INTRODUCTION

Organs and other animal body parts often develop from sets of cells that are determined as a group to form a particular structure. A class of genes, termed field-specific selector genes, has been identified that determine the fates of entire fields of cells and direct the development of whole organs and body structures (reviewed by Carroll et al., 2001; Mann and Morata, 2000). In Drosophila, some field-specific selector genes and the structures whose formation they regulate include Pax6/eyeless (ey) in the eye (Halder et al., 1995; Quiring et al., 1994), tinman (tin) in the heart (Azpiazu and Frasch, 1993; Bodmer, 1993; Frasch, 1999), Distal-less (Dll) in the limbs (Cohen et al., 1989; Gorfinikel et al., 1997), and vestigial/scalloped (vg/sd) in the wing (Halder et al., 1998; Kim et al., 1996; Simmonds et al., 1998; Williams et al., 1991). A second class of selector genes, the Hox genes, act in specific domains along the anteroposterior body axis and in developing appendages to specify their identity but not their formation (Carroll et al., 1995; Lawrence and Morata, 1994; Lewis, 1978; Manak and Scott, 1994; McGinnis and Krumlauf, 1992).

The known selector proteins are transcription factors that exert their prominent effects by regulating presumably large but specific sets of target genes. However, the DNA-binding domains of selector proteins often show promiscuous DNA-binding specificity in vitro. For example, recognition sequences for the homeodomains of Hox proteins are typically only 6 bp long (Biggin and McGinnis, 1997; Ekker et al., 1994; Gehring et al., 1994; Mann, 1995; Mann and Affolter, 1998). Similarly, the consensus sequence bound by the TEA domain of the Scalloped (Sd) protein is 8 bp long but degenerate (reviewed by Jacquemin and Davidson, 1997). Potential binding sites for these proteins are predicted to occur once every 2-4 kb in a random sequence, and therefore may be found in cis-regulatory regions of virtually every gene. However, selector proteins presumably do not regulate all genes in a genome. Furthermore, the activity of many selector proteins, particularly Hox proteins, is not restricted to a single field, but may be required during the development of several structures to regulate distinct sets of target genes (Azpiazu and Frasch, 1993; Bodmer, 1993; Halder et al., 1995; Morata and Sanchez-Herrero, 1999) reviewed in (Mann and Morata, 2000).

Understanding how the target selectivity of selector proteins is determined in vivo is thus fundamental to understanding how they control gene expression and pattern formation.

INTERACTIONS WITH SPECIFIC CO-FACTORS MAY BE A MAJOR DETERMINANT OF SELECTOR PROTEIN TARGET SELECTIVITY. THE DNA-BINDING SPECIFICITY OF TRANSCRIPTION FACTORS IS OFTEN INCREASED BY COOPERATIVE INTERACTIONS WITH SPECIFIC CO-FACTORS THAT ARE ALSO DNA-BINDING PROTEINS. THE HOX PROTEINS AND THEIR PBC (Pbx, ceh-20, Extradenticle (Exd)) AND MEIS (Homothorax (Hth), Meis, Prep) CO-FACTORS PROVIDE A PROMINENT EXAMPLE (MANN,

SUMMARY

The formation and identity of organs and appendages are regulated by specific selector genes that encode transcription factors that regulate potentially large sets of target genes. The DNA-binding domains of selector proteins often exhibit relatively low DNA-binding specificity in vitro. It is not understood how the target selectivity of most selector proteins is determined in vivo. The Scalloped selector protein controls wing development in Drosophila by regulating the expression of numerous target genes and forming a complex with the Vestigial protein. We show that binding of Vestigial to Scalloped switches the DNA-binding selectivity of Scalloped. Two conserved domains of the Vestigial protein that are not required for Scalloped binding in solution are required for the formation of the heterotetrameric Vestigial-Scalloped complex on DNA. We suggest that Vestigial affects the conformation of Scalloped to create a wing cell-specific DNA-binding selectivity. The modification of selector protein DNA-binding specificity by co-factors appears to be a general mechanism for regulating their target selectivity in vivo.

Key words: Selector gene, Co-factor, Transcription, Wing formation network, Drosophila
so that Sd and the Vg-Sd complex bind to different sets of binding sites. The presence or absence of Vg in a particular cell is therefore a key determinant of the set of cis-regulatory elements, which are bound and regulated by Sd. The tissue-specific modification of selector protein DNA-binding specificity by co-factors may be a general mechanism for increasing their target selectivity in vivo.

