Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 6167–6171, July 1992
Developmental Biology

DHR3: A *Drosophila* steroid receptor homolog

(ecdysone-inducible genes/zinc fingers/hormones)

MICHAEL R. KOELLE, WILLIAM A. SEGRAVES*, AND DAVID S. HOGNESS†

Departments of Biochemistry and Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305-5427

Contributed by David S. Hogness, April 6, 1992

ABSTRACT **In *Drosophila* the steroid hormone ecdysone triggers a genetic regulatory hierarchy in which ecdysone combines with a receptor protein to form a complex that induces the transcription of a small class of "early" genes, which encode transcription factors that regulate other genes. We previously reported that one of the early genes, *E75*, encodes members of the steroid receptor superfamily. Using an *E75* hybridization probe, we have identified two additional *Drosophila* genes that encode members of this superfamily. One of these is the ecdysone receptor gene, *EcR*, as previously reported. In this work, we examine the sequence, genomic organization, and developmental expression of the other gene, *DHR3*, which, like *E75*, encodes one of a growing number of "orphan" receptors for which ligands have not yet been identified. The structure of the DHR3 protein is strikingly similar to that of the MHR3 protein (e.g., 97% amino acid identity for the DNA binding domains), another orphan receptor encoded by an ecdysone-inducible early gene of another insect, *Manduca sexta*. The temporal developmental profile for *DHR3* expression closely parallels that for the ecdysone titer and for the ecdysone-inducible *E75* and *E74 Drosophila* early genes. The structural similarity to a *Manduca* early gene and the expression similarities to *Drosophila* early genes suggest that the *DHR3* gene may also belong to the early gene class.**

The most dramatic form of the ecdysone response in *Drosophila melanogaster* occurs at the end of larval life when a pulse of this steroid hormone triggers metamorphosis to the adult fly. The first operational definition of the genes involved in this response came from analyses of the transcription puffs induced by that pulse in salivary gland chromosomes. These analyses led to a model for an ecdysone-activated genetic regulatory hierarchy consisting of "early" and "late" genes responsible for the respective temporal classes of puffs and two sorts of transcription factors, each of which regulates both classes of genes by both positive and negative controls (1). The first factor postulated by the model is an ecdysone receptor, which, when complexed with the hormone, induces the early genes and represses the late genes. By contrast, factors of the second kind, which are encoded by the early genes, induce and repress the late and early genes, respectively. These two +/− regulatory duets account for the rapid induction and subsequent regression of the early puffs and the delayed induction of the late puffs.

The past few years have seen a striking increase in our knowledge of the genes controlling ecdysone response, particularly with respect to the early genes. Three of these genes have been cloned and characterized: *E74* (2–5), *E75* (6, 7), and *BR-C* (8, 9). The molecular and mutational characterization of these genes is not only consistent with the genetic regulatory hierarchy postulated above but indicates that such a hierarchy may be general to all ecdysone target tissues.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

The work reported here derives from the finding that the proteins encoded by the *E75* early gene exhibit sequence similarity to the DNA and ligand binding domains of the receptors for the steroid hormones, thyroid hormone, and retinoic acid and thus belong to the steroid receptor superfamily (refs. 6 and 7; see refs. 10–12 for reviews of this superfamily). The E75 proteins are not ecdysone receptors and the ligand to which they bind has not been identified, placing them in the growing class of "orphan" receptors (12).

In a quest for *Drosophila* genes encoding ecdysone receptors and other members of the steroid receptor superfamily, we have screened a genomic DNA library by hybridization with a probe encoding the E75 DNA binding domain. This paper describes the characterization of *DHR3*, one of two receptor genes isolated in this screen; the other, *EcR*, encodes ecdysone receptors and has been described elsewhere (13).‡

MATERIALS AND METHODS

Low-Stringency Screen. Standard techniques and solutions used for manipulating DNA were as described (13). A 530-base-pair (bp) *Eco*RV/*Xho* I cDNA fragment containing the coding sequences for the E75A DNA binding domain (7) was nick-translated and used to probe Southern blots of *Drosophila* genomic DNA digested with *Eco*RI, *Xho* I, *Eco*RV, or *Bam*HI. Low-stringency hybridizations were performed by (*i*) prehybridization of filters in 5× SSPE (1× SSPE = 0.18 M NaCl/10 mM phosphate, pH 7.4/1 mM EDTA)/10× Denhardt's solution/salmon sperm DNA (100 μg/ml)/0.1% SDS at 65°C for 2 hr; (*ii*) hybridization with denatured probe in 5× SSPE/40% formamide/10× Denhardt's solution/salmon sperm DNA (1 mg/ml)/0.5% SDS for 16 hr at 37°C; and (*iii*) extensive washing in 1× SSPE/0.5% SDS at 50°C. High-stringency hybridizations were performed under similar conditions with hybridization at 42°C in the presence of 50% formamide, and washing in 0.3× SSPE/0.25% SDS at 60°C.

