaginalaasnarghismethismetgintyrproSergijgiprogijprogijyserinrasnvalasngijgijginpr teaegeceasgecaategteaeatgeaeatgeaeteetteeggaggggggggg	2790
	GCTANACATATCGAATTTTCCAAATATGGAGCCTGAAAATAATAATAATGCCAATCCTAGCTTAAAATCAGAAATGAGTAGAACAACTTAAA	360	ePheG InG InG InG InG InVa IN 15 Ason 15G InG InG InLeun 15M 15G InG InAson 15Va IProm 15G InmetG InG InG InG InG InG InG InG InG InG In	2880
	АЛЛАТТАЛСЛАЛАGAATCGAACGCTACAGCTAATTAACTCGACAACTGGTTACCTTTTATTCTTCTAATACATTTTATAATGCACTGCCT	450	Ser na 1 a Ging ing ing ing ing ing ing ing ing ing	2970
	AACAGGTACAGATAGCAAGCACTATATGCTGTCTTACAAAACGATTATATGATATTTTCTTTC	540	¢	
	АЛЛАСАЛАСТСБАТСТССАССАТССТТАТТСТТБТСССАЛБТССТТАТАТАТСТСБСБАТАСТАЛБАТТБАЛТААТБАБТТАТТАЛТА АЛЛАСАЛАСТСБАТСТССАССАТССТТАТТСТТБТСССАЛБТССТТАТАТАТСТСБСБАТАСТАЛБАТТБАЛТААТБАБТТАТТАЛТА	630	DTyrAsaPhedanSyrSyseT yr TyrHetdargSyrG LyMet SyrG Lydla Thr Al SyrrAl SyrrAl a WaldaldargG Lydlad 1 SyrPr CTACAACTTCAACGCTCGTACTACATGCGATCGGGATGTCTGGCGCCATCGGCATCGGCTGTGGCCCGAGGCGCTGrCTCGCC Accl	3060
	GCGGAAGTATGTAACAGAATAAACTACAAAGTGCACATTTTGTTCAATTCAGGCTGGACTGGACTGGAGCATATTAATATTATAATATTA	720	อธิโมร์ยาธิโมเชื้าโมราธิโมProLeuThrProLysสรกสรอธิโมร์ยาProSerLeuCysGlyTleGlyTleGlyGlyProCysสไสไได้สไ GGGCTCCGAGGTCTACGAGCCATTAACACCCAAGAATGACGAAAGTCCGAGTCTGTTBTGGCATCGGCATCGGCGGCGTCTGCGCCATCGG มายามายามายามายามายามายามายามายามายามาย	3150
	ACAAAAATTCAAATTAAACATTCGACACTTGTCTAATTGATTCCTAAATTTGGGGTGCCTGTTTGTT	810	aValGlyGluThrGluAlaAlaAspAspMetAspAspGlyThrSerLysLysThrThrLeuGlvf — intron 3 CGTTGGCGAGACGGAGGGCGACGACATGGACGACGGAACGAGGACGACGACGACGACGACGACGACG	3240
	тесаласабабсалабабттаабттааттббттетастаттбтасалтателабостттттаттаттатесалатбелал	900	NINGIII TGATCTCTTGATTCTGAGTGTGGGGGTTTATAAATTGÅAGCTTTÄÄAGCTTTGTAACTTGTAAACTGTCTGGGTTTGAGATGTTATTCTGAAA	3330
	TCTCTACAAATAAATAAACCTECGACGTTTTAGAACATTCACCTTTTGTCAGTGAGGACAACCTTTCAATACAGCCCGACAGGGGGCTCT Apal I	990		3330
	СТАСТБСТБТСТСТСАСБСССССТББТБАЛАЛСБСТВТВСАСТСАЛТСББТТТБСАБСТТТБССБТАСТБТТСБАТТАЛАЛАСТТТТАЛ Асс 1	1080	C1=1 GTACTTCTATTTCCGATGATGAGATTGGGAGTTCTTCAATATTTAACATTTAACTTATTAAGTTTTTGTTTTCTAAATTAGACATGGC	3420
	аттабаббсаласаттталалаталалтбтссалататттбтсталалтбтаттб ¹ хбасбсттаттбаттттталаттастсалалбал	1170	ATTTCTGAAAGGGAAGTACAAGTGTTAAAGATGTATTTTAATATAGAATTTGTATCAAAGGTTAAGATTTCAACCGTTTGAAAGCCCTTA	3510
	Pil 5'end c53.46.6 TgTTCATCGAGGGAGGGCCGCCAATTGTGCCATCTCTACATCTCTCGCTCATCCCTAAATAACGGCACTCTGCAGATGGCGAAGCAGTGG	1260	GTTTTCAGGGTTTTTTACTTTTTTATTCATGTAATCACTCTTAATACACTGCAAGTTAAAATAGCATTTCTTTGACCAGAAAAATAAGAA	3600
	S'end c53.46.7 Nrul S'end c53.46.5 Atcgcaaaaacgcaaaatggggggaaataagttccggaagggtccggaaggacggaagggggggg	1350	Nsil Ndel	3690
	5: end c3.3.46.9 Metalaciapproproase in Titgecattiticegegegegetteganatteginaagataacgegegegegegegegegegegegegegegegegegeg	1440	TETATGEÄTTTTAANAGTGAAAACAGAETEÄTATGETGATGAACATTTTTAGETAATAATTGTAACAATAATTTAGEAATTAGEAATTIGAA	3690
	PRO-repeat snPheTyrHisHisProLeuProHisThrHisThrHisProHisProHisProHisProHisProHisProHisProHisProHisProHisHisG		Intron 3 → leteuGluProteutysGlyteuAsplysSerCysAspAspGlySerSerAs TTTATTTATGTTCTAAATGCGTTCGCTCTCTCCCTAGATCTTGGASCGTCTGGACAAGAGCTGCGACAGGGGGGAGGGAGGGAGGG	3780
	ACTITTĂ <u>\$\$ATCATCCGCTGCCCCACGCACACGACACCACCGCATCCGCATCCGCATCCGCATCCGCATCGCACCGCACCCACC</u>	1530	PASPMELSerThrG1y11eArgA1aleuA1aG1yThrG1yAsnArgG1yA1aA1aA1aPheA1alysPheG1yLysProSerProProG1nG1 CGACATGAGCACCGGAATAAGAGCLTTAGCAGGAACCGGAAATCGTGGAGCGGCATTTGCCAAATTTGGCAAGCCTTCGCCCCCACAAGG	3870