MATERIALS AND METHODS

Protein production

Full-length 35S-labeled Vg and Sd proteins were produced using the T7 in vitro transcription and translation system (Promega). The T7plink expression vector contains the T7 promoter, the 5' untranslated region of the T7 leader gene fused to a Kozak consensus ATG followed by multiple cloning sites (Dalton and Treisman, 1992). Vg- and Sd-coding regions were PCR amplified and subcloned into T7plink. The Sd protein contained an extra six amino acids at the N terminus (MAGSEF) encoded by the T7plink vector. Sdmyc contained an N-terminal Myc tag (MEQKLISEEDLNNM) fused to the Sd ORF. All Vg proteins containing the Vg N-terminus started with the Methionine of Vg itself and did not contain any extra residues. Vg proteins that had N-terminal deletions started with the vector encoded peptide MAGSEF fused to Vg. The breakpoints, indicated by an asterisk, of the Vg deletions are: ΔSID, 5' breakpoint is DTAASQ*. 3' breakpoint is *NYVHP; position 66, *SVSAN; 5'SID, *QAYQY; 3'SID, PIPAP*; position 73, NAAA*; position 176, *THQTK; and position 274, *GSGQQQ. The HA tag (SMAYPYDPVYASLG) was inserted at position 153 between S and H. The Sd TEA domain was purified as a (His)6-tagged protein on nickel chelate columns (Novagen), as described earlier (Halder et al., 1998; Jaquequin et al., 1996).

DNA probes

DNA probes for EMSAs were labeled with 32P-αATP by fill in reaction of double T overlaps at both ends of double-stranded oligonucleotides using the Klenow fragment of DNA polymerase I. Single strand oligonucleotides were annealed at concentrations of 10 μM in 10 mM Tris pH 7.5 and 50 mM NaCl. Labeled probes were purified over Sephadex G50 columns (Princeton Separations). Different probes were diluted to the same specific activities with cold labeled oligonucleotides. Sequences of the upper strand oligonucleotides were 5’ to 3’: 2xGT (TTCCATACACGAGCAGGAAATGGAGTTACGCCC), 1xGT (TTCCATACACGAGCAGGAAATGGAGTTACGCCC), GTspaceGT (TTCCATACACGAGCAGGAAATGGAGTTACGCCC), cut-564 (TTGTCAGATAATCTGGGAGGAGTCTGTGTGTGTGTGATTGT), cut-341 (TTGTCAGATAATCTGGGAGGAGTCTGTGTGTGTGTGATTGT), 1xGTspaceGT (TTGTCAGATAATCTGGGAGGAGTCTGTGTGTGTGTGATTGT), sal-862 (TTGTCAGATAATCTGGGAGGAGTCTGTGTGTGTGTGATTGT); kni-268 (K. Guss and E. Bier, personal communication; TTTCCCT-CTTACATTTTTCATTTTTCTTTTTCTTTTGCAAGCAGTCTGTGTGTGTGTGATTGT), DSRRF (CCGTTACACAAATGCCAGAATCTGTGTGTGTGTGATTGT); cTTN (TTCCAGAGAATGGCAACACTTTGTGTGTGTGTGATTGT); and αMHC (TTGTCAGATAATCTGGGAGGAGTCTGTGTGTGTGTGATTGT).

Electrophoretic mobility shift assays (EMSA)

EMSA reactions with TNT produced proteins were carried out in 20 μl binding buffer (8% glycerol, 15 mM Heps pH 7.9, 150 mM KCl, 1 mM EDTA, 100 μg/ml bovine serum albumin (BSA)) containing 0.7 μl TNT reaction, 0.3 μg dDNA and 3 fmol DNA probe. Equimolar amounts of 35S-labeled proteins were added by diluting the TNT reactions accordingly with unpurified TNT extract. Binding reactions were incubated for 15 minutes at room temperature and complexes were separated on 5% polyacrylamide gels and standard proteins.
sites with high affinity, and with similar affinity to single, domain of Sd bound cooperatively to all of these doublet sites (Halder et al., 1998). Despite substantial differences in sequence, the TEA domain used are indicated above gels.

uncomplexed Sd. None of the probes are bound by Vg in the absence of Sd. EMSAs with purified TEA domain protein were carried out in essentially the same way, except that 1 fmol probe was used and the buffer contained 100 mM KCl, 2 mM MgCl₂ and no dIdC. TEA domain shifts were run on 6% polyacrylamide gels. For supershifts, 100 ng antibody was added to the binding reaction.