Five genome equivalents of a Canton S *Drosophila* genomic library (14) were screened with the same probe under the same low-stringency conditions. The control probe used on duplicate filters consisted of a mixture of clones Dm4234, Dm4207, and Dm4260 (6). This probe derives from intron sequences adjacent to the exons making up the *E75A* cDNA probe and detects only *E75* clones.

Isolation and Analysis of cDNA and Additional Genomic Clones. Subclones of the *DHR3* genomic clones were used to screen a cDNA library prepared from third-instar tissues treated with ecdysone and cycloheximide (a gift of C. S.

Abbreviation: ORF, open reading frame.

*Present address: Department of Biology, Yale University, New Haven, CT 06511.
†To whom reprint requests should be addressed at: Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305-5427.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M90806).

Thummel, University of Utah). Of the 270,000 primary plaques screened, 20 positive for *DHR3* were detected. cDNA clone DHR3-9, which extends farther both 5′ and 3′ than the other *DHR3* cDNA clones, was sequenced as described (13). The cDNA was sequenced on both strands for the 5′-most 2338 bp, which contains the entire open reading frame (ORF), and the remainder of the long 3′ untranslated region was sequenced on one strand. Additional genomic clones were obtained by screening the Canton S genomic library with *DHR3* cDNA probes. Genomic clones were restriction mapped and probed with labeled *DHR3* cDNAs to localize exon-containing fragments on the genomic restriction map. Genomic exons were sequenced entirely, except in the case of the long 3′ exon, for which genomic and cDNA clones were restriction digested and electrophoresed in parallel to confirm their colinearity.

***In Situ* Hybridization to Polytene Chromosomes.** *In situ* hybridization was carried out with biotinylated DNA probes as described (15).

Northern Blot Analysis. The developmental Northern blot shown is the same filter as used in ref. 4 serially rehybridized with randomly primed *DHR3* and *RP49* probes. The *DHR3* probe was the DHR3-9 cDNA. For *RP49* (16), an *Eco*RI/*Xho* I genomic fragment containing the entire gene was used.

RESULTS

Identification of Genes Encoding *Drosophila* Steroid Receptor Homologs. We used the gene for the previously isolated *Drosophila* steroid receptor homolog E75 (7) as a hybridization probe to search the *Drosophila* genome for additional members of this receptor family. Because the 67 ± 1 amino acid DNA binding domain is the most highly conserved region of the steroid receptors (10), a 530-bp cDNA probe containing the coding sequences for this region of the E75A protein was chosen for the screen. (E75A is one of the isoforms encoded by *E75* via alternative promoters; see ref. 7.) Initially, we probed genomic Southern blots with this DNA to find hybridization conditions under which it would detect a small number of non-*E75* bands. Fig. 1 (lanes 1 and 2) shows representative blots. Under the optimized conditions, we saw between one and five bands per restriction digest that could be detected only at the reduced stringency.

To isolate the sequences responsible for these low-stringency bands, we used the *E75* probe to screen a *Drosophila* genomic library under the same low-stringency conditions, counterscreening duplicate filters with *E75* intron probes to eliminate phage containing inserts from the *E75* gene. Five genome equivalents were screened and 39 non-*E75* containing phage were isolated. The 25 most strongly hybridizing clones were divided into six classes on the basis of restriction mapping and cross-hybridization, with each class containing a set of between one and six independent overlapping genomic inserts.

For each class, a restriction fragment containing the region of hybridization to the *E75* probe was localized by Southern blotting. Hybridization of probes derived from these fragments to genomic Southern blots showed that the fragments can account for all of the low-stringency bands detectable by the *E75* probe (Fig. 1, lane 3). This suggests that these cloned restriction fragments contain the six non-*E75* sequences from the *Drosophila* genome that hybridize most strongly with the E75 DNA binding domain probe.

The six restriction fragments were sequenced to determine whether they derive from candidate receptor genes. Although we observed DNA sequence similarities with the *E75* probe for each of these fragments, only two of the fragments showed predicted amino acid sequences with strong similarity to the DNA binding domains of E75 and other steroid receptor homologs (see ref. 6 for further discussion). We have shown that one of these two fragments derives from the ecdysone receptor gene *EcR* (ref. 13; referred to as DHR23 in ref. 6). We have named the other receptor gene *DHR3* (*Drosophila* hormone receptor 3). *In situ* hybridization of *EcR* and *DHR3* probes to polytene chromosomes mapped these genes to the 42A and 46F positions, respectively, on the right arm of chromosome 2 (data not shown).

