Volume 17 Number 12 1989 Nucleic Acids Research

Characterization of a putative transcription factor gene expressed in the 20-OH-ecdysone inducible puff 74EF in *Drosophila melanogaster*

R.Janknecht, W.Taube, H.-J.Lüdecke+and O.Pongs*

Lehrstuhl für Biochemie, Ruhr-Universität Bochum, Universitätsstrasse 150, D-4630 Bochum, FRG

Received April 4, 1989; Revised and Accepted May 25, 1989 EMBL accession no. X15087

ABSTRACT

Drosophila melanogaster DNA has been cloned which encompasses the major part of the 20-OH-ecdysone inducible puff 74EF. One 20-OH-ecdysone responsive transcription unit was detected which gives rise to two alternative transcripts. The expression of one transcript in salivary glands of 3rd instar larvae is correlated with the 20-OH-ecdysone induced activity of puff 74EF. Corresponding cDNA analysis indicates that the two transcripts are translated into two different proteins which have alternative amino terminal ends. The carboxy terminal domain of the *74E* proteins is similar to the carboxy terminal sequences of *ets*-oncoproteins suggesting that the *74E* proteins represent alternative transcription factors. It is proposed that the activity of the 20-OH-ecdysone inducible puff 74EF leads to a switch in the synthesis of a transcription factor.

INTRODUCTION

20-OH-ecdysone is one of the key hormones in the ontogenesis of *Drosophila melanogaster* [1]. A sharp rise in the 20-OH-ecdysone titre at the end of the development of 3rd instar larvae is correlated with the onset of pupariation. The increase in ecdysteroid hormone titre is responsible for the developmental control of gene expression which is also manifest in a 20-OH-ecdysone dependent puffing activity, visible on the polytene chromosomes in salivary glands of 3rd instar larvae [2]. *In vitro* this puffing activity can be reproduced by culturing puff stage 1 salivary glands in the presence of 20-OH-ecdysone [3]. This results (i) in the regression of intermolt puffs (e.g. at 68C), (ii) in the rapid induction of early puffs (e.g. at 74EF and 75B), and (iii) in the delayed induction of late puffs (e.g. at 78D) [4]. Products of genes expressed in early puffs under the control of 20-OH-ecdysone are necessary for the activation of late puffs [5]. Also, the gene expression in early puffs is apparently autoregulated as shown by studies with *Drosophila* mutants aneuploid for the chromosomal region containing the 74EF and 75B loci [6]. A combination of affinity labeling experiments with indirect immunofluorescence microscopy has demonstrated that ecdysteroid binds *in vivo* and *in vitro* to active puffs [7,8], probably via a 20-OH-ecdysone receptor complex [9].

It has been proposed that early puffs are induced by binding 20-OH-ecdysone-receptor complex and that this binding activates gene expression in early puffs [6]. Recently, a more elaborate hypothesis has been put forward which predicted that the products of 20-OH-ecdysone inducible genes in puffs 74EF and/or 75B are transcription factors [10]. These transcription factors would regulate in conjunction with the 20-OH-ecdysone receptor complex the transient puffing activity during the metamorphosis of 3rd instar larvae to pupae.

© IRL Press

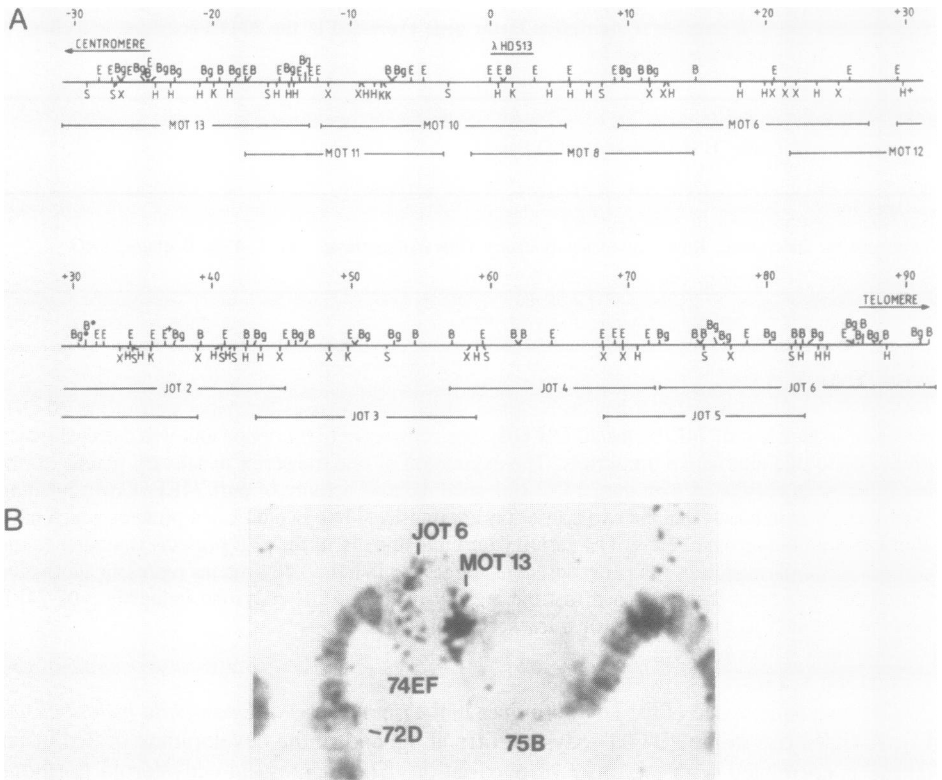

Fig. 1: (A) Restriction map of the isolated genomic DNA. The distance is measured in kb relative to the start site of the chromosomal walk with negative/positive numbers indicating the direction towards the centromere/telomere. Isolated DNA fragments are shown below the restriction map. E=EcoRI, B=BamHI, Bg=BglII, H=HindIII, K=KpnI, S=SalI, X=XbaI. Restriction enzyme sites labeled by + or * were only detected in OregonR or CantonS genomic DNA. (B) *In situ* hybridization of ^{3}H-labeled JOT6 and MOT13 DNA to polytene chromosomes. Only the part of the third chromosome showing hybridization signals is presented.