Co-immunoprecipitations
To precipitate, 10 μl of TNT product were incubated with 400 μl IP buffer (15 mM Hepes pH 7.9, 150 mM KCl, 1 mM EDTA, 1% Triton) and 20 μl of protein A-Sepharose suspension (Amersham Pharmacia Biotech) at 4°C for 20 minutes shaking. Reactions were centrifuged for 2 minutes and 1 μg of antibody (mAB α-Myc, mAB α-HA, both from Babco) was added to the supernatant, which was then incubated on a shaker at 4°C for 60 minutes. Agarose beads were pelleted by centrifugation at 1500 g for 2 minutes. Supernatant was removed and beads were washed four times with 700 μl IP buffer. Bound proteins were eluted and denatured in 40 μl SDS sample buffer (with 200 mM DTT) by incubation at 68°C for 15 minutes. Proteins were separated by standard 12% and 18% SDS-PAGE. Gels were dried and exposed to BiomaxMR film at 70°C -70°C - FlyBase) genes (Guss et al., 2001; Halder et al., 1998). These sites were identified by DNaseI footprinting using the TEA domain of Sd. In these analyses, we were struck by the finding that essential sites occurred most often as tandem double sites, for example, in the cut, spalt and DSRF (bx – FlyBase) genes (Guss et al., 2001; Halder et al., 1998). Despite substantial differences in sequence, the TEA domain of Sd bound cooperatively to all of these doublet sites with high affinity, and with similar affinity to single, nonessential sites and to native single vertebrate TEF-1-binding sites in muscle-specific cis-regulatory elements and the SV40 enhancer (Guss et al., 2001; Halder et al., 1998; Table 1). From these studies, we have inferred a consensus binding site sequence of T⁴A₆G₆G₆T₃A AT G₃T T for the TEA domain of Sd, which is very similar to that of the TEA domain of TEF-1 (Guss et al., 2001; Jacquemin and Davidson, 1997; Jiang et al., 2000).

In contrast to the isolated TEA domain, however, the full-length Sd protein (produced by in vitro translation, see Materials and Methods) did not bind equivalently to all of these sites but rather showed a restricted DNA-binding specificity (Fig. 1; summarized in Table 1). Full-length Sd bound specifically to the doublet site in the DSRF enhancer and to most of the single binding sites (Fig. 1, lanes 14, 18; Table 1), but binding to the cut, sal, kni and other native templates with doublet sites was weak or nearly undetectable (Fig. 1, lanes 2, 6, 10; Table 1). The difference in DNA-binding activity between the TEA domain and Sd protein indicates that there are motifs within the native Sd protein that affect the activity of the TEA domain and restrict its binding to certain sites. We refer to sites that are bound by Sd as A-sites.

The finding that most of the doublet-binding sites were not bound by the full-length Sd protein was surprising, considering that these templates were bound with high affinity by the TEA domain and that these sites are essential for enhancer activity in vivo (Guss et al., 2001). The observations that the activity of these cis-regulatory elements in vivo and in cell culture depends on co-expression of Vg with Sd (Halder et al., 1998), and the finding that Vg and Sd interact physically (Paumard-Rigal et al., 1998; Simmonds et al., 1998), raised the possibility that interaction of Vg with Sd changed its DNA-binding properties and enabled binding to these sites. However, previous Vg-Sd protein interaction studies have been performed in the absence of DNA and the possible effect of the interaction between Vg and Sd on DNA-binding has thus not been addressed. We have tested whether Sd and Vg form a complex on DNA in vivo and whether this complex has different DNA-binding properties from the Sd protein alone.