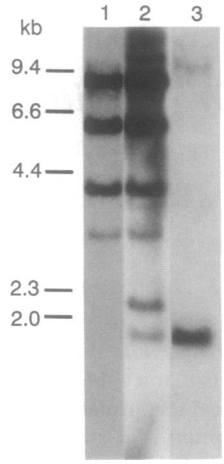

FIG. 1. Low-stringency Southern blots detect *E75*-related sequences. Lane 1, a high-stringency Southern blot with the *E75* DNA binding domain probe hybridized to *Xho* I-digested *Drosophila* DNA. Lane 2, same as lane 1 but probed at low stringency. Additional bands, representing *E75*-related sequences, are detected at 1.8, 2.1, and ≈12 kilobases (kb). Similar blots were prepared with *Drosophila* DNA digested with several other enzymes in order to visualize low-stringency bands that might comigrate with high-stringency bands in a single digest. Lane 3, high-stringency Southern blot of *Xho* I-digested *Drosophila* DNA probed with a *Xho* I fragment from the *EcR* class of clones. This probe detects a 1.8-kb fragment that comigrates with one of the low-stringency bands detected by the *E75* probe, indicating that this band is due to *EcR* sequences. Similar blots with probes from the other clones recovered in the low-stringency screen indicate that each of the low-stringency bands detected by the *E75* probe can be accounted for by one of these clones.

***DHR3* cDNA Sequence.** The *DHR3* clones were used to screen a cDNA library prepared from late third-instar tissues. The longest *DHR3* cDNA clone isolated, DHR3-9, is 4.2 kb long. This cDNA was sequenced (Fig. 2) and found to contain a 487-codon AUG-initiated ORF that predicts a protein homologous to the members of the steroid receptor superfamily. Two additional in-frame AUG codons are found at positions 6 and 9 of the ORF; none of these AUGs shows an excellent match to the general eukaryotic (17) or *Drosophila* (18) translation start consensus sequences. Upstream of this long ORF is a leader sequence containing an additional 57-codon AUG-initiated ORF. The short upstream ORF shows poor *Drosophila* codon usage, while the long ORF has an excellent match to *Drosophila* codon usage (16). The long ORF is followed by a 2.5-kb A+T-rich untranslated 3′ region. The DHR3-9 cDNA does not contain a poly(A) tract at its 3′ end.