This report describes the isolation of 123 kb genomic *Drosophila melanogaster* DNA corresponding to that region on the third chromosome where the 74EF puff is induced by 20-OH-ecdysone. A cDNA has been characterized which is derived from a 20-OH-ecdysone regulated transcription unit localized in the cloned 74EF DNA. The deduced protein sequence contains a domain which is similar in sequence to *ets*-oncoproteins. We suggest that in the early puff at 74EF 20-OH-ecdysone regulates the expression of a transcription factor gene.

MATERIALS AND METHODS

Libraries

Two *Drosophila melanogaster* genomic DNA libraries were used: one was constructed with sheared CantonS DNA (kindly provided by Maniatis *et al.* [11]), the other with OregonR DNA partially digested with Sau3AI and inserted into the BamHI site of EMBL4 [12]. This library was made available to us by H. Jaeckle (Tübingen, FRG). For the isolation

of cDNA two *Drosophila melanogaster* cDNA libraries were screened: one consisted of oligo(dT)-primed 1–2d old OregonR pupa cDNA inserted into lambda gt10, the other of oligo(dT)-primed adult strain Berlin head cDNA inserted into lambda gt11. cDNA libraries were a gift of T. Kornberg (San Francisco, USA) and E. Buchner (Würzburg, FRG), respectively. Libraries were screened according to Benton and Davis [13].

Cloning of DNA

DNA was manipulated according to Maniatis *et al.* [14] using ER1 host-vector systems under L1 containment conditions, as defined in the guidelines of the Federal German Government for recombinant DNA research. pAT153, M13mp8, M13mp10 and pBluescript KS^+ or SK^+ (Stratagene) were used as vectors for subcloning.

Radioactive labeling of DNA

For *in situ* hybridizations, DNA probes were labeled with ^{3}H-dCTP by nick-translation [15]. For plaque hybridizations, DNA probes were labeled with ^{32}P-dCTP by nick-translation or by oligonucleotide primed DNA synthesis [16]. For Northern hybridizations, probes were labeled by the latter method.

Isolation of RNA

Drosophila melanogaster OregonR puff stage 1 larvae were dissected and their salivary glands incubated *in vitro* according to Ashburner [3]. RNA prepared from those salivary glands was a gift of G. Feigl (Bochum, FRG). *Drosophila melanogaster* OregonR RNA of different developmental stages was isolated by the guanidinium-isothiocyanate method and poly(A)$^+$-RNA was prepared by oligo(dT)-cellulose chromatography [14].

Northern blots

12μg of salivary gland RNA or 4μg of poly(A)$^+$-RNA of different developmental stages were electrophoretically separated in 1% agarose gels in 2.2 M formaldehyde, 50 mM 3-[N-morpholino]propanesulfonic acid pH 7.0 and 1 mM EDTA [17]. The RNA was transferred to nylon membrane (Nytran, Schleicher and Schuell, or Biodyne B, Pall) by overnight diffusion-blotting with 20×SSC. The RNA was fixed to the nylon membrane by baking for 2h at 80°C.

Hybridization procedures

In situ hybridization of ^{3}H-labeled DNA to polytene chromosomes of *Drosophila melanogaster* has been described previously [18]. Northern blots were prehybridized in 50% formamide, 750 mM NaCl, 150 mM Tris/HCl pH 7.9, 5 mM EDTA, 0.1% SDS, 5×Denhardt's solution and 0.1 mg/ml denatured salmon sperm DNA for 2h at 42°C. Hybridization took place under the same conditions for 24–36h with ^{32}P-labeled DNA probes. Blots were washed in 1×SSC, 0.1% SDS followed by several washes in 0.1×SSC, 0.1% SDS at 65°C before autoradiography.

DNA-Sequencing

DNA was sequenced by the dideoxynucleotide chain-termination method [19] employing ^{35}S-dATP and using the Sequenase Kit (United States Biochemical) or the T7 Sequencing Kit (Pharmacia). Subclones for sequencing were generated either by cloning of restriction fragments or by constructing DNaseI deletion subclones [20].

RESULTS

Previously, we had isolated 24 kb of *Drosophila melanogaster* DNA which hybridized *in situ* to locus 74F [21]. This DNA was used as a starting point for the chromosomal walk summarized in Fig. 1A. 123 kb of *Drosophila melanogaster* genomic DNA was isolated either from an OregonR (MOT6,8,10,12,13 and JOT6) or a CantonS genomic