	TAN'ISPradi ni Lewi Ini Lewi Probrodi nopretrgas norodheas piled — ini ron 1 <u>Ancatocy</u> cagotto agtigococcaca attoccanatoccito gatitade tradito catoccaco agaga aggoco to to to coccago Ancatoccaca aggoco ag	1620	Net yProGINProProLeug IyMetg Iyg Iyya IAI aleug Iyg IuserAsng Intyrg Incys ThrMetAsp Thr I leMetGINAI atyrAs CCCTCAGCCGCCCCCGGGGATGGGGGGGGGGGGGGTGGGATGGAT	3960
	AAAGCTACAGTACAGATTCCCTATGGTGAACAAACAACCAGTGCGATCACTGATGACCATAAACATTTATTGAGCCGCAGCAAATGTGTT	1710	A A Sacij PprohisargasalaalagiyasnserginphealatyrCysPheasnEhd J?end c53.46.9 Tecccatogaacgacgegaacgacgtgachaggacgaggaggaggaggaggaggaggaggaggaggagga	
	A be 1 TCTAGAACATAGGGCGAAATCTTCTATTATCTTGTTGGACTTTTAAAGTATCGTAGCAGAATCTAAATAACAATTGATATTAATTA	1800	T HDAI ECORV MIUI	
	A () GTTACAGTTAGTATAGTATAATTGTATTGTATTGGGGGGAATGTTCTAATAGTGTTTGGGGAAATGTTCTAAAAGATGTTCTAAT	1890	AGGTTAACCACTGTTGTTCCTGATTGTACAAATACCAAGTGATTGTAGATÅTCTÅCGCGTAGAAAGTTAGGTCTAGTCCTAAGATCCGTG Ngji	
	8a11 Acci		TAAATGGTTCCCAGGGAAGTTTATGTACTAGCCTAGCCAGGCCGCACGGATTCCAGTGCATATCTTAGTGATACTCCAGTTÅACTC Smal	
	GAAATGGACGAATGTTAAACCTGTTGCACTCACACCGAATATCAGTAATGTCTATTTTTCAAAAGCCACATCTATGGCCACTGGGTÂTAC	1980	TATACTITICCC-JCAATACGCTATTCGCCTTAGATGTATCTGGGTGGCTGCTCCACTAAAGCCCGGGAATATGCAACCAGTTACATTTGA	
	ATTATTGACTTAATACACTTCATACAACATATTTCCAAAAACAAGCATTGTTGTCCTGCATGATGATAGTGAAAGTAATATTGCAAGAT Hindiii	2070	GGCCATTTGGGCTTAAGCGTATTCCATGGAAAGTTATCGTCCCACATTTCGGAAATTATATTCCGAGCCAGCAAGAAAATCTTCTCTGTT	
	Intern 1 + Leupheasobiuargingiy TCGGTCCCCGAAGCGAATCGTCCTTTCACGTTTTTATATAAAGACAGTGTACCCCTTGATCTTTGAAGKTTTTCGATGAGCGAACGGGA	2160	ACAATTTGACATAGCTAAAAACTGTACTAATCAAAATGAAAATGATTCTCTTTGGGGGTAATCTCATACAATGATTACCCTTAAAGATCG	
	BlalleAsinTyrAsinTyrIleArgProTyrLeuProAsinGlinHetProLyBroAl — intron 2 GCGATAAACTACAACTACATCCCCTATCTGCCCAACCAAGCGAAGCCAAGCCGAAGCCAAGCCAAGCCAAGCCAAGCCAAGCCAAGCCAAGCCAAGCCAAGCCAAGCCA	2250	AACATTTAAACAATAATATTTGATATGATATTTTCAATTTCTATGCTATGCCAAAGTGTCTGACATAATCAAACATTTGCGCATTCTTTG	
	intron 2 — i — SpraiPheProSecQluGlueuProAspSerLeuValMetArgArgProArgArgThrArgThrThrPheThrSerSerGIn11 ArchatotturtircciccaGacactGccGActtrtaGGGGAcGAcGacGaccacttrtAccActtrtAccActtrtAccActtrtAccAct		ALLAABAA IAB ILABLAAA TUTATTILAA LAAN BUADLAATTUTTILABA TUTABAATA UNABATTITIA AATA UNABATTA UNABATTUTA AATA UNA Suul CATAGGTCACATTCATTTACATCACTAAGAAGAGATTGCAATCTGTTAGGTCAAGTTTAATTTAAATGTGCTGCACCTTTGATGTT	
	Pstl, homeo box	2340	WINGUI A CONTRACTOR CO	
	ea lagi uu evoi lugi amiis paelevoi ingi yang iyy evi tarah aproangi eva taayolevoi erataiyyi eva talevoi yi tarah Aggaggaggi tagaggaggaggaggaggaggaggaggaggaggaggagga	2430	TODI TAGECCATATTTCCGTCCCTTTCTAGAATGAACGAAAACAGTATCTGGTTTTCCCGAAAAATCTTATGAATTTAAAAATGCACTTTATTG	
	aGinvallys 11eTrpPhelysAsnArgArgArgArgMislys 11eGInSerNapG InnislysAspGInSerTyrG1uG1yMetProle CCAGGTGAAGATATGGTTTAAGAACCGTCGGCGTCGTCACAAGATCCAATCGATCAGCACCAGGACCAGTCCTACGAGGGATGCCTCT	2520	Nry1 Sac11 cacatactcacatgcctgccataanatatgattcccgatitttcccggaacatcactaanacatttggaccagctgctgct	
	userProßlyMetLysGinSerAspGiyAspProProSerLeuGinThrleuSerLeuGiyGiyGiyAiaThrProAsnAiaLeuThrPr CTCCCCGGGTATGAAACAGAGCGATGGCGATCCCCCCAGCTTGCAGACTCTTAGCTTGGGTGGAGGAGCCACGCCCCAACCTTGACTC	2610	Acc1 BTTTATTCACCTACCTGAAACCCATACTCTTATCGCCTGATCCTCGCGGCGGTCGCACTATTTAGGTAGACACTGTACAGGCAGCACTAGC	5130
	G			