Co-translation of Sd with Vg produced a Vg-Sd complex
that bound to these other sites (referred to as B-sites). In contrast to Sd alone, complexes containing Sd and Vg bound strongly to the cut, sal and kni elements (Fig. 1, lanes 3, 7, 11). Quantification of the bound complexes showed that Vg increased Sd binding to these doublet sites by about 10-fold. In addition to enabling binding to B-sites, interaction with Vg reduced Sd binding to the single site templates by at least fivefold (Fig. 1). Importantly, we have not observed binding of Vg alone to any of the binding sites described in this report or to any other DNA templates tested (Fig. 1, data not shown). Therefore, Vg binding to Sd switches the DNA target preference of Sd from the single A-sites to the doublet B-sites.

Only two of the eleven templates that we tested were bound by both Sd and Vg-Sd and thus possessed A- and B-site properties. These were the synthetic 2xGT (see below) and the DSRF probes (data not shown). Thus, while most native templates have either A- or B-site character, sites with both A- and B-site properties also occur.

### Table 1. Sequences bound by the TEA domain, Sd and the Vg-Sd complex

<table>
<thead>
<tr>
<th>Site</th>
<th>TEA</th>
<th>Sd</th>
<th>Sd-Vg</th>
</tr>
</thead>
<tbody>
<tr>
<td>cut-564</td>
<td>TCAATGTAATCCGAAATATGGTC</td>
<td>+++</td>
<td>+/-</td>
</tr>
<tr>
<td>cut-341</td>
<td>CAGAATTTATCTGGAAAATTATT</td>
<td>++++</td>
<td>+</td>
</tr>
<tr>
<td>sal-750</td>
<td>TTTCTGAAATCCGAAATATCCCAT</td>
<td>++++</td>
<td>+</td>
</tr>
<tr>
<td>kni-268</td>
<td>CTTCTATACATGCTGTACATCC</td>
<td>++++</td>
<td>+</td>
</tr>
<tr>
<td>sal-862</td>
<td>CATAACTTATTTAAAAAA</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>cTNT</td>
<td>AGAGGAGTTGCAAACA</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>aMHC</td>
<td>CACGGATGAGCTA</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>1xGT</td>
<td>TTGGAGATGTTT</td>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>2xGT</td>
<td>TTGGATGGATGTTA</td>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>GTspaceGT</td>
<td>TTGGATGGATGTTA</td>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>DSRF</td>
<td>TTTAACTATGGCCGAAATTTCTA</td>
<td>++++</td>
<td>++++</td>
</tr>
</tbody>
</table>

The shaded boxes and arrows indicate the Sd binding sites, as inferred from the nucleotides required for TEF-1 binding to cTNT - defined by scanning mutagenesis (Butler and Ordahl, 1999). The black bar above the sequences identifies the region in the DSRF cis-regulatory element protected by the Sd TEA domain from DNaseI digestion (Halder et al., 1998). Circles indicate G-residues in cTNT that, when methylated, interfere with TEF-1 binding (Larkin et al., 1996). Relative binding affinities, as determined by EMSA (Fig. 1) and densitometry, are shown on the right. Sequences are from Guus et al. (2001) (cut, sal, kni), Cooper and Ordahl (1985) (cTNT), Molkentin and Markham (1994) (aMHC), Davidson et al. (1988) (GTIIc, from which 1xGT, 2xGT and GTspaceGT are derived) and Halder et al. (1998) (DSRF).
only a single Sd binding site (1×GT, resembling the native GTⅠ-C-binding site, Table 1) or that had a 10 bp spacer between the two Sd-binding sites (GTspaceGT) that abolishes cooperative binding (Fig. 2; Davidson et al., 1988; Jiang et al., 2000).

Full-length Sd bound to the 2×GT site as a monomer and as a dimer (Sd and Sd2; Fig. 2A, lane 2). Incubation of a co-translated mixture of Sd and Vg produced an additional complex that migrated more slowly (Vg,Sd; Fig. 2A, lane 3), while expression of Vg alone again did not result in any detectable DNA-binding activity (Fig. 2A, lanes 4, 8). Two molecules of the TEA domain bound to the 2×GT probe cooperatively (Fig. 2B, top). Upon titration of the TEA domain, only a small fraction of complexes containing a single TEA molecule was observed, while the shift to two TEA molecules bound occurred abruptly between 0.3 ng and 1 ng of TEA domain added, indicating cooperative binding. However, the GTspaceGT probe was bound by two TEA domains, but in a non-cooperative fashion (Fig. 2B, bottom). At lower TEA domain concentrations, only one TEA molecule bound and with increasing concentration the second site was gradually occupied. However, 2×GT and GTspaceGT bound with similar affinity, as the TEA concentrations required to shift them were approximately the same. Cooperative binding of full-length Sd on GTspaceGT was also reduced, as no complexes containing two Sd molecules were observed, in contrast to Sd binding to the 2×GT template (Fig. 2A, lanes 2, 6). Despite this reduction in Sd cooperativity, a Vg-Sd complex formed on the GTspaceGT probe as efficiently as on 2×GT (Fig. 2A, lanes 5, 7). Therefore, cooperative binding of Sd to DNA is not required for Vg-Sd complex formation on DNA. Importantly, however, two Sd-binding sites are required, as Vg-Sd complexes form on the GTspaceGT template but barely on 1×GT.