74E

```
 taacaaatactcgcagcggttttttgtgctgactttagattttagttttgctgataccgtaagagataaatgacgtgccgcggcgagcga  -   445
cgaattagcggcccaaaacagagaaaaactcaacagccaaaacacacacagaataccagagaaacagaagcaagcaacaagcaaatgttt  -   355
gacccgcgggaaaacaaaatgtaatagctaatgtagcaaaatgtgtttaagtttaagccagaaagttagcatataaaacgtccacaatta  -   265
gtgcaaaaaagttccaaaaacctgatcaattttcgaaatcaaacaaatacctcaagcgtaaagaagtaacaataaaccaacaagcaaaca  -   175
aacaaacaagcaacgaaccgagtgcaaaaacttagtgcaatttctttaataagaaaatttgaagccccaaaaaaacagacataataatc  -    85
atgtctagtgtgtagaacgttgcgtatacgcaatattgccaaacccattaacgcagacatcacgcatacgccgcgtcgtgcaaaatgatt  +     6
                                                  ●●●
                                                                                       M  I      2
atggtgcaacatttggtggccgcctcggcgcacaacttcgccagccaggcggcggcatcgcttgtcaatgtgtccagctccagctctagt  +    96
 M  V  Q  H  L  V  A  A  S  A  H  N  F  A  S  Q  A  A  A  S  L  V  N  V  S  S  S  S  S  S     32
tccagttcgagttcctcgtcctcgctgtcgctgtcttcatcctcatccagctcctcgctgtcctcggccacGCCCACGCCGGTGGCCTCG  +   186
 S  S  S  S  S  S  S  S  L  S  L  S  S  S  S  S  S  S  S  L  S  S  A  T  P  T  P  V  A  S     62
CCCGTGACGCCCACATCTCCACCGCCGGCAGCCGCCGCACCCGCAGAAGCCTCGCCGCCAGCAGGAGCTGAGTTGCAGGAGGATGGGCAG  +   276
 P  V  T  P  T  S  P  P  P  A  A  A  A  P  A  E  A  S  P  P  A  G  A  E  L  Q  E  D  G  Q     92
CAGGCTAAGACTCAGGAGGATCCTACTATGAAGGACCAAGACATGCTGGAAAAGACGAGGCAAGAGGTGAAGGATCCAGTTAATGTGGAG  +   366
 Q  A  K  T  Q  E  D  P  T  M  K  D  Q  D  M  L  E  K  T  R  Q  E  V  K  D  P  V  N  V  E    122
GAACCAGGTGCCATAGTTGATACCGAGTCAGTGATGGCGCGGCAAAGTCCAAGTCCAGTTGCGTCCACCAAAGTCCCCGAGTCACTAGAA  +   456
 E  P  G  A  I  V  D  T  E  S  V  M  A  R  Q  S  P  S  P  V  A  S  T  K  V  P  E  S  L  E    152
GAGATTTCCAACAAATCGCCGCCCGTTCAAGAGGATGAAGAGGAGTCTGAATCCGTAGCCTCCGACTGTCGCGAGTTCAAAGTGCTCTAT  +   546
 E  I  S  N  K  S  P  P  V  Q  E  D  E  E  E  S  E  S  V  A  S  D  C  R  E  F  K  V  L  Y    182
AATCACCTAAGGCAGCAGCAACACCACCACTCTCCCTCCTCGCCAGACAAGACGCGCAGCACCCTGGACGATGTGTCCAAGATCCTCTGG  +   636
 N  H  L  R  Q  Q  Q  H  H  H  S  P  S  S  P  D  K  T  R  S  T  L  D  D  V  S  K  I  L  W    212
GAGCGAAAGCAGCAGCTGCAGCGCAGCTCAGTGATTACCGCAGCACCCACTCTGCAGCCGCAGCAGCACCAGCAGCCCATGAGCGACATC  +   726
 E  R  K  Q  Q  L  Q  R  S  S  V  I  T  A  A  P  T  L  Q  P  Q  Q  H  Q  Q  P  M  S  D  I    242
GAGGACGAGGAGACGCTGGAGGATGTCGATGACGCGGATGCGGACGTGGAGGCAGACGCCGAGGACGAGGAGCTACTGGAGCAGTACCAA  +   816
 E  D  E  E  T  L  E  D  V  D  D  A  D  A  D  V  E  A  D  A  E  D  E  E  L  L  E  Q  Y  Q    272
AACGGCTACGACTCGCCACTGGATCTGTCCCTGGGTGGAGCCACCAGTGCAGCCGCATCAGCTGCCGCTGCCGCCTCGGCGGTCAGTCGC  +   906
 N  G  Y  D  S  P  L  D  L  S  L  G  G  A  T  S  A  A  A  S  A  A  A  A  A  S  A  V  S  R    302
CGACGGGGACGCACCTACTCCGGCACGGAATCCGACGACTCCGCGCAGTGCGAACGGGCGAGGATGCGACTCAAGCCGGAACGCAAGGCG  +   996
 R  R  G  R  T  Y  S  G  T  E  S  D  D  S  A  Q  C  E  R  A  R  M  R  L  K  P  E  R  K  A    332
                                              ▼
GAGAGATCGGCGGCCTACAAGAAGAGCTTGATGAAGCGATATTACACCGAAATTCCTATTGTCAAGCAATCGACAAGTCCGGCACCGCAG  +  1086
 E  R  S  A  A  Y  K  K  S  L  M  K  R  Y  Y  T  E  I  P  I  V  K  Q  S  T  S  P  A  P  Q    362
CAGCAACTGCAGCAGCAGCACCACCTGCAACAGCAGCAACAGCAGCAGCCGCACAACGGCTCCACCTTCGCCGGCGCCACCGCTCTGCTC  +  1176
 Q  Q  L  Q  Q  Q  H  H  L  Q  Q  Q  Q  Q  Q  P  H  N  G  S  T  F  A  G  A  T  A  L  L    392
CACATAAAGACGGAGCAAAATACGCTGCTCACTCCGCTCCAGCTGCAGCAGCAGCAACAACAGCAGCAGGGTCTACACGGTGCGGCAGGC  +  1266
 H  I  K  T  E  Q  N  T  L  L  T  P  L  Q  L  Q  Q  Q  Q  Q  Q  Q  Q  G  L  H  G  A  A  G    422
AACGGAGGCAGCAGCAACGGCAACAATGCCCACCAGCAGCAACAACCGCTGGCCATCCCACAACGCCCGCTGCTCCACAATCTGCTTAGC  +  1356
 N  G  G  S  S  N  G  N  N  A  H  Q  Q  Q  Q  P  L  A  I  P  Q  R  P  L  L  H  N  L  L  S    452
                                                       ▼
GGCGGTGCCATCCACAATCCGCACCACAGAAACTACACGACGGCCACAACAGGTTCATTTCCGCCCAGTCCCGCGGACAGCGGCGTCTCC  +  1446
 G  G  A  I  H  N  P  H  H  R  N  Y  T  T  A  T  T  G  S  F  P  P  S  P  A  D  S  G  V  S    482
GACGTGGACAGCTCCAGTTCCGGCGGCCAGCCCTGCGCCGACGAGCTGAAGGCGCGGCTCGGGATGCCACCAGCCACCTCGGCGTCGGCA  +  1536
 D  V  D  S  S  S  S  G  G  Q  P  C  A  D  E  L  K  A  R  L  G  M  P  P  A  T  S  A  S  A    512
GCGGCCGCGGCAGCAGCGGCGGCGGCGGCAGCGGCACACCTGCACACGGGCACCTTTCTGCATCCGAATTTATACCAGAACAATGCGGCA  +  1626