Fig. 3. DNA sequence of the *bcd* gene and corresponding amino acid sequence of putative *bcd* gene products. The genomic DNA sequence from the first *Sall* site upstream of the 5' end to ~ 300 bp beyond the 3' end [site of poly(A) addition] of nearly full-length cDNAs is numbered from 1 to 5130 nt. Above the DNA sequence, the amino acid sequence corresponding to the longest ORF is shown. Deviations from the genomic sequence found in one of the sequenced cDNAs are indicated below (nucleotide) and above (amino acids) the corresponding positions. Of the few nucleotide changes between genomic and cDNAs, three result in altered amino acids, two in exon 3 where genomic Pro and Phe are changed to Ala in c53.46.6 and Ser in c53.46.9, respectively, and one in exon 4 where a Leu is changed to a Met in c53.46.6. Since at least one cDNA sequence is identical to the genomic sequence at these positions, we consider it more probable that these changes are due to mistakes during cDNA synthesis rather than due to true polymorphisms. Intron boundaries are marked by vertical lines, the alternative splice acceptor sites of the second intron by dashed vertical lines. The positions of the 5' and 3' ends of c53.46.6 and 7 coincide with and include a stretch of poly(A). The polyadenylation signal is underlined. The sequences corresponding to the PRD-repeat and the homeo domain are boxed.

