Article No. jmbi.1999.2963 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 291, 71–82

JMB

The Helix-Loop-Helix Proteins dAP-4 and Daughterless Bind Both *in Vitro* and *in Vivo* to SEBP3 Sites Required for Transcriptional Activation of the *Drosophila* Gene *Sgs-4*

Kirst King-Jones, Günter Korge and Michael Lehmann*

Institut für Genetik der Freien Universität Berlin, Arnimallee 7, D-14195, Berlin, Germany

The expression of *Sgs* genes in the salivary gland of the third instar larva of *Drosophila* is a spatially restricted response to signalling by the steroid hormone 20-hydroxyecdysone. For *Sgs-4*, we have previously demonstrated that its strictly tissue and stage-specific expression is the result of combined action of the ecdysone receptor and secretion enhancer binding proteins (SEBPs). One of these SEBPs, SEBP2, was shown to be the product of the homeotic gene *fork head*. Together with SEBP3, SEBP2 appears to be responsible for the spatial restriction of the hormone response of *Sgs-4*. Here, we show that SEBP3 is a heterogeneous binding activity that consists of different helix-loop-helix (HLH) proteins. We cloned the *Drosophila* homologue of human transcription factor AP-4 (dAP-4) and identified it as one of these HLH proteins. The dAP-4 protein shows great similarity to its human and *Caenorhabditis* counterparts within the bHLHZip domain, the second leucine zipper dimerization motif, and a third region of unknown function. The expression pattern of dAP-4 indicates that it is a ubiquitously expressed HLH protein in *Drosophila*. As a second component of SEBP3 we identified the Daughterless (Da) protein, which is also ubiquitously expressed and binds to SEBP3 sites independent of dAP-4. Since both dAP-4 and Da can be detected *in situ* at transposed *Sgs-4* transcriptional control elements in polytene salivary gland chromosomes, we propose that each of the two proteins contributes to the transcriptional control of *Sgs-4*.

Keywords: AP-4; Daughterless; *Drosophila melanogaster*; HLH factors; *Sgs* genes

**Corresponding author*

Introduction

Cellular differentiation during development of multicellular organisms is accompanied by the allocation of specific sets of transcription factors to each cell type. Subsets of these transcription factors, in an interplay with signals acting on the cell from the outside, often control specific physiological functions exerted by differentiated cells. It is believed that genes responsible for these physiological functions recruit, by their transcriptional control regions, combinations of signal-controlled and constitutive transcription factors, and are thus activated in a spatially and temporally restricted manner. A particularly well-defined system to test this concept are the salivary gland secretion protein genes *(Sgs)* of *Drosophila melanogaster*, which are expressed exclusively in the salivary gland of the third instar larva, and thus exhibit a simple, strictly tissue-specific and developmental stage-specific expression pattern. The *Sgs* genes, exemplified by *Sgs-4*, are activated and probably also repressed in response to signalling by the steroid hormone 20-hydroxyecdysone (for a review, see Lehmann, 1996). The spatial and temporal specificity of the hormone response appears to be established by three putative transcription factors, termed secretion enhancer binding proteins

Abbreviations used: dAP-4, *Drosophila* activator protein 4; SEBP, secretion enhancer binding protein; Sgs, salivary gland secretion protein; bHLH, basic helix-loop-helix; bHLHZip, bHLH leucine zipper; Da, Daughterless; FKH, fork head.

E-mail address of the corresponding author: MLehm@genetik.biologie.fu-berlin.de

0022-2836/99/310071–12 $30.00/0

(SEBPs) 1-3, which bind together with the heterodimeric ecdysone receptor to closely spaced sites within the transcriptional control region of *Sgs-4* (Hofmann & Lehmann, 1998; Lehmann & Korge, 1995, 1996). Knockouts of these sites by point mutations indicate that they contribute in a synergistic manner to the transcriptional activation of *Sgs-4*. These studies suggest that SEBP1 is required for a transcriptional switch between *Sgs-4* and a neighboring gene, *Pig-1* (Hofmann & Lehmann, 1998). While the gene encoding SEBP1 has yet to be identified, SEBP2 was shown to be the product of the homeotic gene *fork head* (*fkh*) (Lehmann & Korge, 1996). The *fkh* gene product (FKH) is required for development of the ectodermal portions of the gut, the Malpighian tubules and the salivary gland (Jürgens & Weigel, 1988). The restricted expression of *fkh* in tissues of third instar larvae suggests that it is a major tissue specifier of *Sgs-4* expression. SEBP3, finally, binds to a high and a low-affinity site, each of which is closely associated with SEBP2/FKH binding sites in the *Sgs-4* transcriptional control region. The salivary gland-specific expression of SEBP3 in third instar larvae suggests that it determines, together with SEBP2/FKH, the tissue-specificity of the *Sgs-4* hormone response.

An E-box, a core motif generally recognized by basic helix-loop-helix (bHLH) proteins (Church *et al*., 1985; Ephrussi *et al*., 1985), is present in the high-affinity SEBP3 site, indicating that SEBP3 may belong to this family of transcription factors. Members of the bHLH family (for reviews, see Littlewood & Evan, 1994; Murre *et al*., 1994) possess a dimerization interface, the HLH motif, that allows homodimerization, as well as heterodimerization with other HLH proteins. Dimerization is a prerequisite for DNA-binding by a conserved basic region located immediately N-terminal to the HLH motif. Members of the bHLHZip subgroup of bHLH proteins, like Myc and AP-4, possess a leucine zipper as a second dimerization motif immediately C-terminal to the HLH motif. HLH proteins can be categorized into two major classes (Murre *et al*., 1989). In the first class (class A), which includes for instance USF and Daughterless (Da), proteins are ubiquitously expressed (Cronmiller & Cummings, 1993; Sirito *et al*., 1994). In contrast, expression of proteins of the second class (class B), such as the HLH proteins of the *achaete-scute* complex, is restricted to certain cell types. It is believed, however, that tissue-specific gene activation by these HLH proteins is achieved only after heterodimerization with class A HLH proteins. Ubiquitously expressed HLH proteins should therefore be expected to have pleiotropic functions. This is confirmed by studies on class A HLH proteins like Da, which proved to be involved in processes as diverse as oogenesis, sex determination and neurogenesis (Caudy *et al*., 1988a,b; Cline, 1983, 1989; Cummings & Cronmiller, 1994; Vaessin *et al*., 1994).

Here, we show that SEBP3 is a heterogeneous binding activity that contains two HLH proteins, Da and the *Drosophila* homologue of human transcription factor AP-4 (dAP-4). Since dAP-4 is widely expressed in different tissues and developmental stages, it belongs *bona fide*, like Da, to the class A subgroup of ubiquitously expressed HLH proteins. Da and dAP-4 form two independent but similarly sized protein-DNA complexes in which they are probably heterodimerized with an as yet unknown class B HLH protein(s) of a more restricted expression pattern. Binding of dAP-4 and Da to the SEBP3 sites of *Sgs-4* is detected *in vitro*, and *in situ* in the polytene salivary gland chromosomes, suggesting that each of the two HLH proteins is involved in the transcriptional control of *Sgs-4*.

Results

SEBP3 binding sites are putative target sites of the *Drosophila* homologue of human transcription factor AP-4

To elucidate the molecular nature of SEBP3, we searched the TRANSFAC database (Wingender *et al*., 1996) for transcription factors with recognition sites similar to those of SEBP3. This search revealed a striking similarity between the high-affinity SEBP3 site of *Sgs-4* and binding sites of human transcription factor AP-4 (Figure 1). AP-4 is a bHLHZip protein that contains a second leucine zipper in addition to the leucine zipper motif of the bHLHZip region (Figure 2(a)). Since similarity between the SEBP3 and AP-4 sites clearly extends beyond the E-box core motif (Figure 1), we decided to take a closer look at the relationship between SEBP3 and the *Drosophila* homologue of AP-4 (referred to as dAP-4). A database search revealed that a P1 clone containing the *dAP-4* gene had been sequenced by the Berkeley *Drosophila* Genome Project. Likewise, the *AP-4* gene of *Caenorhabditis elegans* had been sequenced by the *Caenorhabditis* genome project (Figure 2(b)). The predicted amino acid sequence of dAP-4 shows great similarity to the human and *C. elegans* homologues within the bHLHZip domain (69 % identity with human AP-4, 56 % with *Caenorhabditis* AP-4), the second leucine zipper (35 % identity with human AP-4, 24 % with *Caenorhabditis* AP-4), and a third region of unknown function, which we call the TIV motif (65 % identity with human AP-4, 60 % with *Caenorhabditis* AP-4). A database search with the TIV motif did not reveal any match with other proteins that we consider to be of significance (see Discussion).