 A  A  A  A  A  A  A  A  A  A  A  A  H  L  H  T  G  T  F  L  H  P  N  L  Y  Q  N  N  A  A    542
                                       ▼
AATTCGCTGCGGAACATATGGAATCGCAGTGTTGGGGTGCCCGACAACTACTACGGAAGCAGTGGCGCCGGCAGTGGGGGCACCCAGCCA  +  1716
 N  S  L  R  N  I  W  N  R  S  V  G  V  P  D  N  Y  Y  G  S  S  G  A  G  S  G  G  T  Q  P    572
GGTGGTCCGGGAAACCCCCAAACGCCTGGCTACCTGACAACGTCCTACTTCAACGCACCGACAGCGGCAACGGCAGCGGCGTCCCAGCGA  +  1806
 G  G  P  G  N  P  Q  T  P  G  Y  L  T  T  S  Y  F  N  A  P  T  A  A  T  A  A  A  S  Q  R    602
GGAACCACAATCAATGGCTATCATTCCCTGCACCAGCAGCAGCAGCAGCAACAGCAGTCGCAGCAATCGCAGCAGCAACAGCAACTGGCC  +  1896
 G  T  T  I  N  G  Y  H  S  L  H  Q  Q  Q  Q  Q  Q  Q  S  Q  Q  S  Q  Q  Q  Q  Q  L  A    632
CACCAGCAACTGTCGCACCAGCAGCAGCAGGCGTTGCATCAGCAACTGTCGCACCAACAGCAGCAGCAGCAACAGCAACAGCAGCAACAC  +  1986
 H  Q  Q  L  S  H  Q  Q  Q  Q  A  L  H  Q  Q  L  S  H  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  H    662
CCGCACTCCCAACTGAACGGTCCGCATCCGCACTCACATCCGCACTCTCATCCGCATTCGCATCCGCACGCGGGCCAGCACACGCATTCC  +  2076
 P  H  S  Q  L  N  G  P  H  P  H  S  H  P  H  S  H  P  H  S  H  P  H  A  G  Q  H  T  H  S    692
ACAATCGCTGCAGCGGCGGCAGCGGCAGCAGCATCTGTCGTCAGCAGCAGCAGCAGCGCCGTTGCAGCGGCAGCCATGCTGAGCGCCAGT  +  2166
 T  I  A  A  A  A  A  A  A  A  A  S  V  V  S  S  S  S  S  A  V  A  A  A  A  M  L  S  A  S    722
```

```
GCAGCGGCAGCAGCAACGGCGGCAGCTGCGGCGGGCGGATCCCAGAGTGTTATCCAACCGGCGACGTCCAGCGTCAGCTACGATCTCTCC  +  2256
 A  A  A  A  A  T  A  A  A  A  A  G  G  S  Q  S  V  I  Q  P  A  T  S  S  V  S  Y  D  L  S      752
TACATGCTGGAGCTGGGCGGATTCCAGCAGCGGAAGGCGAAGAAGCCGCGCAAACCGAAGCTGGAGATGGGCGTAAAGCGGCGCAGCCGG  +  2346
 Y  M  L  E  L  G  G  F  Q  Q  R  K  A  K  K  P  R  K  P  K  L  E  M  G  V  K  R  R  S  R      782
GAGGGATCCACCACCTATCTGTGGGAGTTCCTCCTCAAACTGCTCCAGGATCGCGAATACTGTCCGCGTTTCATCAAGTGGACGAACCGG  +  2436
 E  G  S  T  T  Y  L  W  E  F  L  L  K  L  L  Q  D  R  E  Y  C  P  R  F  I  K  W  T  N  R      812
GAGAAGGGCGTCTTCAAGCTGGTCGACTCGAAGGCTGTGTCCCGCCTGTGGGGCATGCACAAGAACAAGCCGGACATGAACTACGAGACG  +  2526
 E  K  G  V  F  K  L  V  D  S  K  A  V  S  R  L  W  G  M  H  K  N  K  P  D  M  N  Y  E  T      842
ATGGGCAGGGCGCTGAGGTACTACTACCAGCGCGGCATACTGGCCAAGGTGGATGGCCAGCGGCTGGTCTACCATTTCGTGGATGTGCCC  +  2616
 M  G  R  A  L  R  Y  Y  Y  Q  R  G  I  L  A  K  V  D  G  Q  R  L  V  Y  H  F  V  D  V  P      872
AAGGACATCATCGAGATTGACTGCAACGGTGTGTGAAGTGATGAGGACGAGGAGGACGACGAGGCCTGTAAATATAGTTCGATCCCAGAT  +  2706
 K  D  I  I  E  I  D  C  N  G  V  *                                                            883
CGCAGACGGATCGAGCTGGAATGGCAATAGATTTTGGGTGGGGTATAAATCAACCGGCTAGCGGCAATAGTGGCTGAAGGCCACGCGAGA  +  2796
TAGGGATTTCTTTGTACATGCGTTGAGTTTCTTGGTTAGCCAGCGATAAGCGTAGGTTACACACATAGCTGATAGTCCCAATACGCGTAC  +  2886
ATATACTTATATATATATATACATATATATATAAATATATATTCACACCCTGTGTCCTGTACCCTATACATACACCCAGCTAAATCGACG  +  2976
GTAAATAATGATCTATTCATGGGCGTTAGTGTTGTACTTGTAACTGTAGATTTAACTGTAAACTGTAACTGTAATGAATATTGTTATTAT  +  3066
TGGTTATTATTATTATTATTATTACTATTACTATTATTAATATTATAATTATTATTGCTACTGCTATGGCGGAAGGAAGAGGACGAAGGA  +  3156
CGACCGAATTGATTCATATTGATATTCGATAGATGTGTGTAATATGTAAATGAGTACGAACCCCTATAGCTGCAAGAATTTCTCTATCTC  +  3246
TATTGATCTAATGCTAAATCTAAATGTATATGAGTATGTATATGCCTAGCCTAATGACCTTTCAACTGTCCCCTAGCCCAACCCGCTGCC  +  3336
CAAGGACAAGCTCAGTGCAGATGCAGATATGGGTCTATGTACGAACATATGTATATGGAGATCTGAACGATCTGCACGGGTGTCTGTGCT  +  3426
TGGGCTAATTGTTTTTGTTGCTTATTTTATAAACGAAGAGAGATGGAAATTTATTATTTCATAGTTTTAGCTATTTGTACGACTACGTTT  +  3516
AGGATATTTATTTAAGTGTTTCTCCAGCTAATCCCCCGAAACTTTCCCCAAAACAACAAAAACACACACATCCTAACTGGAGCTAAGAAC  +  3606
CCAGCAAACTGGTGGTCTAAACTTTCTTGTAATTATTATTTGATAAGTTTTGTATATTTACAACATCAACAGCAGAAAAGCTCACAGGAT  +  3696
AAACCCGAAGAAGGCACTCCTTGAGCCCTCGTTTAAGCAGTCTTAGAACACCAAACCAAACTCAATTCTCTCGATCTGAAGGTTATTTAT  +  3786
AAAGCATAAAGTTAAACTACCGTATTCTTGTTAACGTCACACTGCCACAAATCGAAACAAGCCTTACATAATGGAACAGATCTGGACATT  +  3876
TAATCGAACCAAACCGAGATACTGACCCATAGTTTCAACAATACTGCAACCGCAATCCTCAATGCAATATACCGACAACTTCTACAACAA  +  3966
TTTCCAAACACTGCCCAAAACGTGCAATCTCTTGAAAGGATACCTTGAATAGGAAAGAGAATAAAACCAACTCCTAGTTCTTAAATCTAA  +  4056
TTCTTTAATAGTTGCGTTCTCTAAGTACCTTTATCTATTATTATTATTATTATGTGTACGTGAAGATGAAATCTACAATATATCATTTTG  +  4146
TTTATGATGATTGGTGATTGATGATGCAAATAAACAAAGTGAGTCTTGAGTTCAACTCTGTTCATTGTAAAAAATTTTTTATGATTATTT  +  4236
TTTATTTTTTAAATAAAATTATTATTTCAAAAAAAAAAA                                                    +  4275
```