divergence to nucleotides 2650 (bcd^{E1}) and 2400 (bcd^{E2}) within the 3rd exon. The size reduction of the respective *Hind*III restriction fragments in genomic Southern blots suggests deletions of ~180 bp (bcd^{E1}) and 260 bp (bcd^{E2}) (Figure 1b). Both deletions, based on their size and position, must include parts of the homeobox domain. They result in the strong *bcd* mutant phenotype.

bcd transcript distribution in oocytes and early embryos

The distribution of *bicoid* transcripts was determined in oocytes and early embryos by *in situ* hybridization using ³⁵S-labelled *bcd* anti-sense RNA probes.

The wild-type ovaries *bcd* mRNA can be clearly detected during vitellogenic stages of oogenesis and in small oocytes

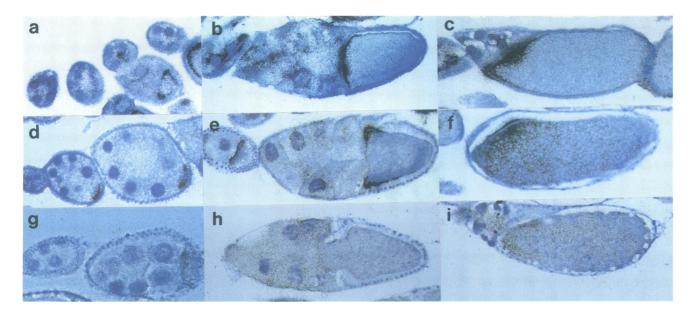


Fig. 4. Distribution of *bcd* transcripts in individual egg chambers of wild-type and mutant ovaries. *bcd* anti-sense RNA probe was hybridized *in situ* to sectioned ovaries of (a,b,c) OregonR, (d,e,f) swa¹⁴/Df(1)JF5 and (g,h,i) exu^{PJ}/exu^{QR} females. Photomicrographs show egg chambers at stage 5-7 (left), 10 (middle) and stages 13/14 (right). The anterior of the oocyte is always oriented to the left.

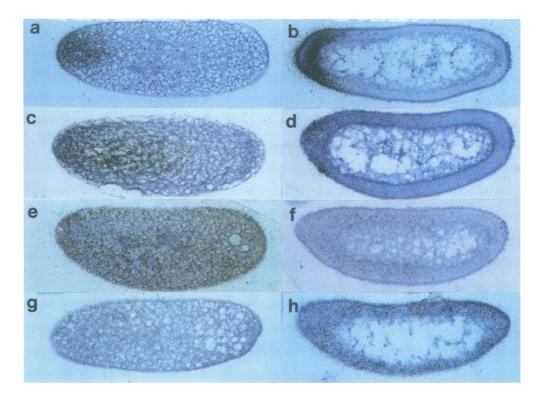


Fig. 5. Distribution of *bcd* transcripts in wild-type and mutant embryos. Photomicrographs show embryo sections that were prepared and hybridized under similar conditions as the ovarian sections shown in Figure 4. In all cases anterior is to be left and dorsal uppermost. The maternal genotypes of the embryos correspond to that of the oocytes in Figure 4.

it is already localized anteriorly (Figure 4a). During subsequent stages of oogenesis, intensive label is observed around the nurse cell nuclei and most strongly at the anterior edges of the developing oocytes. By contrast, virtually no hybridization is observed within the oocyte, the remaining cortical regions or in the follicle cells (Figure 4b). The pattern of transcription thus confirms the conclusions drawn from pole cell transplantation experiments, that the *bcd* gene is expressed in the germline (Frohnhöfer, 1987).