dAP-4 and SEBP3 have very similar DNA-binding specificity

In order to test our hypothesis that SEBP3 is identical with dAP-4, we constructed an expression plasmid that encodes the full-length dAP-4 protein

origin	sequence
SV40: 293/306	GAACCAGCTGTGGA
Py I: 18/31	AAAACAGCTGTTGT
Enk: -91/-78	GCGTCAGCTGCAGC
Py II: 5176/5189	CGGCCAGCTGCGGT
Tax: -133/-146	CCACCAGCTGCCGT
consensus	A/G NCAGCTG T/C NG
Sgs-4: -426/-413 (strong)	CCAACAGCTGCGGT
Sgs-4: -348/-335 (weak)	AAAACATACCTTTC
Sgs-4^{d}: -348/-335 (no binding)	↓ T

Figure 1. Comparison of experimentally characterized AP-4 recognition sequences with SEBP3 binding sites of *Sgs-4*. The AP-4 sites were identified in simian virus 40 (SV40), polyoma virus (Py I and Py II) (Mermod *et al.*, 1988), the human *proenkephalin* gene (*Enk*) (Comb *et al.*, 1988), and bovine leukemia virus *Tax* gene (Unk *et al.*, 1994). The position of each of the sequences is given after the colon, and a consensus sequence is shown. The central E-box and a G located three nucleotides further 3′, which are found in all the AP-4 sites and the strong SEBP3 site of *Sgs-4*, are shaded. Nucleotides of the AP-4 sites that match nucleotides at equivalent positions of the strong SEBP3 site are underlined. SEBP3 binding sites of *Sgs-4* are shown below the consensus sequence. *Sgs-4*^{d} is the *Sgs-4* allele of the strain Samarkand. This allele differs from that of strain Oregon R, on which the studies described here and elsewhere (Hoffmann & Lehmann, 1998; Lehmann & Korge, 1995, 1996) are based, by a C to T transition at position −344 within the weak SEBP3 binding site (Hofmann *et al.*, 1987). When this transition is introduced into an oligonucleotide containing this binding site (Org-344), the resulting oligonucleotide (Sam-344) loses its ability to bind SEBP3 (see Figure 3).

(see Materials and Methods). This plasmid served to produce recombinant dAP-4 in bacteria for DNA-binding studies. In addition, we isolated 12 *dAP-4* cDNAs from embryonic and third larval instar cDNA libraries, to confirm the predicted structure of the *dAP-4* gene. The sequences and sizes of these cDNAs are consistent with the transcript structures deduced from the genomic sequence and from Northern analysis (Figure 2(a), and data not shown).

The DNA-binding specificity of the bacterially expressed dAP-4 protein was analysed and compared to that of SEBP3 using the mobility shift assay. These experiments showed that dAP-4 binds to several wild-type and mutated SEBP3 binding sites with the same specificity as SEBP3. As expected, it binds to the high-affinity SEBP3 site of *Sgs-4*. This binding is prevented by two point mutations that also interfere with SEBP3 binding (Figure 3, lanes 2-4 and 8-10). Most importantly, dAP-4 also binds to the low-affinity SEBP3 binding site (see Figure 1), which lacks a canonical E-box and is of little similarity to the high-affinity site (Figure 3, lanes 5-6 and 11-12). Moreover, dAP-4 and SEBP3 show an equally diminished binding affinity to the high-affinity SEBP3 site when the central 5′-GC dinucleotide of the E-box is reversed to 5′-CG (oligonucleotide O5ME, Figure 3, lanes 7 and 13). Thus, the mobility shift data do not indicate any differences in the binding specificities of dAP-4 and SEBP3. However, they clearly reveal a difference in the electrophoretic mobilities of the two proteins. Since human AP-4 binds, like HLH proteins in general, to DNA as a dimer, this probably holds true for the bacterially produced dAP-4 protein. Since the latter migrates clearly faster than SEBP3, the simplest interpretation of our results is that dAP-4 is a component of SEBP3, but that it is heterodimerized in SEBP3 with a partner of greater molecular mass.

dAP-4 is a ubiquitously expressed bHLHZip protein in *D. melanogaster*

To further test the hypothesis that dAP-4 is part of SEBP3, we produced two types of polyclonal antibodies against the dAP-4 protein. One antibody (SA4114) is directed towards an N-terminal 274 amino acid residue polypeptide, which includes the bHLHZip domain, and the other (SA2720) towards a 498 residue C-terminal polypeptide lacking this domain. Using SA2720, we first tested whether dAP-4 can be detected in salivary glands of third instar larvae. In Western blot analyses, the antibody reacts with an 85 kDa protein that is present in nuclear extracts from salivary glands and embryos (Figure 4(a), lanes 1-2). The molecular mass of this protein is somewhat greater than the 67 kDa predicted from the *dAP-4* sequence data. However, dAP-4 obtained by *in vitro* translation as well as bacterially produced dAP-4 show the same reduced mobility during SDS-PAGE (Figure 4(a), lanes 3 and 7), supporting the conclusion that the protein detected in nuclear extracts is indeed dAP-4. To further analyse the intracellular and tissue distribution of dAP-4, we performed whole-mount *in situ* immunostainings of tissues of third instar larvae. These stainings indicate that dAP-4 is a nuclear protein. We could detect strong nuclear staining in all tissues that we examined, among them the salivary glands, the fat body, the ring gland (Figure 4(b)), the nervous system, the gut, the Malpighian tubules, the tracheal epithelium, and the imaginal discs as well as other imaginal progenitor cells (data not shown). The dAP-4 protein thus belongs *bona fide* to the class A of ubiquitously expressed HLH proteins. This conclusion is supported by a developmental Northern analysis (data not shown), which detects *dAP-4* mRNA in all stages examined (embryo, first, second and third instar larva, prepupa, adult).