Fig. 2: DNA sequence and predicted amino acid sequence of *74E* protein. Nucleotides are numbered in the 5′−3′ direction, beginning with the first ATG triplet encoding the putative methionine initiation amino acid. The nucleotides on the 5′-side of residue 1 are indicated by negative numbers. The number of the nucleotide residue at the right end of each line is given. Small letters correspond to the 5′-extension of cDNA sequence by genomic DNA, capital letters to cDNA sequence. The sequence of the 74E/1 cDNA clone was from nucleotide +996 to +4275, the sequence of the 74E/2 cDNA clone from nucleotide +168 to +4064. Polyadenylation signals (AATAAA) are underlined. Splice junctions of introns are indicated by arrows. The deduced *74E* amino acid sequence is shown below the nucleotide sequence. Amino acid residues are numbered beginning with the putative initiation methionine. Numbers of the last residues in a line are given on the right-hand side. The termination codon TGA at the end of ORF is marked by an asterisk. Dots denote an upstream in frame stop codon in the 5′-non-translated sequence.

library (MOT11 and JOT2,3,4,5). Some restriction site polymorphisms have been detected between the cloned OregonR and CantonS DNA (Fig. 1A). The boundaries as well as the orientation of the chromosomal walk with respect to the centromere were delimited by *in situ* hybridization experiments. ^{3}H-labeled MOT13 DNA and ^{3}H-labeled JOT6 DNA either separatedly or together were hybridized to squash preparations of salivary gland polytene chromosomes of *Drosophila melanogaster* 3rd instar larvae. As shown in Fig. 1B, ^{3}H-labeled MOT13 DNA hybridized to the proximal end and ^{3}H-labeled JOT6 DNA to the distal end of the puff at 74EF. Thus, the chromosomal walk encompassed the major part of the DNA in the puff at 74EF.