In very early embryos (stage 1 and 2, intravitelline cleavage) the *bcd* mRNA is localized in a cone-shaped region in the anterior 20% of the egg cytoplasm. The decrease in *bcd* mRNA concentration toward the more posterior regions appears to be abrupt rather than graded and at regions up to 80% egg length (0% = posterior pole) almost no label can be detected above background level (Figure 5a). Con-

current with the migration of the nuclei of the egg cortex (stages 3 and 4), the *bcd* mRNA becomes concentrated in the cortical cytoplasm in the form of an anterior cap (Figure 5b). It remains present in the cortical plasm throughout the perivitelline cleavage stages, and then disappears after the last cleavage division during the period of nuclear elongation (stage 5a) before blastoderm cells are formed.

We have also analyzed the distribution of *bcd* mRNA in the 11 alleles by *in situ* hybridization to *bcd* mutant embryos. We observed essentially normal *bcd* mRNA distribution in all 11 alleles, although the absolute amount was very much reduced in the allele bcd^{33-5} (data not shown).

Mutants affecting the localization of bcd RNA

The mechanism which localizes the *bcd* mRNA anteriorly presumably involves factors which act *in trans* on the *bcd* mRNA. Mutations that affect the localization of transplantable bcd^+ activity, like *swa* and *exu*, were therefore likely candidates for functions exerting their effects via defective *bcd* mRNA localization (Frohnhöfer and Nüsslein-Volhard, 1987). To test this hypothesis, we analyzed the distribution of *bcd* mRNA in ovaries and early embryos from *swa* and *exu* mutant females.

In *swa* ovaries (Figure 4d-f), *bcd* RNA appears to be transcribed at normal levels and is also localized anteriorly until mid oogenesis (about stage 10,11). At late stages of oogenesis and at early embryonic stages, strict localization is abolished and *bcd* transcripts become distributed throughout the egg. However, the concentration remains higher in the anterior half such that a shallow gradient is formed (Figure 5c,d).

In *exu* oocytes, *bcd* transcripts are never clearly localized and at late oogenic and early embryonic stages they appear essentially uniformly distributed (Figures 4g-i, 5e). During syncytial blastoderm stages, a reduction in the concentration of *bcd* mRNA in the posterior part of the *exu* embryos is observed resulting in a shallow concentration gradient with a high point at the anterior egg pole (Figure 5f), similar to that seen in *swa* embryos.

Genetic evidence suggests that the decrease in the concentration of *bcd* mRNA in the posterior region of the egg observed in swa and late exu embryos is dependent on the activity of the genes of the posterior center (Nüsslein-Volhard et al., 1987; Frohnhöfer and Nüsslein-Volhard, 1987). The elimination of this activity by mutation (e.g. in exu - vasaor exu - osk double mutants) leads to the formation of ectopic anterior structures at the posterior egg pole (Schüpbach and Wieschaus, 1986; Frohnhöfer and Nüsslein-Volhard, 1987). We have therefore analyzed the distribution of bcd mRNA in double mutants of exu and vasa. While young vasa -exuembryos show the same low and uniform *bcd* mRNA level as exu embryos, the label does not disappear from the posterior pole at later stages (Figure 5g,h). This indicates that the posterior activity destablizes the *bcd* mRNA, and explains the inhibitory effect on anterior development in transplantation experiments of posterior pole plasm to the anterior (Lehmann, 1985; Frohnhöfer et al., 1986. Nüsslein-Volhard et al., 1987).

Discussion

The most striking feature of the *bicoid* gene is the anterior localization of a transplantable, head-inducing activity

(Frohnhöfer and Nüsslein-Volhard, 1986). The distribution of bcd mRNA under various genetic conditions mimics that of the bcd^+ activity inferred from transplantation experiments and fate mapping studies (Frohnhöfer and Nüsslein-Volhard, 1986, 1987). This suggests that the bcd^+ activity is the mRNA. As the distribution of the bcd protein product is directly determined by that of the RNA (Driever and Nüsslein-Volhard, 1988), a contribution of the protein to the transplantable activity cannot be ruled out. Injection experiments using pure bcd RNA and bcd protein respectively will be required to resolve this issue.