Comparison of the predicted DHR3 protein sequence to the sequence data base and to individual members of the steroid receptor superfamily shows that the DHR3 protein contains the two conserved domains (underlined in Fig. 2) characteristic of steroid receptor superfamily members (10, 11). The more N-terminal and the more C-terminal of these conserved domains are referred to as the C and E regions, respectively, according to the nomenclature of Krust *et al.* (19). The C region is a 67 ± 1 amino acid sequence that has been shown to function as a Zn finger DNA binding domain in vertebrate receptors (20, 21). The E region is an ≈225-amino acid

```
   1  GAATTCATTCAACTGCAAAGAGCAGCCAAATTGCGCATACGCCGCGTATGGCCGTCGGTGTGAGTGCCCGTGTTCATCAGCGGTTGCATCAACTGATACCAAGTGTACATAACTACAGCTACAA
 125  TTGCAACTATTTCACCAATCAACGGCAGCGGCAACAACATCAGCAACAGCACCGGCAAACGTTTGAAACGTCACCAAAGCTTCGCATTTCCCACTAATAATT   ATG TAT ACG CAA CGT
                                                                                                                  MET Tyr Thr Gln Arg

 242  ATG TTT GAC ATG TGG AGC AGC GTC ACT TCG AAA CTG GAA GCA CAC GCA AAC AAT CTC GGT CAA AGC AAC GTC CAA TCG CCG GCG GGA CAA AAC
   6  MET Phe Asp MET Trp Ser Ser Val Thr Ser Lys Leu Glu Ala His Ala Asn Asn Leu Gly Gln Ser Asn Val Gln Ser Pro Ala Gly Gln Asn
                    {gtaaag··v··tcacag}
 335  AAC TCC AGC GGT TCC ATT AAA GCT CAA ATT GAG ATA ATT CCA TGC AAA GTC TGC GGC GAC AAG TCA TCC GGC GTG CAT TAC GGA GTG ATC ACC
  37  Asn Ser Ser Gly Ser Ile Lys Ala Gln Ile Glu Ile Ile Pro Cys Lys Val Cys Gly Asp Lys Ser Ser Gly Val His Tyr Gly Val Ile Thr

 428  TGC GAG GGC TGC AAG GGA TTC TTT CGA AGA TCG CAA AGC TCC GTG GTC AAC TAC CAG TGT CCG CGC AAC AAG CAA TGT GTG GTG GAC CGT GTT
  68  Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Gln Ser Ser Val Val Asn Tyr Gln Cys Pro Arg Asn Lys Gln Cys Val Val Asp Arg Val
                                                                  {gtctgt··v··ttgcag}
 521  AAT CGC AAC CGA TGT CAA TAT TGT AGA CTG CAA AAG TGC CTA AAA CTG GGA ATG AGC CGT GAT GCT GTA AAG TTC GGC AGG ATG TCC AAG AAG
  99  Asn Arg Asn Arg Cys Gln Tyr Cys Arg Leu Gln Lys Cys Leu Lys Leu Gly MET Ser Arg Asp Ala Val Lys Phe Gly Arg MET Ser Lys Lys


 614  CAG CGC GAG AAG GTC GAG GAC GAG GTA CGC TTC CAT CGG GCC CAG ATG CGG GCA CAA AGC GAC GCG GCA CCG GAT AGC TCC GTA TAC GAC ACA
 130  Gln Arg Glu Lys Val Glu Asp Glu Val Arg Phe His Arg Ala Gln MET Arg Ala Gln Ser Asp Ala Ala Pro Asp Ser Ser Val Tyr Asp Thr
                                                {gtgcag··v··actcag}
 707  CAG ACG CCC TCG AGC AGC GAC CAG CTG CAT CAC AAC AAT TAC AAC AGC TAC AGC GGC GGC TAC TCC AAC AAC GAG GTG GGC TAC GGC AGT CCC
 161  Gln Thr Pro Ser Ser Ser Asp Gln Leu His His Asn Asn Tyr Asn Ser Tyr Ser Gly Gly Tyr Ser Asn Asn Glu Val Gly Tyr Gly Ser Pro

 800  TAC GGA TAC TCG GCC TCC GTG ACG CCA CAG CAG ACC ATG CAG TAC GAC ATC TCG GCG GAC TAC GTG GAC AGC ACC ACC TAC GAG CCG CGC AGT
 192  Tyr Gly Tyr Ser Ala Ser Val Thr Pro Gln Gln Thr MET Gln Tyr Asp Ile Ser Ala Asp Tyr Val Asp Ser Thr Thr Tyr Glu Pro Arg Ser
                                  {gtaaag··v··ctccag}      (C)