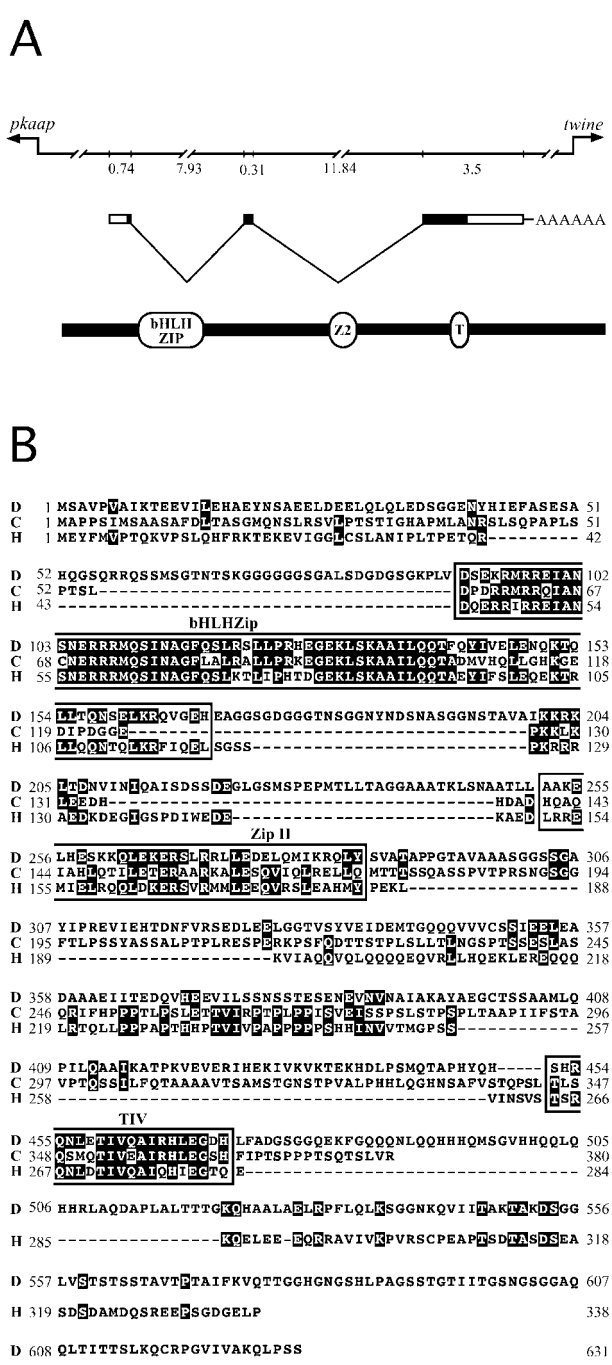

Figure 2. Genomic organization of the *Drosophila AP-4* gene, structure of gene products, and comparison with human and nematode AP-4. (a) The *dAP-4* transcription unit is located on the second chromosome at 35F1-2 flanked distally by *twine* and proximally by a gene that is homologous to the human *pkaap* (protein kinase A anchoring protein) gene. It spans about 25 kb and is split by two introns of 7.9 and 11.8 kb, respectively. Structural motifs of the dAP-4 protein, outlined below the primary transcript structure, are the bHLHZip domain and a second leucine zipper motif (Z2). In addition, dAP-4 contains a third conserved region, the TIV motif (T), which is also present in the human and nematode AP-4 proteins (see below). (b) Comparison of the predicted amino acid sequences of the *Drosophila* (D), *Caenorhabditis* (C) and human (H) AP-4 proteins (Hu *et al*., 1990). Note that dAP-4 is almost twice as large as the human and nematode counterparts. Filled boxes indicate amino acid residues that are identical in at least two of the aligned sequences. The bHLHZip domain, the second leucine zipper (Zip II), and the TIV motif are marked by open boxes. Although the bHLHZip domain of the *Caenorhabditis* protein contains an incomplete leucine zipper motif, the high degree of similarity (see the text) between all three proteins within these domains infers that the *Drosophila* and *Caenorhabditis* proteins represent the homologues of human AP-4.

SEBP3 is a heterogeneous binding activity containing dAP-4 as one of at least two DNA-binding components

After having verified that dAP-4 is indeed present in the salivary gland, we asked whether anti-dAP-4 antibodies can interfere with SEBP3 complex formation or alter the mobility of existing SEBP3-DNA complexes. Figure 5(a) (lanes 8-9) shows that both types of anti-dAP-4 antibodies show only a partial reaction with SEBP3 from salivary glands. In a parallel experiment, the same amount of antibody proved to be sufficient to almost completely interfere with an embryonic binding activity that migrates with a greater mobility than SEBP3 (Figure 5(a), lanes 3-4). Since this binding activity shows the same mobility as bacterially produced dAP-4, it most likely represents a dAP-4 homodimer. The specificity of the antibody reactions is indicated by the formation of supershifts by antibody SA2720, which recognizes epitopes outside of the DNA-binding domain (Figure 5(a), lanes 3 and 8), while antibody SA4114, which is directed towards the DNA-binding domain, essentially interferes with complex formation (Figure 5(a), lanes 4 and 9). However, the antibody reaction with the SEBP3 complex was not complete and could not be improved by raising antibody concentrations (data not shown), suggesting that SEBP3 is a heterogeneous binding activity composed of at least two different DNA-binding proteins.

A test of this hypothesis was facilitated by detection of a new SEBP3 binding site in the upstream region of *Sgs-1*. (Roth *et al*., 1999). When an oligonucleotide containing this site (O1) was used to compete for SEBP3 binding to the high-affinity SEBP3 site of *Sgs-4*, a weak complex remained that corresponds to the upper portion of the SEBP3 complex (Figure 5(b), lane 10). This complex has the same slightly reduced mobility as the complex that remained after addition of anti-dAP-4 antibodies (cf. Figure 5(a), lanes 8 and 9) and it clearly migrates more slowly than the complex remaining after competition with oligonucleotide O5ME (Figure 5(b), lane 11). O5ME and the SEBP3 binding site of *Sgs-1* thus appear to preferentially bind two components of SEBP3, SEBP3A and SEBP3B, which migrate with slightly different mobilities.

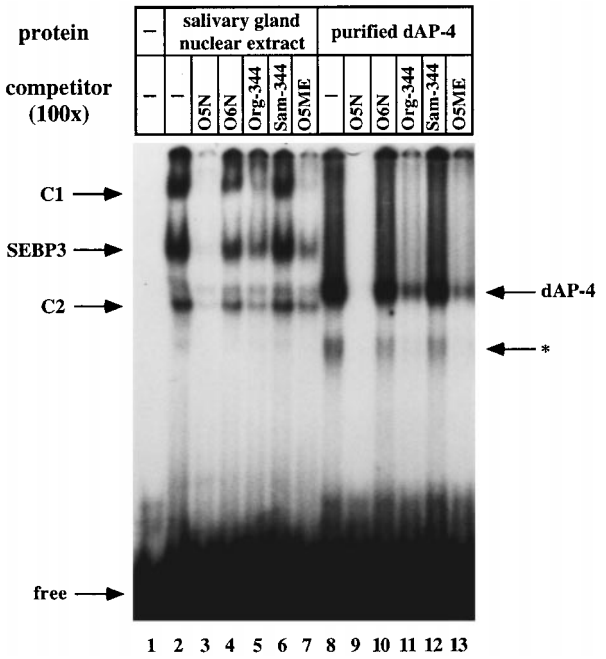

Figure 3. dAP-4 binds to DNA with a specificity very similar to that of SEBP3. Binding of bacterially produced and affinity-purified dAP-4 to a radiolabelled oligonucleotide containing the strong SEBP3 site of *Sgs-4* (O5N) was compared to binding of salivary gland nuclear extract proteins to the same oligonucleotide using the mobility shift DNA-binding assay. To probe for differences in the SEBP3 and dAP-4 binding specificities, different unlabelled competitor oligonucleotides were added to the binding reactions at a 100-fold molar excess. O6N contains two base exchanges within the strong SEBP3 binding site, but is otherwise identical with O5N. Org-344 contains the weak SEBP3 site of *Sgs-4*. Sam-344 is identical with this oligonucleotide, except that it contains the base exchange of the *Sgs-4*^{d} allele (see Figure 1). In O5ME, which is otherwise essentially identical with O5N (see Materials and Methods), the central 5′-GC-dinucleotide of the E-box is reversed to 5′-CG. The competition patterns do not indicate that SEBP3 and dAP-4 differ in their binding specificity. However, note that dAP-4 migrates clearly faster than SEBP3. The salivary gland nuclear extract forms two additional complexes, C1 and C2, which appear to have similar sensitivity towards competitor DNAs as the SEBP3 complex. These complexes have not been observed in previous studies in which SEBP3 binding was analysed using a different oligonucleotide that contained a SEBP2 site in addition to the strong SEBP3 site (Lehmann & Korge, 1995, 1996). Complex C1 may have been masked in these studies by a slowly migrating complex formed by non-specific binding proteins. Since it behaves very similarly towards competitor DNAs as SEBP3, C1 may represent a higher-order complex in which SEBP3 is associated with additional protein(s). Complex C2 may have escaped notice in earlier studies, since it is formed only in the presence of high concentrations of NP-40 and DTT (see Materials and Methods). It shows a similar mobility as bacterially expressed dAP-4, but antibody studies indicate that dAP-4 is not responsible for formation of this complex (see Figures 5(a) and 6). Bacterially expressed dAP-4 forms a fast-migrating complex, marked by an asterisk (*), which is probably formed by the truncated product that can be detected on Western blots (see Figure 4(a)). The major complex formed by bacterially produced dAP-4 has the same mobility as dAP-4 contained in nuclear extracts from embryos (cf. Figure 5).

If it is true that SEBP3B is the dAP-4-containing component of SEBP3, as suggested by its apparent absence after addition of anti-dAP-4 antibodies (Figure 5(a), lanes 8 and 9), the SEBP3A complex, preferentially remaining after competition with the SEBP3 site of *Sgs-1*, should not react with the anti-dAP-4 antibodies, whereas the SEBP3B complex, preferentially remaining after competition with O5ME, should show such a reaction. Figure 6(a) demonstrates that this is indeed the case. In the presence of a 100-fold molar excess of oligonucleotide O1, the remaining SEBP3-DNA complex was not significantly disturbed by addition of either anti-dAP-4 or preimmune antibodies (Figure 6(a), lanes 3-5). In striking contrast, the complex remaining after competition with oligonucleotide O5ME was clearly diminished by the anti-dAP-4 but not by the preimmune antibodies (Figure 6(a), lanes 6-8). These results strongly suggest that SEBP3B, but not SEBP3A, does contain dAP-4 as a DNA-binding component.