The 8.7-kb fragment employed for the p-element mediated transformation contains all the sequences required for correct stage- and tissue-specific expression of the gene. As the *bcd* gene is transcribed in all nurse cells long before the transcript is translated (Driever and Nüsslein-Volhard, 1988), we do not expect stringent transcriptional control such as that of zygotic segmentation genes for which 5 control regions far exceeding 4 kb have been reported (Hiromi *et al.*, 1985). Instead, much of the region specific function of the *bcd* gene is achieved by the intracellular localization of its mRNA and by translational control (Driever and Nüsslein-Volhard, 1988).

bcd transcripts are detected in the nurse cells of the growing follicles, while no transcripts can be detected within or surround the oocyte nucleus [as, e.g. in the case of fs(1)K10 (Haenlin et al., 1987)]. Within the nurse cell complex, we do not observe a spatially differential pattern of transcriptional activity, and the simplest interpretation of our in situ data is that the bcd gene is transcribed in all nurse cells and transported to the anterior of the oocyte together with the bulk of maternal mRNA via the ring canals. Here, the bcd RNA appears to be specifically trapped and it remains localized at the anterior of the oocyte and the early embryo. This anchoring process requires a special mechanism in which presumably cis- and trans-acting factors are involved. We expect the *bcd* RNA to contain sequences which are recognized by specific proteins involved in the anchoring process.

It will be interesting to unravel the structural features of the bcd mRNA required for localization, e.g. by in vitro mutagenesis. The *bcd* alleles all show apparently normal RNA localization. As a fortuitous piece of evidence, the deletions in the two alleles bcd^{E1} and bcd^{E2} show that the region immediately adjacent to the homeodomain is not important for mRNA localization. The bcd mRNA sequence does not reveal any striking feature which may be suggestive of a role in specific localization. In computer searches for secondary structures within the RNA no large stretches of high selfcomplementarity were obvious which could provide loop formation potentially required for specific anchoring complexes. However, complex modes of secondary structure involving interrupted stretches of self-complementarity would not have been detected. Two thirds of the transcribed sequence provide an open reading frame and are presumably translated into a 54-kd protein.

The simplest mechanism for achieving *bcd* mRNA localization is the formation of a specific RNA – protein complex which is immobilized in the oocyte and later in the embryo by binding to the cytoskeleton of the egg cell. Formation of such a complex would prevent diffusion and provide mRNA stabilization. The gene products of the maternal defect genes *exu* and *swa* are excellent candidates

for proteins involved in forming complexes with the bcd RNA and thereby trapping it at the anterior end of the oocyte upon entry. Phenotypic analysis and transplantation experiments have suggested that mutations in exu and swa result in a low level and almost homogenous distribution of the bcd^+ activity (Schüpbach and Wieschaus, 1986). Here we show that it is the distribution of the bcd mRNA that is affected in these mutants. The phenotypes of double mutants between exu and bcd suggest a very specific involvement of exu in complexing bcd mRNA: in a bcd mutant background, exu is without phenotypic effect. swa mutants on the other hand, have phenotypic features which are not suppressed in bcd mutants (Frohnhöfer and Nüsslein-Volhard, 1987). Abnormalities of nuclear migration in early swa embryos at low temperature suggests that the swa product might be a component of the cytoskeleton (Frohnhöfer and Nüsslein-Volhard, 1987). An attractive hypothesis is that the exu protein binds both the bcd RNA as well as the swa protein, and the swa protein, and the swa protein is in turn attached to the cytoskeleton of the oocyte. If an excess of both the *exu* and *swa* products is distributed evenly in the oocyte, bcd mRNA transported into the oocyte from the anteriorly located nurse cells would immediately be trapped in the anterior region of the oocyte. No additional mechanism for localization would be required and the polarity of the oocyte would thus directly determine that of the embryo.

In *exu* mutant embryos, *bcd* mRNA is present initially up to the very posterior tip of the egg reaching the realm of action of the genes for the posterior centre. Later, in syncytial blastoderm, a shallow gradient, presumably caused by degradation of the mRNA in the posterior half of the embryo, is formed. Elimination of the posterior activity, in our case by employing the mutant *vasa*, leaves the *bcd* mRNA intact and thus explains the formation of anterior structures at the posterior egg pole observed in *exu* vasa embryos (Schüpbach and Wieschaus, 1986; Frohnhöfer and Nüsslein-Volhard, 1987). This suggests that the inhibitory influence of the posterior center on *bcd*⁺ activity observed in transplantation experiments (Nüsslein-Volhard and Frohnhöfer, 1987; Lehmann, 1985) is mediated by affecting mRNA stability.