 893  ACA ATA ATC GAT CCC GAA TTT ATT AGT CAC GCG GAT GGC GAT ATA AAC GAT GTG CTG ATC AAG ACG CTG GCG GAG GCG CAT GCC AAC ACA AAT
 223  Thr Ile Ile Asp Pro Glu Phe Ile Ser His Ala Asp Gly Asp Ile Asn Asp Val Leu Ile Lys Thr Leu Ala Glu Ala His Ala Asn Thr Asn
                                        {gtgagt··v··gtgcag}   (G)
 986  ACC AAA CTG GAA GCT GTG CAC GAC ATG TTC CGA AAG CAG CCG GAT GTG TCA CGC ATT CTC TAC TAC AAG AAT CTG GGC CAA GAG GAA CTC TGG
 254  Thr Lys Leu Glu Ala Val His Asp MET Phe Arg Lys Gln Pro Asp Val Ser Arg Ile Leu Tyr Tyr Lys Asn Leu Gly Gln Glu Glu Leu Trp
          (C)
1079  CTG GAC TGC GCT GAG AAG CTT ACA CAA ATG ATA CAG AAC ATA ATC GAA TTT GCT AAG CTC ATA CCG GGA TTC ATG CGC CTG AGT CAG GAC GAT
 285  Leu Asp Cys Ala Glu Lys Leu Thr Gln MET Ile Gln Asn Ile Ile Glu Phe Ala Lys Leu Ile Pro Gly Phe MET Arg Leu Ser Gln Asp Asp
 {gtgag··v··cctag}
1172  CAG ATA TTA CTG CTG AAG ACG GGC TCC TTT GAG CTG GCG ATT GTT CGC ATG TCC AGA CTG CTT GAT CTC TCA CAG AAC GCG GTT CTC TAC GGC
 347  Gln Ile Leu Leu Leu Lys Thr Gly Ser Phe Glu Leu Ala Ile Val Arg MET Ser Arg Leu Leu Asp Leu Ser Gln Asn Ala Val Leu Tyr Gly

                                                  (G)
1265  GAC GTG ATG CTG CCC CAG GAG GCG TTC TAC ACA TCC GAC TCG GAA GAG ATG CGT CTG GTG TCG CGC ATC TTC CAA ACG GCC AAG TCG ATA GCC
 378  Asp Val MET Leu Pro Gln Glu Ala Phe Tyr Thr Ser Asp Ser Glu Glu MET Arg Leu Val Ser Arg Ile Phe Gln Thr Ala Lys Ser Ile Ala
                                                                  {gtgcg··v··ccttag}
1358  GAA CTC AAA CTG ACT GAA ACC GAA CTG GCG CTG TAT CAG AGC TTA GTG CTG CTC TGG CCA GAA CGC AAT GGA GTG CGT GGT AAT ACG GAA ATA
 409  Glu Leu Lys Leu Thr Glu Thr Glu Leu Ala Leu Tyr Gln Ser Leu Val Leu Leu Trp Pro Glu Arg Asn Gly Val Arg Gly Asn Thr Glu Ile

1451  CAG AGG CTT TTC AAT CTG AGC ATG AAT GCG ATC CGG CAG GAG CTG GAA ACG AAT CAT GCG CCG CTC AAG GGC GAT GTC ACC GTG CTG GAC ACA
 440  Gln Arg Leu Phe Asn Leu Ser MET Asn Ala Ile Arg Gln Glu Leu Glu Thr Asn His Ala Pro Leu Lys Gly Asp Val Thr Val Leu Asp Thr
                        {gtacgt··v··ttccag}
1544  CTG CTG AAC AAT ATA CCC AAT TTC CGC GAT ATT TCC ATC TTG CAC ATG GAA TCG CTG AGC AAG TTC AAG CTG CAG CAC CCG AAT GTC GTT TTT
 471  Leu Leu Asn Asn Ile Pro Asn Phe Arg Asp Ile Ser Ile Leu His MET Glu Ser Leu Ser Lys Phe Lys Leu Gln His Pro Asn Val Val Phe

1637  CCG GCG CTG TAC AAG GAG CTG TTC TCG ATA GAT TCG CAG CAG GAC CTG ACA TAA CAAGAGCAGCAGCCGTTCCTGGAGACGACCGCGGACGATGTTGCCGAGGAT
      Pro Ala Leu Tyr Lys Glu Leu Phe Ser Ile Asp Ser Gln Gln Asp Leu Thr  -

1742  GCGGCTGCCGCCGGATGTGTCCTGCCGCCGGTGGCGCCCCCTGCCGGGCAGCAACCAGCGCTGCTCGAGGACTGAGGGCCGCAGGATGTGGCAACAATAATTATTTGAGTAAACACTGCACTGC
1866  GCATGCAGCAGATACAAGAACTTTATCATGATTTAAGCTAGCATACAACCAAGGATGTGATCCTCGCCAAGGACTCACTTAAAAAGAACTCTATCTATATACATATATAAATTATATATGACAG
1990  AGCGGATGACGCAAAGGGAAGGGAAAATATTTCAAAAATATTGTTAACTCAGTTAAGACTTTTGCTTCGTAGAGAACCGAAACCGAAACCGATTGCATTTCGAGCAAGGGGCATCAAACTGATT
2114  TTCGAGGTTATACTATACATATATACACACAAACACACACACACACACACACATATATATATATGTAACTTCCAAACTTTCATATCCTGGCCCGAGCAGATCAGATCGTCTAAGTACTTAAA
2238  ACCAAGCGAAATTCTCTACACCGCACAACCCAGGACCCGTAGACCCCAATAATTCAGTTCGGTTAGTGTTAACCCCAGAAAGCCCGATTCCGATCCCGCCT...
```