Daughterless and dAP-4 are two independent binding components of SEBP3

Even though SEBP3A does not contain dAP-4, our competition studies (Figure 3) show that the DNA-binding component of SEBP3A should recognize DNA with a binding specificity very similar to that of dAP-4. Interestingly, it has been shown that the mammalian homologues of the *Drosophila* Da protein, E12 and E47, have a binding specificity very similar to that of AP-4 (Nelson *et al*., 1990). Moreover, it is known that Da is expressed in the larval salivary gland (Cronmiller & Cummings, 1993). We therefore tested the effects of anti-Da antibodies and of a combination of anti-dAP-4 and anti-Da antibodies on formation of the SEBP3-DNA complex. Figure 6(b) shows that an affinity-purified polyclonal anti-Da antibody clearly reduced the amount of the SEBP3 complex (lanes 5 and 7). When this antibody was used together with the anti-dAP-4 antibody SA2720, the SEBP3 complex was almost completely abolished (Figure 6(b), lane 6). These results strongly suggest that Da is indeed the DNA-binding component of SEBP3A. Since Da is a ubiquitously expressed bHLH protein that exerts its specific functions through heterodimerization with other HLH proteins of a more restricted expression pattern (see Introduction), this is probably the case in the salivary gland. As we hitherto could detect SEBP3 binding activity only in the salivary gland, expression of this partner, like expression of the putative dimerization partner of dAP-4, should be confined to this tissue or to only a few larval tissues.

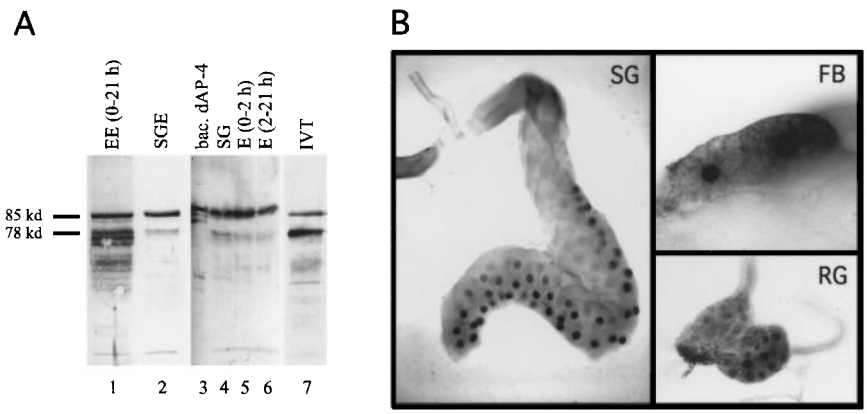

Figure 4. dAP-4 is a nuclear protein present in various tissues of third instar larvae and in embryos. (a) dAP-4 is detected by Western analysis as an 85 kDa protein in nuclear extracts from embryos (EE, lane 1) and salivary glands (SGE, lane 2). The protein can also be detected when extracts from whole embryos of different stages and whole salivary glands are used (lanes 4-6). Lane 3 shows the Western analysis of bacterially produced and affinity-purified dAP-4 and lane 7 an autoradiography of [^{35}S]methionine-labelled dAP-4 produced by *in vitro* translation. A 78 kDa protein, which is the predominating product of the *in vitro* translation reaction, is detected as a minor product in embryos and salivary glands. The size of this product is consistent with the use of an alternative start codon 60 codons downstream of the first AUG of the open reading frame. Note that additional minor bands that are observed with the embryo nuclear extract are absent when whole embryos are used, suggesting that these bands represent degradation products. (b) Whole-mount immunodetection of dAP-4 with anti-dAP-4 antibody SA2720 in tissues dissected from third instar larvae. A strong nuclear staining is observed in the salivary gland (SG), the fat body (FB) and the ring gland (RG). Nuclear staining is absent when preimmune serum is used instead of the anti-dAP-4 antiserum (not shown). The seemingly uneven distribution of dAP-4 in salivary gland and ring gland rather reflects local differences in staining conditions than true differences in the expression of dAP-4 within these tissues.

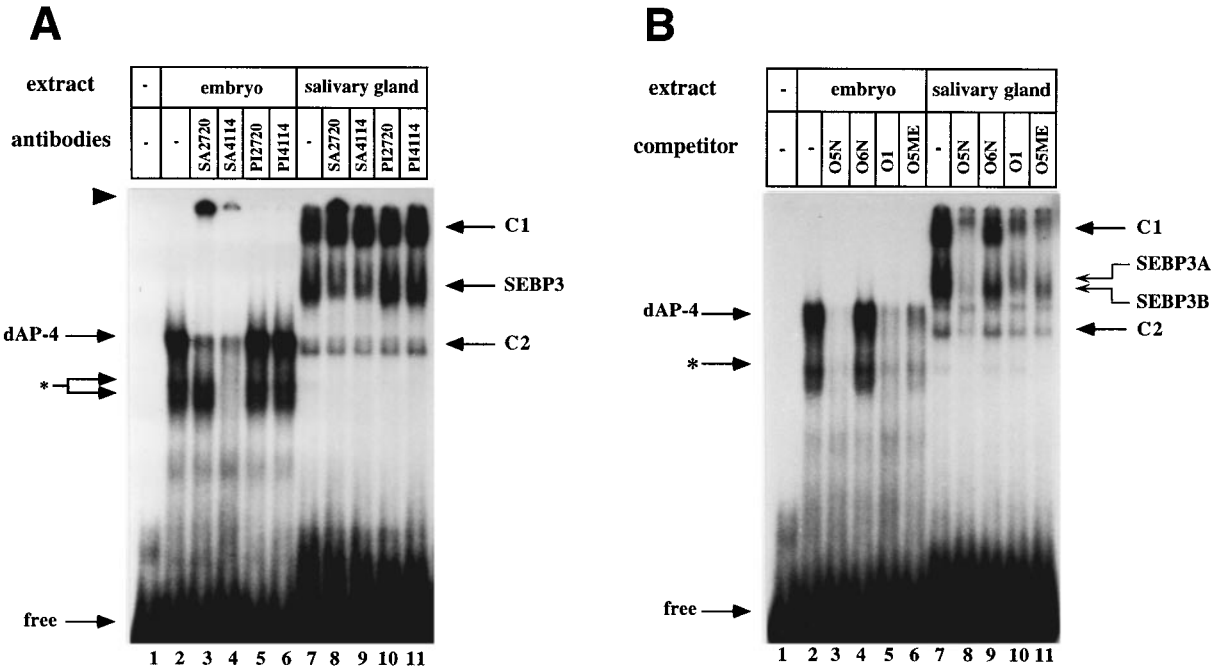

Figure 5. SEBP3 is a heterogeneous binding activity consisting of SEBP3A and SEBP3B. (a) Nuclear extract from embryos and salivary glands was incubated with radiolabelled oligonucleotide O5N in the absence or presence of the indicated anti-dAP-4 (SA) and preimmune (PI) antibodies. Protein-DNA complexes and free DNA were then separated in a mobility shift gel. Either of the two anti-dAP-4 antibodies almost completely reacts with an embryonic binding activity, indicating that this activity is formed by dAP-4. Two faster migrating complexes, marked by an asterisk (*), are disturbed only by antibody SA4114, indicating that they are formed by degradation products that lack epitopes recognized by SA2720. From the three complexes formed by salivary gland nuclear extract, only the SEBP3 complex is diminished in the presence of the dAP-4 antibodies. The arrowhead marks supershifted complexes formed only in the presence of antibody SA2720. (b) Nuclear extract from embryos and salivary glands was incubated with radiolabelled O5N in the absence or presence of the indicated competitor olignucleotides. O1 contains a SEBP3 binding site of *Sgs-1*. The other oligonucleotides are described in the legend to Figure 3. Note that O1 selectively competes for formation of the lower portion of the SEBP3-DNA complex, while O5ME preferentially competes for formation of the upper portion of the complex (lanes 10-11). Since O1 contains a better binding site for dAP-4 than O5ME (compare lanes 5 and 6), this suggests that SEBP3 includes a dAP-4-containing component, SEBP3B, which migrates slightly faster than a second component of this binding activity, SEBP3A.