Exploratory Northern blot experiments were carried out by systematically hybridizing cloned 74EF DNA to RNA extracted from salivary glands of 3rd instar larvae incubated

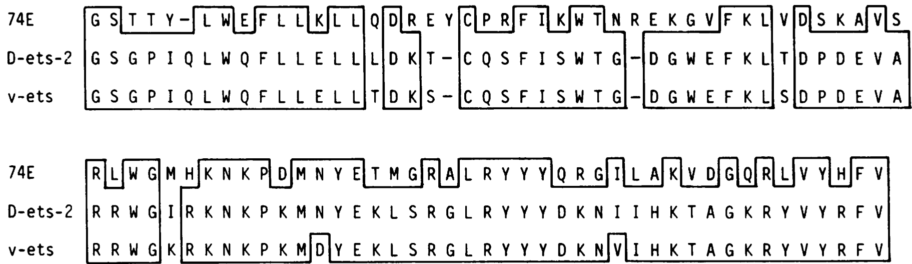

Fig. 3: Comparison of amino acids 784–869 of *74E* protein to the carboxy terminal sequences of *Drosophila D-ets-2* [27] and avian erythroblastosis virus E26 *v-ets* protein [26]. Identical amino acids are boxed.

for 2 to 6h with or without 20-OH-ecdysone in tissue culture medium. A number of transcripts were detected and were analyzed by characterizing corresponding cDNA (data not shown) [22]. However, only one transcript was detected, the synthesis of which was apparently regulated by 20-OH-ecdysone. This transcript was derived from DNA cloned in JOT5. Since this DNA is located in the distal part of the 74EF puff, the locus 74E harbors the corresponding transcription unit. Transcript(s) derived from this transcription unit have been designated 20-OH-ecdysone inducible *74E* transcripts.

Two corresponding cDNA clones, 74E/1 and 74E/2, were isolated from an adult head and an 1–2d old pupa cDNA library, respectively. Both sequences overlap by 3069 bp. The combined cDNA sequence is 4108 bp long (Fig. 2). It contains three polyadenylation signals (AATAAA [23]) followed by nine adenine nucleotides at the 3′-end which are not present in the corresponding genomic DNA sequence. Thus, the 3′-end of the cDNA

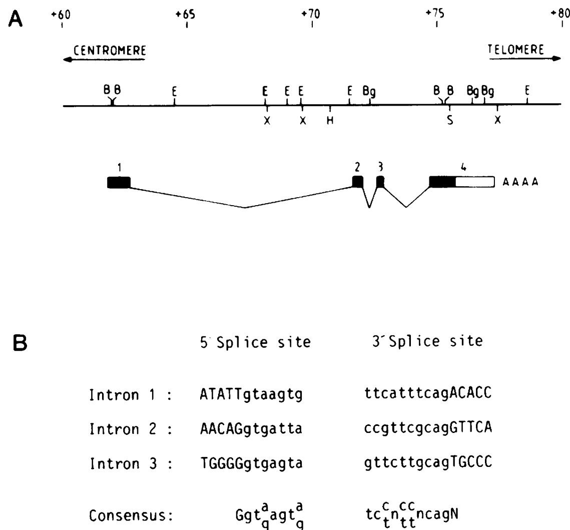

Fig. 4: (A) Structure of *74E* transcription unit. The restriction map is as in Fig. 1. Calibration is in kb. Intervening sequences are indicated by thin lines underneath, exons by numbered boxes. Open box indicates non-translated region, filled boxes correspond to the *74E* cDNA open reading frame. Lengths are to scale. (B) Comparison of the exon/intron (capital letters/small letters) boundaries to the *Drosophila* splice consensus [29].

sequence shown in Fig. 2 should be the 3'-terminus of the *74E* transcript. The longest open reading frame derived from the cDNA sequence started at the second nucleotide. We did not succeed in isolating further cDNA molecules extending the 5'-sequence of the cDNA. Therefore, genomic DNA was isolated and sequenced which extended the cDNA sequence (Fig. 2). 167 nucleotides upstream of the first cDNA nucleotide an in frame ATG is present in the genomic DNA which matches the *Drosophila* start consensus sequence (C/A)AA(A/C)ATG [24]. 36 nucleotides in front of this ATG is located an in frame stop codon. Accordingly, the *74E* open reading frame would be 2649 nucleotides long as shown in Fig. 2.

The proposed amino acid sequence of the *74E* protein is composed of 883 amino acids with a calculated molecular weight of 94.8 kDa. The protein can be divided into three domains: an acidic amino terminus (amino acids 1−280), a middle domain rich in Gln, Ala, and Ser as a consequence of *opa* elements [25], and a basic carboxy terminus (amino acids 760−883). Only 17 of the 69 basic amino acids, but 50 of the 80 acidic amino acids occuring in *74E* protein are located in the amino terminal domain. Viceversa, 28 of the 69 basic amino acids, but only 14 acidic amino acids are located in the carboxy terminal domain. The sequence encompassing amino acids 784−869 of the carboxy terminal domain is similar to the carboxy terminal sequences of *ets*-oncoproteins (Fig. 3). 43 or 45 amino acids of this sequence (50−52%) are found at identical positions in the sequences of erythroblastosis virus E26 *v-ets* protein [26] and of *Drosophila D-ets-2* protein [27], respectively. If one allows for conservative amino acid substitutions [28], the similarity between the sequence of *74E* protein and *ets*-oncoproteins increases to 67%.

Comparison of the cDNA sequence with the corresponding genomic DNA sequence revealed 11 nucleotide differences outside the cDNA's open reading frame. Two base pair substitutions were detected in the coding region, but only one of them (G versus T at nucleotide 2601) would alter the derived *74E* amino acids sequence by replacing His 867 by Gln. The sequence differences are most likely explained by the use of different *Drosophila* stocks in the construction of genomic and cDNA libraries. The alignment of cDNA with the genomic DNA sequence showed 3 points of non-colinearity, indicative of 3 intervening sequences in the primary *74E* transcript (Fig. 4A). The splice junctions for these intervening sequences are located at positions +1039, +1408, and +1663. The exon/intron boundaries conform to the consensus sequences of *Drosophila* splice sites (Fig. 4B) [29].