FIG. 2. Sequence of the DHR3-9 cDNA. The conserved C and E regions are underlined in the deduced protein sequence. In the 5′ leader, the upstream ATG codon is underlined. Positions of introns and the splice donor and acceptor sequences are indicated above the cDNA sequence. The 5′-most 2.3 kb of sequence is shown. The remainder of this 4.2-kb cDNA was sequenced on one strand (not shown). Four silent, third-position differences between the cDNA and genomic clones are indicated above the cDNA sequence.

domain that functions as a hormone binding domain in these receptors (22–25). The presence of E-region homology establishes DHR3 as a bona fide member of the steroid receptor superfamily, as opposed to the *Drosophila* knirps (26), knirps-related (27), and egon (28) proteins, which show homology to the C but not the E region. Quantitative consideration of the sequence similarities between the DHR3 protein and other bona fide members of the steroid receptor superfamily is reserved for the *Discussion*.

***DHR3* Genomic Organization.** We have mapped *DHR3* exons as described. The deduced gene structure is shown in Fig. 3, and the splice junctions are indicated on the cDNA sequence in Fig. 2. These splice junctions all conform to the splice donor and acceptor consensus sequences (29). The

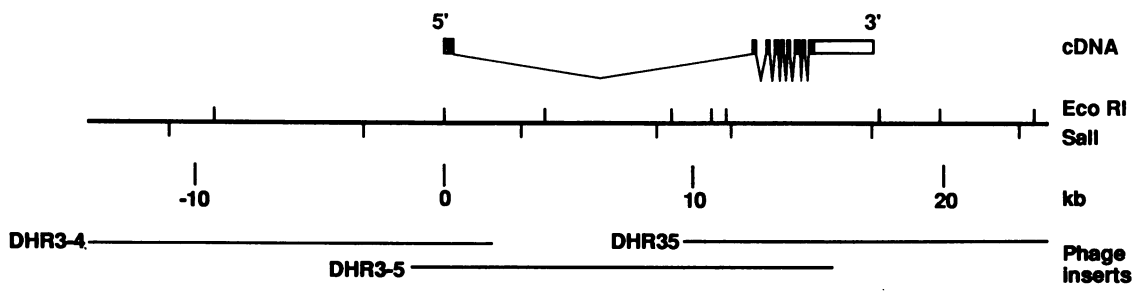

FIG. 3. Genomic organization of *DHR3*. The scale is indicated in kb, with an arbitrary 0 point for the map coordinates. A restriction map is indicated above the scale. Exons mapped by comparison of cDNA and genomic clones are shown at the top; solid boxes represent the coding region of the cDNA.

sequenced *DHR3* cDNA is derived from nine exons spread over 18 kb, with the ORF beginning in the first exon and ending in the ninth. Because the 5′ and 3′ ends of the transcripts were not mapped, it should be emphasized that the gene may have additional noncoding exons at its 5′ or 3′ end. In addition, Northern blots (see below) suggest the existence of multiple *DHR3* transcripts, whose structures relative to that of the sequenced cDNA remain unknown.

The *DHR3* and *EcR* genes identified in our low-stringency screen are unusual among steroid receptor superfamily genes in that the sequences encoding the DNA binding domains are not interrupted by an intron. This would result in increased hybridization of the E75 DNA binding domain cDNA probe with genomic clones of these loci and would favor the isolation of these genes, as opposed to other *Drosophila* steroid receptor superfamily genes (12).

Developmental Expression of *DHR3*. We have examined the expression pattern of DHR3 mRNA during development. A *DHR3* cDNA probe was hybridized to a Northern blot bearing samples of RNA prepared from populations representing 3-hr intervals during embryonic development and 12-hr intervals thereafter until adulthood. Fig. 4 shows peaks of *DHR3* transcription at midembryogenesis (E) during each of the first two larval stages or instars (L1 and L2), at the end of the third larval instar (L3) extending into prepupal (PP) development, and during pupal (P) development. The correspondence between these peaks and the six commonly accepted peaks in the ecdysone titer (mid-E, L1, L2, terminal L3, PP, and P; refs. 30 and 31) suggests that *DHR3* transcription may be ecdysone induced. We also note in this respect that the temporal profile of *DHR3* expression is similar to that for *E74A* (4) and *E75A* (6), two early gene transcription units that are ecdysone inducible. While this similarity is not exact, it is as close as that between the two known early genes.

At least three *DHR3* transcripts of approximately 5.5, 7, and 9 kb were detected, of which the 9-kb transcript was observed only during pupal development. This is also reminiscent of the *E74* and *E75* early genes (3, 7) and of the *BR-C* early gene (8), where multiple mRNAs result from overlapping transcription units and, in the case of BR-C, from alternative splicing. At present, however, the structural difference between the three *DHR3* transcripts is unknown.

DISCUSSION

We have carried out a screen that has yielded genes for two additional *Drosophila* members of the steroid receptor superfamily, *DHR3* and the *EcR* gene that encodes the ecdysone receptor (13). Like the *E75* gene, which was used as a hybridization probe to identify them, these genes encode proteins that exhibit sequence homology to both the DNA binding and hormone binding domains characteristic of members of this family of ligand-regulated transcription factors.

Four other *Drosophila* genes have been reported that also encode proteins exhibiting sequence homology to both domains (32–35), bringing the current total for the rapidly growing bona fide *Drosophila* members of this superfamily to seven. With the exception of EcR, these proteins are orphan receptors in that their ligands have not yet been identified. This creates a discrepancy between the number of steroid receptor homologs identified in *Drosophila* and the number of known hormones that might serve as ligands for them since there are currently only two types of insect hormones that might be expected to function via such receptors: the ecdysteroids and the juvenile hormones (12, 36). Response to the ecdysteroid 20-hydroxyecdysone, which we have referred to here as ecdysone, is mediated by the products of the ecdysone receptor gene *EcR* (13). We have previously made the argument that the other known *Drosophila* members of the steroid receptor superfamily are unlikely to function as ecdysone receptors because none of the E-region homologies among the *Drosophila* group exceeds 24% identity, whereas vertebrate receptors for the same ligand show at least 75% sequence identity in the E region and even receptors with distinct but overlapping ligand specificities exhibit E-region identities above 50% (13).