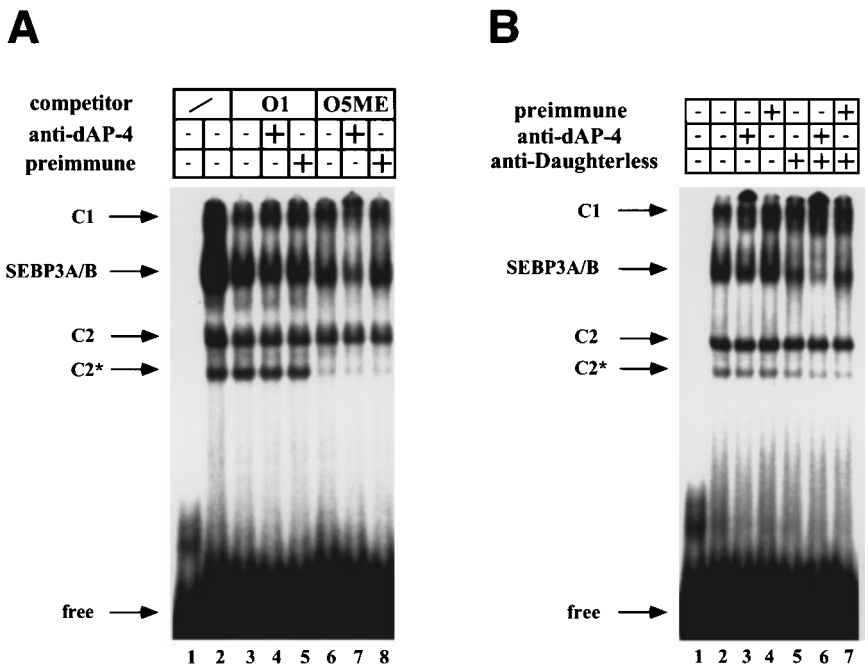

Figure 6. Daughterless is a component of SEBP3A and dAP-4 is involved in the formation of SEBP3B. (a) Salivary gland nuclear extract was equilibrated for mobility shift analysis with radiolabelled O5N in the absence or presence of a 100-fold molar excess of oligonucleotides O1 or O5ME. The contribution of dAP-4 to formation of the complexes that persist in the presence of competitor DNA was then assayed by addition of anti-dAP-4 (SA4114) or preimmune antibodies as indicated. A clear antibody reaction is observed only with the SEBP3 complex that remains visible in the presence of O5ME and that has been marked as SEBP3B in Figure 5(b). In this assay, an additional complex, marked as C2*, is particularly prominent that usually does not show up or appears only as a faint band in our mobility shift assays. Note that the protein responsible for formation of this complex shows strong binding to O5ME but virtually no binding to O1, and thus clearly differs in its binding specificity from dAP-4. (b) Salivary gland nuclear extract was incubated with radiolabelled O5N in the absence or presence of anti-dAP-4 antibody SA2720, an affinity-purified anti-Da-antibody or preimmune antibodies (PI2720) in the indicated combinations. The resulting complexes were then separated in a mobility shift gel. The SEBP3-DNA complex is only attenuated in the presence of either anti-dAP-4 or anti-Da antibody alone (lanes 3 and 5). However, a combination of both antibodies almost completely abolishes complex formation (lane 6). The weak complex remaining in lane 6 was reproducibly observed, suggesting that a third independent component beside dAP-4 and Da might be involved in SEBP3 complex formation.

Both Daughterless and dAP-4 can be detected at transposed *Sgs-4* regulatory elements *in situ*

It has been shown that both the high-affinity and the low-affinity SEBP3 site are required for full transcriptional activation of *Sgs-4* (Hofmann *et al*., 1987; Lehmann & Korge, 1995). Here, we show that two different HLH proteins that are expressed in the salivary gland can bind at least to the high-affinity SEBP3 site *in vitro*. Therefore, the question has to be addressed of whether binding of one of these proteins or of both is indeed responsible for the transcriptional activation of *Sgs-4*. Since homozygous *dAP-4* and *da* null mutants escape an analysis of *Sgs* expression by their early lethality, we asked whether bound dAP-4 and/or Da can be detected at the *Sgs-4* regulatory region *in vivo* at the time when the *Sgs* genes are actively transcribed. To answer this question, we performed immunofluorescence stainings of salivary gland polytene chromosomes from third instar larvae using anti-dAP-4 and anti-Da antibodies. Since these antibodies stain hundreds of loci on the polytene chromosomes (Figure 7(a), and data not shown; cf. Cronmiller & Cummings, 1993), we used transgenic lines carrying P-elements with four copies of the *Sgs-4* regulatory region to be able to unequivocally correlate immunofluorescence signals with the presence of the *Sgs-4* regulatory region. Figure 7 shows that specific fluorescence signals were obtained with both anti-Da and anti-dAP-4 antibodies *in situ* at these transposed *Sgs-4* elements. It should be noted, however, that signals obtained with the anti-dAP-4 antibody were clearly weaker than those produced by the anti-Da-antibody, suggesting that Da might be preferentially bound *in vivo*. These data are consistent with a model in which both Daughterless and dAP-4 are involved in the transcriptional control of *Sgs-4*.

Discussion

In this study, we show that two ubiquitously expressed HLH proteins, Da and dAP-4, are independent DNA-binding components of SEBP3, a previously identified DNA-binding activity that appears to be required for transcriptional activation of the *Drosophila* gene *Sgs-4*. Since SEBP3 binding activity is specific for the salivary gland, we postulate that both Da and dAP-4 are heterodimerized in SEBP3 with another HLH protein of a more restricted expression pattern. Alternatively, SEBP3A and/or SEBP3B may represent ternary or higher-order complexes in which Da and/or dAP-4 homodimers are bound by one or more additional factors. In particular, this model has to be considered for the dAP-4-containing SEBP3B complex, as the leucine zipper motifs of human AP-4 have been shown to promote homodimer formation rather than heterodimerization (Hu *et al*., 1990). In addition, human AP-4 proved to be unable to form heterodimers with several other HLH or leucine

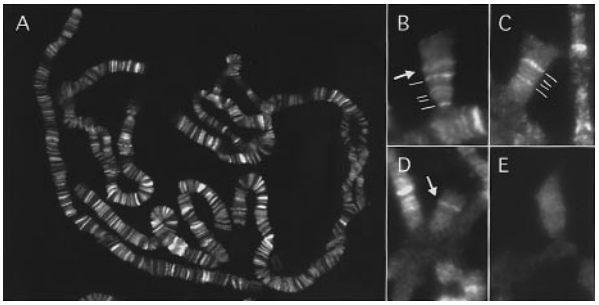

Figure 7. Both dAP-4 and Da bind *in situ* to the transposed *Sgs-4* regulatory region in polytene salivary gland chromosomes. (a) Immunofluorescence staining of a complete set of polytene salivary gland chromosomes of a wild-type larva using the anti-dAP-4 antibody SA2720. (b) and (d) Immunofluorescence staining of the fourth chromosome of a transformant strain carrying an integration of four copies of the *Sgs-4* regulatory region in this chromosome at position 102D3-5. In (b) anti-dAP-4 antibody SA2720 was used for staining, and in (d) a monoclonal anti-Da-antibody. (c) and (e) Control stainings of fourth chromosomes of wild-type larvae using these antibodies. White arrows mark immunofluorescence signals coming from the P-element integration site bearing the *Sgs-4* elements. At integration sites of P-elements lacking the *Sgs* elements, such signals are not observed. For localization of the immunofluorescence signals, chromosomes were counterstained with Hoechst 33258 (not shown). Four thin lines mark signals that are observed also when wild-type chromosomes are stained with the anti-dAP-4 antibody. In contrast, the anti-Da antibody does not stain specific sites on the fourth chromosome of wild-type animals.

zipper proteins, including the mammalian homologue of Da, E12, and the Id proteins Id1 and Id2 (Hu *et al*., 1990; Sun *et al*., 1991). In contrast to AP-4, it is well-established that Da acts through binding to other HLH proteins. In neurogenesis and sex determination the function of Da depends on the formation of heterodimers with proteins encoded by the *achaete-scute* complex (Cabrera & Alonso, 1991; Van Doren *et al*., 1991), and it has been shown that Da can interact with a subset of HLH proteins from the *Enhancer of split* complex (Alifragis *et al*., 1997).

The salivary gland-specific expression of SEBP3 suggests that it is, together with FKH, responsible for the tissue specificity of the *Sgs-4* hormone response. Interestingly, such a spatial restriction of gene expression through cooperation of HLH proteins and FKH has recently been shown to take place during differentiation of the neural tip cell that controls cell proliferation in the Malpighian tubules of *Drosophila*. The expression of *Krüppel* is restricted to the tip cell by direct binding of FKH and proneural HLH proteins to a common *cis*-acting element (Hoch & Jäckle, 1998). In this system, the spatial information provided by FKH, which is expressed in all Malpighian tubule cells, is further refined by the tip cell-specific expression of the proneural HLH proteins. Since FKH is not exclusively expressed in the salivary glands of third instar larvae (Kuzin *et al*., 1994), a similar mechanism may operate to confine glue gene expression to the salivary glands. A test of this hypothesis will require the identification of the heterodimerization partner(s) of Da and dAP-4 in the salivary gland and the analysis of their expression patterns.