In other dipteran species (Chironomous sp.) anterior development can be supressed by various treatments of the anterior egg pole (Yajima, 1960; Kalthoff and Sander, 1968; Kandler-Singer and Kalthoff, 1976). As a consequence, the embryos develop a bicaudal (not a bicoid) phenotype. Kalthoff and co-workers (Kalthoff and Sander, 1968; Kandler-Singer and Kalthoff, 1976; Kalthoff, 1979) characterized an 'anterior determinant' as a UV- and RNasesensitive entity localized in a cone-shaped form at the anterior egg pole of Smittia embryos in cleavage stages. Using a mutant condition in Chironomous sp., it was suggested that isolated poly(A)⁺ RNA could restore normal development when injected into the anterior egg pole (Kalthoff and Elbetieha, 1986). Although a comparison of data from different species obtained with different methods is difficult, it seems likely that the bcd^+ activity corresponds to the 'anterior determinant' of Kalthoff. The reason why in other dipteran species the elimination of the anterior activity causes complete mirror image duplications of the posterior pattern is not obvious. In Drosophila, bcd elimination only causes duplication of the telson, while the abdomen has normal

polarity. Transplantation experiments only show that the tendency for abdominal development of the anterior egg region is greatly increased in bcd^- embryos.

Materials and methods

DNA and RNA techniques

The genomic clones containing *bcd* have been obtained by Scott *et al.* (1983). For transformation, the 8.7-kb *Eco*RI fragment containing the *bcd* transcription unit was cloned into the P^{PA} vector (Goldberg *et al.*, 1983). Embryos from a Adhⁿ², *bcd*^{E1}/TM3 stock were injected with the DNA. Transformants were selected on 6% ethanol and the sites of insertion determined using standared genetic and *in situ* hybridization techniques.

The isolation of *bcd* cDNAs from a λ gt 10 cDNA library of poly(A)⁺ RNA from 0–4 h old embryos has been described (Frigerio *et al.*, 1986). Southern blot analysis was performed according to standard procedures, probing with a *bcd* cDNA prove c53.46.6. S1 nuclease protection assays were performed according to Murray (1986). For Northern analysis, poly(A)⁺ RNA (2 µg each), isolated from follicles, 0–2 h, 2–4 h, 4–8 h, 8–12 h, 12–17 h old embryos, from first instar larvae (1–2 days), from early (3–4 days) and late (6–7 days) third instar larvae, and from adult females, as previously described (Frei *et al.*, 1985), were run in a 1.1% agarose gel containing formaldehyde. The RNA was transferred to a nitrocellulose filter and hybridized with nick-translated c53.46.6 DNA as described (Kilchherr *et al.*, 1986).

DNA was sequenced by the dideoxynucleotide method (Sanger *et al.*, 1977; Barnes *et al.*, 1983) using the M13 vectors mWB3296 or its counterpart, mWB3226, which contains the M13mp18 polylinker in opposite orientation. Both were derived from M13 vector as described (Hiromi *et al.*, 1985). Most genomic and cDNA sequences have been read on both strands. The minor fraction (5%) of genomic DNA (cDNA) sequences determined only on one strand has been analyzed on both strands of the cDNAs (genomic DNA).

In situ hybridization

In situ hybridization was performed according to Ingham *et al.* (1985). Paraffin sections of ovaries and embryos were hyridized with ³⁵S-labelled anti-sense RNA probes obtained by transcription of cDNA c53.46.6 cloned in bluescript vector using T3 RNA polymerase.

Fly strains; mutants

The wild-type strain was Oregon R. Mutant alleles were swa^{14} (Gans *et al.*, 1975) in trans to Df(1)JF5 (Stephenson and Mahowald, 1987), exu^{PJ} and exu^{QR} (Schüpbach and Wieschaus, 1986) and bcd^{E1} , bcd^{E2} , bcd^{GB} , $bcd^{33\cdot5}$, $bcd^{111\cdot3}$, bcd^{E4} , $bcd^{23\cdot16}$, bcd^{-13} and bcd^{085} (Frohnhöfer and Nüsslein-Volhard, 1986). For the *bcd* alleles, eggs were collected from females heterozygous for the respective allele and Df(3R)LIN (Frohnhöfer and Nüsslein-Volhard, 1986). All fly stokes carried suitable markers.