A more general application of this argument leads to the expectation that the *Drosophila* receptors will bind different ligands. Hence, even if juvenile hormone functions via a receptor that is a member of the steroid receptor superfamily, it would be expected to account for no more than one of the six *Drosophila* orphan receptors. This arithmetic raises the possibility that there may exist more steroid-like hormones in insects than have been discovered by classical physiological methods. Alternatively, the low-level E-region similarity may result from other functions of the domain.

Among the *Drosophila* group, the DHR3 protein is most homologous to E75A, exhibiting 64% and 24% sequence identities for the C and E regions, and its closest relative among the vertebrate receptors is the human retinoic acid receptor hRARα, where the respective identities are 65% and 20% (37, 38). However, the receptor most closely related to DHR3 is found in the tobacco hornworm, *Manduca sexta* (39). This receptor is referred to as *Manduca* hormone receptor 3 (MHR3) because of this strong similarity. DHR3

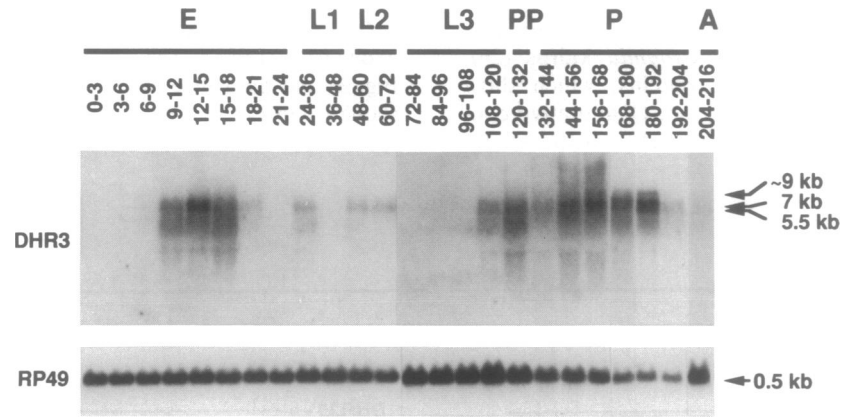

FIG. 4. Developmental profile of *DHR3* transcription. RNA prepared at various times during development was fractionated by gel electrophoresis, blotted, and probed with DNA from *DHR3* or *RP49* (16). The *RP49* gene encodes a ribosomal protein and is used as a control for loading of RNA. The hours and stages of development are indicated above the lanes. E, embryos; L1, L2, and L3, the three larval instars; PP, prepupae; P, pupae; A, adults.

and MHR3 exhibit 97% and 68% sequence identities in their putative DNA binding (C region) and ligand binding (E region) domains, making it likely that the two receptors have the same or very similar functions and are activated by the same or very similar ligands.

Interestingly, *MHR3* transcription is induced by ecdysone in a primary response, indicating that it is a target of an ecdysone–receptor complex in *Manduca* (39). This property, and the likelihood from its primary structure that MHR3 functions as a transcription factor, classifies *MHR3* as an early gene in an ecdysone-triggered genetic regulatory hierarchy, analogous to the *E74*, *E75*, and *BR-C* early genes in *Drosophila* (2–9). The strong structural similarity, and inferred functional similarity, between the DHR3 and MHR3 proteins raises the question of whether *DHR3* also functions as an early gene in the ecdysone-triggered *Drosophila* hierarchies. As indicated earlier, the temporal developmental profile of *DHR3* transcription (Fig. 4) suggests that this transcription may be ecdysone induced, although providing no hint as to whether such a response would be primary or secondary. Preliminary results from experiments in which third-instar tissues in organ culture were exposed to ecdysone, or to ecdysone plus the protein synthesis inhibitor cycloheximide, and assayed for DHR3 mRNA, indicated that *DHR3* is indeed induced by ecdysone in a primary response (M.R.K., unpublished experiments). The *DHR3* gene is located within the ecdysone-inducible transcription puff at chromosomal position 46F, which has characteristics suggesting that control of its expression may combine aspects of both classical early and late gene regulation (40).

We thank the members of the Stanford Biochemistry Department for helpful advice and discussions. This work was supported by a National Science Foundation grant (D.S.H.), a National Science Foundation Graduate Fellowship (M.R.K.), and a National Institutes of Health Graduate Fellowship (W.A.S.).