Expression of 74E transcripts

The developmental profile of *74E* mRNA expression was studied by Northern blot experiments. A BamHI/BglII fragment of cDNA 74E/1 containing exon 4 sequences was used as hybridization probe. It hybridized to two transcripts, approximately 5.0 and 6.2 kb long (Fig. 5A). Although each lane in the Northern blot experiment in Fig. 5A did not contain the same quantity of poly(A)$^+$-RNA as indicated by the control hybridization with β1-tubulin cDNA [30], it is obvious that both *74E* transcripts are predominantly present in late 3rd instar larvae and in early pupae.

Small amounts of the 5.0 kb transcript were also present in 12- 18h embryos, early 2nd instar larvae, and late pupae whereas small amounts of the 5.0 kb and the 6.2 kb transcript were detected in 18−24h embryos and late 2nd instar larvae. Corresponding faint bands were visible in the original autoradiographic exposure in the lanes containing poly(A)$^+$-RNA of these developmental stages. The results shown in Fig. 5A indicate that the 5.0 kb transcript is synthesized ahead of the 6.2 kb transcript during 3rd larval instar

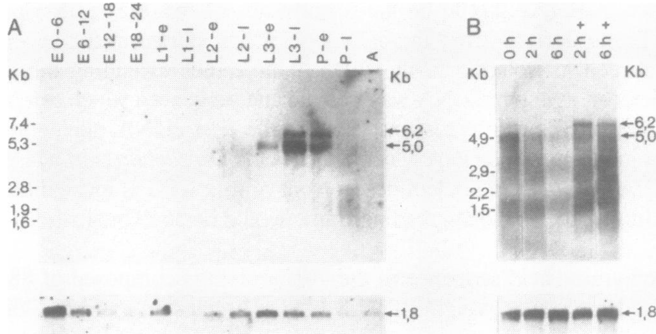

Fig. 5: Expression of *74E* transcripts. A ^{32}P-labeled cDNA probe from exon 4 was used for Northern hybridization. Size markers in kb are shown on the left side. To check the amounts and the integrity of RNA Northern blots were reprobed with probes to β1-tubulin [30]. The results are shown under each corresponding lane. The size of β1-tubulin mRNA is 1.8 kb. (A) poly(A)$^+$-RNA from different developmental stages. E0−6, E6−12, E12−18, E18−24: 0−6h, 6−12h, 12−18h, 18−24h old embryos. L1, L2, L3: 1st, 2nd, 3rd larval instar. P: pupa. A: adult. e = early and l = late, to denote the first and second half of the developmental stage, respectively. In case of L3 e denotes larvae crawling in the nutrient agar and l denotes larvae which have left the nutrient agar. (B) RNA from puff stage 1 salivary glands of 3rd instar larvae incubated *in vitro* for the indicated time span without or in the presence of 5×10^{-6}M 20-OH-ecdysone (+).

development. This suggests that the expression of both *74E* transcripts is regulated differently.

The possibility that the appearance of one or both *74E* transcripts in 3rd instar larvae might be correlated with a corresponding peak of ecdysteroid titre in larval development [31] was investigated by studying *in vitro* the effect of 20-OH-ecdysone on *74E* mRNA expression. For this purpose, salivary glands corresponding to puff stage 1 were dissected from 3rd instar larvae and were explanted in tissue culture medium. Incubations were carried out for 2 or 6h at 25°C in the presence of 5×10^{-6}M 20-OH-ecdysone. Control incubations were carried out without added 20-OH-ecdysone. Subsequently, *74E* mRNA expression was analyzed by Northern blot experiments with total RNA isolated from the salivary glands (Fig. 5B). Apparently, the non-specific background on the resultant autoradiogram was caused by rRNA and/or RNA present in the total RNA preparations, which was not isolated with the poly(A)$^+$-RNA used in the developmental Northern blot experiments in Fig. 5A. Evidently, the synthesis of the 5.0 kb transcript is not influenced by 20-OH-ecdysone. The amount of the 5.0 kb transcript decreased in salivary glands which were incubated with or without 20-OH-ecdysone. On the other hand, the synthesis of the 6.2 kb transcript is regulated by 20-OH-ecdysone. The amount of the 6.2 kb transcript increases in salivary glands upon incubation with 20-OH-ecdysone. Incubations of 2h apparently induced larger quantities of the 6.2 kb transcript than incubations of 6h. This suggests that the expression of the 6.2 kb transcript is transiently induced by 20-OH-ecdysone in explanted salivary glands. This transient increase follows closely the activity of the 74EF puff which is rapidly induced by 20-OH-ecdysone, reaches a maximal size at about 2−4h and then regresses.

DISCUSSION

We have analyzed the structure of *74E* cDNA. It was used to study the expression of two *74E* transcripts in the 20-OH-ecdysone regulated puff 74EF in salivary glands of *Drosophila melanogaster* 3rd instar larvae. As probes corresponding to the first exon of *74E* cDNA

hybridized only to 5.0 kb mRNA in Northern blot experiments (data not shown) the isolated *74E* cDNA is derived from the 5.0 kb transcript. The structure of an alternative *74E* cDNA has been described which corresponds to the 6.2 kb transcript [32]. It shares with our cDNA sequence exons 2–4 but disposes of alternative exons at its 5'-end. Thus, the two *74E* transcripts have different exon compositions which could be generated by transcription from two different promoters and/or by alternative splicing mechanisms. The different exon compositions give rise to two alternative reading frames of which proteins are deduced with different amino terminal domains. The deduced proteins have in common most of the middle and the carboxy terminal domain.