We show here that both dAP-4 and Da are able to bind *in vitro* to the strong SEBP3 site that has previously been shown to be required for full transcriptional activation of *Sgs-4* (Lehmann & Korge, 1995). Also, the weak SEBP3 site, which contributes to the transcriptional activation of *Sgs-4* as well, albeit to a lesser extent (Hofmann *et al*., 1987), is recognized *in vitro* by both dAP-4 and Da (this study, and data not shown). While the *in vitro* binding data do not indicate that either of the proteins is bound preferentially, the *in situ* antibody staining data seem to indicate that most of the SEBP3 sites are occupied by Da (SEBP3A) *in vivo*. However, it is not clear whether the observed difference in signal intensities reflects different amounts of antigen bound to the *Sgs-4* elements, or simply the different quality of the antibodies used. Whatever the reason for the stronger response to the anti-Da-antibodies is, the polytene chromosome stainings clearly show that also dAP-4 is bound to the transposed elements. The *in situ* data therefore support the notion that both proteins are involved in the transcriptional control of *Sgs-4 in vivo*.

The binding sites recognized by dAP-4 and Da harbour a classical E-box motif (high-affinity SEBP3 site) and a non-canonical E-box related motif (low-affinity SEBP3 site). Interestingly, the SV40 enhancer also contains a high-affinity canonical E-box motif and a low-affinity non-E-box site, which are both recognized by human AP-4 (Mermod *et al*., 1988). Such binding of HLH proteins to non-E-box motifs has also been observed with other HLH proteins, like for instance the Myc-Max heterodimer (Blackwell *et al*., 1993). Although Myc-Max binds with lower affinity to these non-canonical sites as to canonical E-boxes *in vitro*, some of them seem to be the predominant *in vivo* binding sites (Grandori *et al*., 1996). As some of the non-canonical sites are not recognized by other HLH proteins, which are otherwise able to bind to canonical Myc-Max sites, it has been proposed that non-canonical sites contribute to the specificity of target gene recognition by HLH proteins (Grandori & Eisenman, 1997). It may be speculated that the weak SEBP3 site serves a similar role. But why should two independent HLH protein-containing complexes be recruited by the SEBP3 sites? In the light of the Myc-Max studies, we cannot exclude that indeed each of the two SEBP3 sites preferentially binds either SEBP3A or SEBP3B *in vivo*, possibly leading to a synergistic activation of *Sgs-4* transcription. Another possibility is that the two HLH proteins have additive

effects ensuring mass production of the *Sgs*-encoded glue proteins by the polytene salivary gland tissue. The number of binding sites that have to be saturated to achieve maximal transcription in this tissue is about 1000-fold higher than in a diploid tissue. Recruitment of more than one HLH protein may thus have been necessary for maximal transcriptional activation. Although we favour this explanation, we cannot exclude, for instance, that one of the HLH complexes acts positively while the other acts negatively. Further functional analyses will help to distinguish between these and other possibilities.

An interesting result of our study is the identification of the TIV motif, a well-conserved AP-4 protein region of as yet unknown function that constitutes a promising target for mutational analyses. A database search revealed that a sequence (LETIVQ) present in the TIV motif of dAP-4 is also found in vertebrate transcription factor NF-E2 (Andrews *et al*., 1993; Chan *et al*., 1993). However, since the corresponding regions of NF-E2-related transcription factors do not show such a similarity to the TIV motif, the limited similarity between dAP-4 and NF-E2 is likely to be insignificant. Apart from the TIV motif, the comparison of the AP-4 homologues of *Drosophila* and *Caenorhabditis* with human AP-4 reveals the expected high degree of similarity within the bHLHZip domain and the second leucine zipper, but not outside of these regions.

One puzzling finding of our study is that dAP-4 from salivary glands obviously does not bind to SEBP3 sites as a homodimer, at least not without being associated with other proteins (see above). This is in striking contrast to embryonic and bacterially produced dAP-4, which give complexes that are most likely formed by homodimers. One possible explanation for this finding is that dAP-4 monomers are completely titrated off by a high-affinity heterodimerization partner in the salivary gland. If this is true, an excess of this partner in the salivary gland should lead to the formation of additional slowly migrating complexes when salivary gland nuclear extract is mixed with nuclear extract from embryos. However, such mixing experiments led to negative results (data not shown). As an alternative explanation, it may therefore be considered that dAP-4 is modified post-translationally in the salivary gland leading to a form that retains its ability to form heterodimers but is unable to form homodimers. These considerations make it clear that a search for interacting partners of dAP-4 and also of Da in the salivary gland should be the next step taken to clarify the mechanisms by which these two HLH proteins contribute to the transcriptional control of *Sgs-4* and other *Sgs* genes.

Materials and Methods

Flies and strains

Unless otherwise stated, the nucleotide sequences of synthetic oligonucleotides used in this work correspond to sequences of the *Sgs-4* allele of the wild-type strain Oregon R (Stanford) of *D. melanogaster*. Likewise, tissues and embryos as well as genomic DNA for DNA amplification by PCR were from this strain. For P-element transformations, we used the recipient strain *Ko; ry*^{506}, which produces almost no Sgs-4 protein and carries the *ry*^{506} allele as a marker for the selection of transformants (Krumm *et al*., 1985).

Preparation of nuclear extracts

Nuclear extracts from larval salivary glands were prepared as described (Lehmann & Korge, 1995). For the preparation of nuclei from staged embryos (0-2 hours, 2-20 hours), embryos were collected and rinsed with 0.7% (w/v) NaCl, 0.04% (v/v) Triton X-100, blotted dry and dechorionated for 90 seconds with 2.25% (v/v) sodium hypochlorite. All subsequent steps were carried out at 0-4 °C. Approximately 1 g of dechorionated embryos was homogenized in 5 ml of buffer H (10 mM Tris-HCl, 1.5 mM $MgCl_2$, 0.1 mM EDTA, 50 mM KCl, 0.25 M saccharose, pH 7.5 at 20 °C). Washing and extraction procedures for embryonic nuclei were the same as for salivary gland nuclei, except that 5 ml buffer of H and 1 ml of E500 (10 mM Tris-HCl, 1 mM EDTA, 650 mM NaCl, 5% glycerol, pH 7.5 at 20 °C) were used. Both buffers were supplemented with 1 mM DTT, 1 mM PMSF and 1 mM benzamidine.

Construction of expression vectors and production of dAP-4 fusion proteins

Recombinant dAP-4 was used for two purposes: (1) generation of polyclonal antibodies; (2) testing of DNA-binding properties in mobility shift assays. Antibodies were generated against an N-terminal polypeptide of 272 amino acid residues (APDB), which contains the DNA-binding domain, and against a C-terminal polypeptide of 498 amino acid residues (APEX3) encoded by exon III, which lacks the DNA-binding domain. A fusion protein containing the full-length 631 residue dAP-4 protein was produced for gel shift analyses.

For construction of expression plasmids, appropriate *dAP-4* DNA fragments were cloned into pET vectors (Novagen). To obtain the APDB-encoding construct, a full-length (3.9 kb) *dAP-4* cDNA was cut with *Afl*III and *Xho*I. The resulting 818 bp fragment was cloned into pET21d cut with *Nco*I and *Xho*I. The APEX3-encoding construct is based on a 1494 bp PCR fragment that was directly amplified from genomic DNA using *Pwo* polymerase (Eurogentec). For amplification, we used primer EX3KY (5′-GCGGCCATCCTACAACAGACCTTC-CAATAC-3′), which matches the 5′ end of exon III, and primer ENDX3 (5′-ATTGAATCCGGCCGACGGCAA-CTGCTTGGC-3′), which represents the end of the coding region, but is modified at two positions (underlined) to introduce an *Eag*I site. The PCR fragment was cut with *Eag*I and cloned into vector pET21c (*Bam*HI/filled-in, *Eag*I). A plasmid encoding full-length dAP-4 was constructed as follows. First, a 1353 bp *Kpn*I/*Eag*I-fragment was isolated from the APEX3-encoding construct. Then,

a 598 bp PCR fragment representing the 5′ end of *dAP-4* was amplified from an embryonic cDNA library (Clontech) using primers APBEG (5′-CAAATACACAAACACGATTAAACATGTCCGC-3′) and APGG (5′-CCGCTGGCATTGGAATCATTGTAGTTTCCG-3′), and cut with *Afl*III and *Kpn*I. The resulting 537 bp fragment was ligated to the 1353 bp fragment and subsequently cloned into *Nco*I/*Eag*I-cut pET21d.