1. Ashburner, M., Chihara, C., Meltzer, P. & Richards, G. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 655–662.
2. Burtis, K. C. (1985) Ph.D. thesis (Stanford Univ., Stanford, CA).
3. Burtis, K. C., Thummel, C. S., Jones, C. W., Karim, F. D. & Hogness, D. S. (1990) *Cell* **61**, 85–99.
4. Thummel, C. S., Burtis, K. C. & Hogness, D. S. (1990) *Cell* **61**, 101–111.
5. Urness, L. D. & Thummel, C. S. (1990) *Cell* **63**, 47–61.
6. Segraves, W. A. (1988) Ph.D. thesis (Stanford Univ., Stanford, CA).
7. Segraves, W. A. & Hogness, D. S. (1990) *Genes Dev.* **4**, 204–219.
8. DiBello, P. R., Withers, D. A., Bayer, C. A., Fristrom, J. W. & Guild, G. M. (1991) *Genetics* **129**, 385–397.
9. Guay, P. S. & Guild, G. M. (1991) *Genetics* **129**, 169–175.
10. Evans, R. M. (1988) *Science* **240**, 889–895.
11. Green, S. & Chambon, P. (1988) *Trends Genet.* **4**, 309–314.
12. Segraves, W. A. (1991) *Cell* **67**, 225–228.
13. Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P. & Hogness, D. S. (1991) *Cell* **67**, 59–77.
14. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
15. Zuker, C. S., Cowman, A. F. & Rubin, G. M. (1985) *Cell* **40**, 851–858.
16. O'Connell, P. O. & Rosbash, M. (1984) *Nucleic Acids Res.* **12**, 5495–5513.
17. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
18. Cavener, D. R. (1987) *Nucleic Acids Res.* **15**, 1353–1361.
19. Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J. M. & Chambon, P. (1986) *EMBO J.* **5**, 891–897.
20. Green, S. & Chambon, P. (1987) *Nature (London)* **325**, 75–78.
21. Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R. & Sigler, P. B. (1991) *Nature (London)* **352**, 497–505.
22. Giguere, V., Hollenberg, S. M., Rosenfeld, M. G. & Evans, R. M. (1986) *Cell* **46**, 645–652.
23. Kumar, V., Green, S., Staub, A. & Chambon, P. (1986) *EMBO J.* **5**, 2231–2236.
24. Rusconi, S. & Yamamoto, K. R. (1987) *EMBO J.* **6**, 1309–1315.
25. Gronemeyer, H., Turcotte, B., Quirin, S. C., Bocquel, M. T., Meyer, M. E., Krozowski, Z., Jeltsch, J. M., Lerouge, T., Garnier, J. M. & Chambon, P. (1987) *EMBO J.* **6**, 3985–3994.
26. Nauber, U., Pankratz, M. J., Kienlin, A., Seifert, E., Klemm, U. & Jackle, H. (1988) *Nature (London)* **336**, 489–492.
27. Oro, A. E., Ong, E. S., Margolis, J. S., Posakony, J. W., McKeown, M. & Evans, R. M. (1988) *Nature (London)* **336**, 493–496.
28. Rothe, M., Nauber, U. & Jackle, H. (1989) *EMBO J.* **8**, 3087–3094.
29. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
30. Richards, G. (1981) *Mol. Cell. Endocrinol.* **21**, 181–197.
31. Handler, A. M. (1982) *Dev. Biol.* **93**, 73–82.
32. Mlodzik, M., Hiromi, Y., Weber, U., Goodman, C. S. & Rubin, G. M. (1990) *Cell* **60**, 211–224.
33. Oro, A. E., McKeown, M. & Evans, R. M. (1990) *Nature (London)* **347**, 298–301.
34. Pignoni, F., Baldarelli, R. M., Steingrimsson, E., Diaz, R. J., Patapoutian, A., Merriam, J. R. & Lengyel, J. A. (1990) *Cell* **62**, 151–163.
35. Lavorgna, G., Ueda, H., Clos, J. & Wu, C. (1991) *Science* **252**, 848–851.
36. Richards, G. (1981) *Biol. Rev.* **56**, 501–549.
37. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. (1987) *Nature (London)* **330**, 444–450.
38. Giguere, V., Ong, E. S., Sequi, P. & Evans, R. M. (1987) *Nature (London)* **330**, 624–629.
39. Palli, S. R., Hiruma, K. & Riddiford, L. M. (1992) *Dev. Biol.* **150**, 306–318.
40. Walker, V. K. & Ashburner, M. (1981) *Cell* **26**, 269–277.