Acknowledgements

We thank Matthew Scott for the clones of the *Antp*-walk, our colleagues for interest and discussion and D.Kaiser and V.Koch for secretarial assistance. This work was supported by Swiss National Science Foundation Grant No. 3.600-0.84 and 3.348-0.86 to M.N. and the Leibniz program of the Deutsche Forschungsgemeinschaft (C.N.-V).

References

- Barnes, W.M., Bevan, M. and Son, P.H. (1983) *Methods Enzymol.*, 101, 98-122.
- Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986) Cell, 47, 1033-1040.
- Doyle,H.J., Harding,K., Koey,T. and Levine,M. (1986) Nature, 323, 76-79.
- Driever, W. and Nüsslein-Volhard, C. (1988) Cell, in press.
- Frei, E., Baumgartner, S., Edström, J.-E. and Noll, M. (1985) *EMBO J.*, 4, 979-987.
- Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M. (1986) Cell, 47, 735-746.
- Frohnhöfer, H.G. (1987) Thesis, University of Tübingen.
- Frohnhöfer, H.G. and Nüsslein-Volhard, C. (1986) Nature, 324, 120-125.

Frohnhöfer, H.G. and Nüsslein-Volhard, C. (1987) Genes Dev., 1, 880-890.

Frohnhöfer, H.G., Lehmann, R. and Nüsslein-Volhard, C. (1986) Embryol. Exp. Morphol., 97a, 169-179.

- Gans, M., Audit, C. and Masson, M. (1975) Genetics, 81, 683-704.
- Goldberg, D.A., Posakony, J.W. and Maniatis, T. (1983) *Cell*, **34**, 59–65. Haenin, M., Roos, C., Cassab, A. and Mohier, E. (1987) *EMBO*, **6**, .
- 801-807. Hiromi,Y., Kuroiwa,A. and Gehring,W.J. (1985) *Cell*, **43**, 603-613. Ingham,P.W., Howard,K.R. and Ish-Horowicz,D. (1985) *Nature*, **318**,
- 439-445.
- Kalthoff, K. (1979) Symp. Soc. Dev. Biol., 37, 97-126.
- Kalthoff, K. and Sander, K. (1968) Wilhelm Roux's Arch. Dev. Biol., 161, 129-146.
- Kalthoff,K. and Elbetieha,A. (1986) *Embryol. Exp. Morphol.*, **97** Suppl., 81–196.
- Kandler-Singer, I. and Kalthoff, K. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 3739–3743.
- Kilchherr, F., Baumgartner, S., Bopp, D., Frei, E. and Noll, M. (1986) *Nature*, **321**, 493–499.
- Lehmann, R. (1985) Ph.D. Thesis, Tübingen.
- Lehmann, R. and Nüsslein-Volhard, C. (1986) Cell, 47, 141-152.
- McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A. and Gehring, W.J. (1984) Nature, **308**, 428-433.
- Mohler,H. and Wieschaus,E.F. (1985) Cold Spring Harbor Symp., 50, 105-112.
- Murray, M.G. (1986) Anal. Biochem., 151, 165-170.
- Nüsslein-Volhard, C. (1977) Wilhelm Roux's Arch. Dev. Biol., 52, 166.
- Nüsslein-Volhard, C., Frohnhöfer, H.G. and Lehmann, R. (1987) *Science*, **238**, 1675–1681.
- Sander, K. (1976) Adv. Inst. Physiol., 12, 125-238.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467.
- Scott, M.P., Weiner, A.J., Hazelrigg, T.I., Polisky, B.A., Pirotta, V., Scalenghe, F. and Kaufman, T.C. (1983) *Cell*, **35**, 763-776.
- Schüpbach, T. and Wieschaus, E. (1986) Wilhelm Roux's Arch. Dev. Biol., 195, 302-317.
- Stephenson, E. and Mahowald, A.P. (1987) Dev. Biol., 24, 1-8.
- Wakimoto, B.T., Turner, F.R. and Kaufman, T.C. (1984) Dev. Biol., 102, 147-172.
- Wharton, K.A., Johansen, K.M., Xu, T. and Artavanis-Tsakonas, S. (1985) *Cell*, **43**, 567–581.
- Yajima, H.J. (1960) Embryol. Exp. Morphol., 8, 198-215.

Received on February 11, 1988