Recombinant protein was produced essentially following the protocol of the manufacturer (Novagen). BL21 cells, transformed with the appropriate pET constructs, were grown in LB/carbenicillin, and expression of fusion protein was induced at an A_{600} of 0.6-1.0 by adding 1 mM IPTG. After three hours, cells were harvested, resuspended in binding buffer (0.5 M NaCl, 5 mM imidazole, 20 mM Tris-HCl, pH 7.9) and disrupted by sonification. To isolate native protein, the homogenate was centrifuged at 18,000 rpm (SS-34), and dAP-4 fusion protein was purified from the supernatant by affinity chromatography on a His-bind matrix-column (Novagen, 8 ml), run by an FPLC system (Pharmacia). Prior to immunization of rabbits, the buffer of the fusion protein-containing fraction was replaced by PBS through gel-filtration. The pellet obtained after centrifugation of the homogenate was resuspended in binding buffer containing 6 M guanidine hydrochloride to dissolve insoluble fusion protein deposited in the inclusion bodies. The resulting solution was cleared by centrifugation (18,000 rpm, SS-34) and a subsequent filtration step (Millex-GV/0.22 μm, Millipore), before it was passed over the His-bind matrix-column under denaturing conditions. Denatured dAP-4 fusion protein eluted from this column was further purified by SDS-PAGE prior to immunization.

Antibodies

Production of polyclonal antibodies SA4114, directed towards antigen APDB, and SA2720, directed towards antigen APEX3, was carried out by Eurogentec. APDB was provided in its native form and APEX3 as a denatured protein purified by SDS-PAGE (see above). Before use in mobility shift assays, anti-dAP-4 and preimmune antibodies were purified on protein A-Sepharose as described (Lehmann & Korge, 1996).

Monoclonal anti-Da antibodies and an affinity-purified polyclonal anti-Da antiserum were a kind gift from Claire Cronmiller. The antigen used for production of both antibodies was a ß-gal fusion protein containing a 430 residue polypeptide from the N-terminal half of the Da protein that lacks the HLH domain. These antibodies showed no cross-reaction with proteins other than Da on Western blots (Cronmiller & Cummings, 1993).

Western blot and *in vitro* translation

Western analysis was carried out according to Harlow & Lane (1988) using anti-rabbit-immunoglobulin antibodies conjugated to alkaline phosphatase as secondary antibodies.

A cell-free *in vitro* transcription/translation system (TNT T7/T3 Coupled Reticulocyte Lysate System; Promega) and [^{35}S]methionine were used for synthesis of full-length dAP-4 by *in vitro* translation. Labelled protein was separated by SDS-PAGE and visualized by autoradiography.

Immunohistochemistry and immunofluorescence staining of polytene chromosomes

Detection of dAP-4 with polyclonal antisera (1:500) was based on the protocol for ß-galactosidase detection as described by Bier *et al*. (1989). As secondary antibody, a biotin-conjugated anti-rabbit antibody (Jackson) (1:200) was used and, after addition of ExtrAvidin (Sigma, 1:200), the staining reaction was carried out with HCl-diaminobenzidine and H_2O_2 in 50 mM Tris-HCl (pH 7.5).

Immunofluorescence staining of polytene chromosomes was performed as described (Lehmann & Korge, 1996).

Electrophoretic mobility shift assay

Binding reactions were as described (Lehmann & Korge, 1995), except that equilibration was on ice for 30 minutes, and a modified binding buffer was used (35-75 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, 5 mM DTT, 10 % glycerol, 0.4 % NP-40, pH 7.5 at 4 °C). The elevated concentration of DTT and NP-40 proved to be beneficial for binding of bacterially produced dAP-4. For reasons of consistency, the modified buffer was also used with binding reactions containing nuclear extract proteins. In supershift experiments, purified antibodies (1 μl) were added 30 minutes after mixing of all other components, and incubation was continued for another 30 minutes.

The double-stranded oligonucleotide O5N, which covers the *Sgs-4* upstream region from −436 to −406 and thus contains the strong SEBP3 binding site, was used as ^{32}P-labelled probe in all mobility shifts. Radiolabelling was performed by filling-in single-stranded 5′-ACGT overhangs on both ends of the oligonucleotide using Klenow fragment. The following double-stranded oligonucleotides were used for competition experiments: O6N, differs by two point mutations in positions −423 (G for A) and −422 (T for C) from O5N. O5ME is the same as O5N, except that it is slightly shorter (−432 to −405), and the central 5′-GC of the E-box (−420/ − 419) is reversed to 5′-CG. The weak SEBP3 binding site is represented by the Org-344 oligonucleotide, which corresponds to positions −359 to −329 of the *Sgs-4* upstream region. A mutated version, Sam-344, is identical with Org-344, except for a single C to T transition at −344. The O1 oligonucleotide (5′-AGAGATTTCACAGATG GTCGT-3′) is derived from the upstream region (−1000 to −980) of the *Sgs-1* gene, which has been cloned recently (Roth *et al*., 1999), and contains an AP-4 binding motif (underlined) that slightly deviates from the consensus sequence.

Data base searches and sequence alignments

The program MatInspector 3.0 (Quandt *et al*., 1995), which is based on the TRANSFAC data base (Heinemeyer *et al*., 1998; Wingender *et al*., 1997), was used to search for transcription factor binding sites related to the SEBP3 sites of *Sgs-4*. Database searches with human AP-4 as a DNA or protein query sequence were performed with BLAST at EMBL (Altschul *et al*., 1990). These searches yielded a highly significant match to P1 clone DS02740 (accession number L49408), sequenced by the Berkeley *Drosophila* Genome Project, as well as a match to cosmid clone F58A4, sequenced by the *C. elegans* genome project (Wilson *et al*., 1994). Coding sequences were assembled with the programs

GeneMark at EBI (Borodovsky & McIninch, 1993) and FGENED at BCM (Solovyev *et al*., 1994). Multiple alignments were carried out with PIMA at BCM (Ladunga *et al*., 1996) and SeqVu 1.1.

Sequence accession numbers

The nucleotide sequence of a 3.9 kb cDNA encompassing the complete coding region of *dAP-4* has been deposited in the GenBank™ database under accession number AF158371. The amino acid sequence of *Caenorhabditis* AP-4 can be accessed through the Swiss Protein database under Swiss-Prot number P34474, and the amino acid sequence of human AP-4 through the GenBank™ database under accession number NP_003214.

Acknowledgements

We are grateful to Claire Cronmiller for generously providing monoclonal and affinity-purified polyclonal anti-Da antibodies. We gratefully acknowledge the assistence of Ruth Brockmann and Madeleine Brünner in performing the polytene chromosome stainings, and of Karla Dünnbier in preparing nuclear extracts from salivary glands. We thank Günther Roth for critical reading of the manuscript.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft to G.K. (SFB 344) and M.L. (LE 870/2-2).

References

Alifragis, P., Poortinga, G., Parkhurst, S. M. & Delidakis, C. (1997). A network of interacting transcriptional regulators involved in *Drosophila* neural fate specification revealed by the yeast two-hybrid system. *Proc. Natl Acad. Sci. USA,* **94**, 13099-13104.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

Andrews, N. C., Kotkow, K. J., Ney, P. A., Erdjument-Bromage, H., Tempst, P. & Orkin, S. H. (1993). The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the *v-maf* oncogene. *Proc. Natl Acad. Sci. USA,* **90**, 11488-11492.

Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Carretto, R., Uemura, T. & Grell, E., *et al*. (1989). Searching for pattern and mutation in the *Drosophila* genome with a P- lacZ vector. *Genes Dev.* **3**, 1273-1287.

Blackwell, T. K., Huang, J., Ma, A., Kretzner, L., Alt, F. W., Eisenman, R. N. & Weintraub, H. (1993). Binding of myc proteins to canonical and noncanonical DNA sequences. *Mol. Cell. Biol.* **13**, 5216-5224.

Borodovsky, M. & McIninch, J. D. (1993). GeneMark: parallel gene recognition for both DNA strands. *Comput. Chem.* **17**, 123-133.