The 5.0 kb transcript is already present in salivary glands of puff stage 1, i.e. before the puff 74EF is active. This transcript is not 20-OH-ecdysone inducible in explanted salivary glands. On the contrary, the 5.0 kb transcript disappears upon incubations of salivary glands *in vitro*. However, the synthesis (amount) of the 6.2 kb transcript transiently increases in explanted salivary glands in the presence of 20-OH-ecdysone (Fig. 5B). This transient increase correlates with the transient activity of puff 74EF induced by 20-OH-ecdysone. Therefore, we propose that the primary result of 20-OH-ecdysone regulated 74EF puff activity would be a decrease in the synthesis of the 5.0 kb transcript concomitant with an increase in the synthesis of the 6.2 kb transcript. A corollary of these observations is that 20-OH-ecdysone does not activate together with the 74EF puff the expression of a hitherto silent gene, but induces the synthesis of an alternative transcript and hence of an alternative gene product.

The common carboxy terminal domain of the *74E* gene products is similar to the carboxy terminal sequences of *ets*-oncoproteins (Fig. 3). *ets*-oncoproteins are nuclear proteins [33,34] and may bind to DNA [35]. Together with our previous proposition that 20-OH-ecdysone inducible gene(s) of puff 74EF encode transcription factor(s) [10], these observations suggest that *74E* gene product(s) function as transcription factor(s). Then, the 20-OH-ecdysone induced puff activity at 74EF would lead to a transcription factor switch. Such a switch would have important implications for models to understand the time-dependent sequence of puffing activities controlled by gene products of early puffs. They autoregulate their own expression as well as stimulate the activity of late puffs [6,10]. This functional duality could be accomplished easily by the synthesis of alternative transcription factors. Firstly, they would compete with existing ones for the same DNA binding sites and secondly, they would interact with other, i.e. different transcription factors. Both mechanisms might be the molecular basis for the time-dependent sequence of gene activities seen during the development of 3rd instar larvae into pupae.

ACKNOWLEDGEMENTS

We thank M. Rentrop for providing MOT phages. This work was supported by the Deutsche Forschungsgemeinschaft.

*To whom correspondence should be addressed

+Present address: Institut für Humangenetik, Universitätsklinikum Essen, Hufelandstrasse 55, D-4300 Essen, FRG

REFERENCES

1. Karlson, P. (1966) Naturwissenschaften **53**, 445–453.
2. Becker, H.J. (1959) Chromosoma **10**, 654–678.
3. Ashburner, M. (1972) Chromosoma **38**, 255–281.

4. Ashburner, M. (1973) Dev. Biol. **35**, 47–61.
5. Ashburner, M. (1974) Dev. Biol. **39**, 141–157.
6. Walker, V.K. and Ashburner, M. (1981) Cell **26**, 269–277.
7. Gronemeyer, H. and Pongs, O. (1980) Proc. Natl. Acad. Sci. USA **77**, 2108–2112.
8. Dworniczak, B., Seidel, R. and Pongs, O. (1983) EMBO J. **2**, 1323–1330.
9. Schaltmann, K. and Pongs, O. (1982) Proc. Natl. Acad. Sci. USA **79**, 6–10.
10. Pongs, O. (1988) Eur. J. Biochem. **175**, 199–204.
11. Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) Cell **15**, 687–701.
12. Kaiser, K. and Murray, N.E. (1985) In Glover, D.M. (ed), DNA Cloning–A Practical Approach, IRL Press, Oxford, Vol 1, pp. 1–47.
13. Benton, W.D. and Davis, R.W. (1977) Science **196**, 180–182.
14. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
15. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. **113**, 237–251.
16. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13.
17. Gerard, G.F. and Miller, K. (1986) Focus **8**:3, 5–6.
18. Baumann, A., Krah-Jentgens, I., Müller, R., Müller-Holtkamp, F., Seidel, R., Kecskemethy, N., Casal, J., Ferrus, A. and Pongs, O. (1987) EMBO J. **6**, 3419–3429.
19. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463–5467.
20. Lin, H.-C., Lei, S.-P. and Wilcox, G. (1985) Anal. Biochem. **147**, 114–119.
21. Möritz, T., Edström, J.E. and Pongs, O. (1984) EMBO J. **3**, 289–295.
22. Rettberg, P. (1987) Ph.D. Thesis, Ruhr-Universität Bochum, FRG.
23. Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell **41**, 349–359.
24. Cavener, D.R. (1987) Nucl. Acids. Res. **15**, 1353–1361.
25. Wharton, K.A., Yedvobnick, B., Finnerty, V.G. and Artavanis- Tsakonas, S. (1985) Cell **40**, 55–62.
26. Nunn, M.F., Seeburg, P.H., Moscovici, C. and Duesberg, P.H. (1983) Nature **306**, 391–395.
27. Pribyl, L.J., Watson, D.K., McWilliams, M.J., Ascione, R. and Papas, T.S. (1988) Dev. Biol. **127**, 45–53.
28. Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Silver Springs, MD, Vol. 5, Suppl. 3, pp. 345–352.
29. Snyder, M., Hunkapiller, M., Yuen, D., Silvert, D., Fristrom, J. and Davidson, N. (1982) Cell **29**, 1027–1040.
30. Bialojan, S., Falkenburg, D. and Renkawitz-Pohl, R. (1984) EMBO J. **3**, 2543–2548.
31. Richards, G. (1981) Molec. cell. Endocr. **21**, 181–197.
32. Burtis, K.C. (1986) Ph. D. Thesis, Stanford University, California (USA).
33. Boulukos, K.E., Pognonec, P., Begue, A., Galibert, F., Gesquiere, J.C., Stehelin, D. and Ghysdael, J. (1988) EMBO J. **7**, 697–705.
34. Bhat, N.K., Fisher, R.J. Fujiwara, S., Ascione, R. and Papas, T.S. (1987) Proc. Natl. Acad. Sci. USA **84**, 3161–3165.
35. Golay, J., Introna, M. and Graf, T. (1988) Cell **55**, 1147–1158.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.