Cabrera, C. V. & Alonso, M. C. (1991). Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* **10**, 2965-2973.

Caudy, M., Grell, E. H., Dambly-Chaudiere, C., Ghysen, A., Jan, L. Y. & Jan, Y. N. (1988a). The maternal sex determination gene *daughterless* has zygotic activity necessary for the formation of peripheral neurons in *Drosophila*. *Genes Dev.* **2**, 843-852.

Caudy, M., Vassin, H., Brand, M., Tuma, R., Jan, L. Y. & Jan, Y. N. (1988b). *daughterless*, a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to *myc* and the *achaete-scute* complex. *Cell,* **55**, 1061-1067.

Chan, J. Y., Han, X. L. & Kan, Y. W. (1993). Isolation of cDNA encoding the human NF-E2 protein. *Proc. Natl Acad. Sci. USA,* **90**, 11366-11370.

Church, G. M., Ephrussi, A., Gilbert, W. & Tonegawa, S. (1985). Cell-type-specific contacts to immunoglobulin enhancers in nuclei. *Nature,* **313**, 798-801.

Cline, T. W. (1983). The interaction between *daughterless* and *sex-lethal* in triploids: a lethal sex-transforming maternal effect linking sex determination and dosage compensation in *Drosophila melanogaster*. *Dev. Biol.* **95**, 260-274.

Cline, T. W. (1989). The affairs of *daughterless* and the promiscuity of developmental regulators. *Cell,* **59**, 231-234.

Comb, M., Mermod, N., Hyman, S. E., Pearlberg, J., Ross, M. E. & Goodman, H. M. (1988). Proteins bound at adjacent DNA elements act synergistically to regulate human *proenkephalin* cAMP inducible transcription. *EMBO J.* **7**, 3793-3805.

Cronmiller, C. & Cummings, C. A. (1993). The *daughterless* gene product in *Drosophila* is a nuclear protein that is broadly expressed throughout the organism during development. *Mech. Dev.* **42**, 159-169.

Cummings, C. A. & Cronmiller, C. (1994). The *daughterless* gene functions together with *Notch* and *Delta* in the control of ovarian follicle development in *Drosophila*. *Development,* **120**, 381-394.

Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. (1985). B lineage-specific interactions of an immunoglobulin enhancer with cellular factors *in vivo*. *Science,* **227**, 134-140.

Grandori, C. & Eisenman, R. N. (1997). Myc target genes. *Trends Biochem. Sci.* **22**, 177-181.

Grandori, C., Mac, J., Siebelt, F., Ayer, D. E. & Eisenman, R. N. (1996). Myc-Max heterodimers activate a DEAD box gene and interact with multiple E box-related sites in vivo. *EMBO J.* **15**, 4344-4357.

Harlow, E. & Lane, D. (1988). *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L. & Kolchanov, N. A. (1998). Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucl. Acids Res.* **26**, 362-367.

Hoch, M. & Jäckle, H. (1998). *Krüppel* acts as a developmental switch gene that mediates *Notch* signalling-dependent tip cell differentiation in the excretory organs of *Drosophila*. *EMBO J.* **17**, 5766-5775.

Hofmann, A. & Lehmann, M. (1998). The transcriptional switch between the *Drosophila* genes *Pig-1* and *Sgs-4* depends on a SEBP1 binding site within a shared enhancer region. *Mol. Gen. Genet.* **259**, 656-663.

Hofmann, A., Keinhorst, A., Krumm, A. & Korge, G. (1987). Regulatory sequences of the *Sgs-4* gene of *Drosophila melanogaster* analysed by P element-mediated transformation. *Chromosoma,* **96**, 8-17.

Hu, Y. F., Luscher, B., Admon, A., Mermod, N. & Tjian, R. (1990). Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. *Genes Dev.* **4**, 1741-1752.

Jürgens, G. & Weigel, D. (1988). Terminal versus segmental development in the *Drosophila* embryo: the role of the homeotic gene *fork head*. *Roux's Arch. Dev. Biol.* **197**, 345-354.

Krumm, A., Roth, G. E. & Korge, G. (1985). Transformation of salivary gland secretion protein gene *Sgs-4* in *Drosophila*: stage- and tissue-specific regulation, dosage compensation, and position effect. *Proc. Natl Acad. Sci. USA,* **82**, 5055-5059.

Kuzin, B., Tillib, S., Sedkov, Y., Mizrokhi, L. & Mazo, A. (1994). The *Drosophila trithorax* gene encodes a chromosomal protein and directly regulates the region-specific homeotic gene *fork head*. *Genes Dev.* **8**, 2478-2490.

Ladunga, I., Wiese, B. A. & Smith, R. F. (1996). FASTA-SWAP and FASTA-PAT: pattern database searches using combinations of aligned amino acids, and a novel scoring theory. *J. Mol. Biol.* **259**, 840-854.

Lehmann, M. (1996). *Drosophila Sgs* genes: stage and tissue specificity of hormone responsiveness. *Bioessays,* **18**, 47-54.

Lehmann, M. & Korge, G. (1995). Ecdysone regulation of the *Drosophila Sgs-4* gene is mediated by the synergistic action of ecdysone receptor and SEBP 3. *EMBO J.* **14**, 716-726.

Lehmann, M. & Korge, G. (1996). The *fork head* product directly specifies the tissue-specific hormone responsiveness of the *Drosophila Sgs-4* gene. *EMBO J.* **15**, 4825-4834.

Littlewood, T. D. & Evan, G. I. (1994). Transcription factors 2: helix-loop-helix. *Protein Profile,* **1**, 635-709.

Mermod, N., Williams, T. J. & Tjian, R. (1988). Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription *in vitro*. *Nature,* **332**, 557-561.

Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D. & Lassar, A. B., *et al*. (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell,* **58**, 537-544.

Murre, C., Bain, G., van Dijk, M. A., Engel, I., Furnari, B. A., Massari, M. E., Matthews, J. R., Quong, M. W., Rivera, R. R. & Stuiver, M. H. (1994). Structure and function of helix-loop-helix proteins. *Biochim. Biophys. Acta,* **1218**, 129-135.

Nelson, C., Shen, L. P., Meister, A., Fodor, E. & Rutter, W. J. (1990). Pan: a transcriptional regulator that binds chymotrypsin, insulin, and AP-4 enhancer motifs. *Genes Dev.* **4**, 1035-1043.

Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucl. Acids Res.* **23**, 4878-4884.

Roth, G. E., Wattler, S., Bornschein, H., Lehmann, M. & Korge, G. (1999). Structure and regulation of the salivary gland secretion protein gene *Sgs-1* of *Drosophila melanogaster*. *Genetics,* in the press.

Sirito, M., Lin, Q., Maity, T. & Sawadogo, M. (1994). Ubiquitous expression of the 43- and 44-kDa forms of transcription factor USF in mammalian cells. *Nucl. Acids Res.* **22**, 427-433.

Solovyev, V. V., Salamov, A. A. & Lawrence, C. B. (1994). Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Nucl. Acids Res.* **22**, 5156-5163.

Sun, X. H., Copeland, N. G., Jenkins, N. A. & Baltimore, D. (1991). Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol. Cell. Biol.* **11**, 5603-5611.

Unk, I., Kiss-Toth, E. & Boros, I. (1994). Transcription factor AP-4 participates in activation of bovine leukemia virus long terminal repeat by p34 *Tax*. *Nucl. Acids Res.* **22**, 4872-4875.

Vaessin, H., Brand, M., Jan, L. Y. & Jan, Y. N. (1994). *daughterless* is essential for neuronal precursor differentiation but not for initiation of neuronal precursor formation in *Drosophila* embryo. *Development,* **120**, 935-945.

Van Doren, M., Ellis, H. M. & Posakony, J. W. (1991). The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development,* **113**, 245-255.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T. & Cooper, J., *et al*. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature,* **368**, 32-38.

Wingender, E., Dietze, P., Karas, H. & Knuppel, R. (1996). TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucl. Acids Res.* **24**, 238-241.

Wingender, E., Kel, A. E., Kel, O. V., Karas, H., Heinemeyer, T., Dietze, P., Knuppel, R., Romaschenko, A. G. & Kolchanov, N. A. (1997). TRANSFAC, TRRD and COMPEL: towards a federated database system on transcriptional regulation. *Nucl. Acids Res.* **25**, 265-268.

Edited by M. Yaniv

(Received 19 March 1999; received in revised form 21 June 1999; accepted 25 June 1999)