The Sd-interaction domain of Vg is sufficient for binding to Sd but not for complex formation on DNA

To identify domains within Vg that may be important for Sd interaction and complex formation on DNA, we first searched
for conserved domains in Vg homologs and other proteins. A Vg homolog from the mosquito (Jim Williams and S. B. C., unpublished) shows strong conservation of the first 79 amino acids and of the region beginning with the previously identified Sd-interaction domain (SID) to the very C terminus (Fig. 3A,B). The region from position 80 to 280 in the *Drosophila* protein shows no or only moderate similarity to mosquito Vg. A Vg homolog from vertebrates, Tondu (Tdu; Vaudin et al., 1999), shares the first part of the SID but no other domains (Fig. 3A,B). We have identified a novel vertebrate Vg homolog.

![Image](https://example.com/image1.png)

**Fig. 4.** Vg protein domains outside the Sd interaction domain are required for Vg-Sd complex formation on DNA but not for Vg-Sd interaction in solution. (A) EMSA of the binding to the *cut* template by the Vg deletion mutant proteins co-expressed with Sd. Only the two internal deletions Δ(73-176) and Δ(73-274) are as efficient in DNA complex formation as full-length Vg (lanes 3,10,11). The deletions 66-C, SID-C and N-SID showed partial activity, while all other deletions were not active. Lane 1, TNT extract; lane 2, Sd only; lanes 3-13 Sd co-expressed with the indicated Vg mutants. Labeling as in Fig. 1. (B) SDS-PAGE of Co-IPs of Sd and Vg mutants. The 35S-labeled proteins produced by TNT (T) and the precipitated proteins (IP) are shown next to each other for each Vg mutant. Lanes 1 and 2: anti-Myc antibody does not precipitate untagged Vg or Sd. Lanes 3-13 Myc-tagged Sd co-expressed with the indicated Vg mutants. The top band is Sd, lower bands are Vg mutant proteins.

![Image](https://example.com/image2.png)

**Fig. 5.** Vg and Sd form multimeric complexes on DNA but heterodimers in solution. (A) Lanes 1-3: EMSA with VgΔHA and/or VgΔ(73-274) co-expressed with Sd binding to the *cut* template. Black arrowhead indicates complexes with intermediate mobility, presumably comprising the two Vg forms, Sd and DNA. Open arrowhead indicates complexes with the same mobility as those observed by co-expressing full-length Vg and Sd only. Lanes 4-6: same reactions incubated with anti-HA antibody. Complexes containing full-length Vg (lanes 4,5) as well as the intermediate complexes (lane 5, arrowhead), but not the untagged VgΔ(73-176) (lane 6), are supershifted. VgΔ(73-176) gave the same effect (not shown). (B) SDS-PAGE of Co-IPs of Sd co-expressed with VgΔHA and/or VgΔ(73-176). Precipitation of VgΔHA co-precipitated Sd but not the other Vg form demonstrating that Vg and Sd form complexes in solution that do not contain more than one Vg molecule. The anti-HA antibody did not precipitate the untagged Vg deletion (lanes 5,6). The VgΔ(73-274) protein gave the same result (not shown).
mediate the formation of a complex with Sd on the DNA and were not able to substitute for full-length Vg and to interact with Sd in solution. (Fig. 4A, lane 22). This demonstrates that the SID is necessary and required to correctly position the green domain, so that it can interact with C-terminal residues to form the Vg-Sd-DNA complex.

Using these domain boundaries as guidelines, we constructed a series of Vg protein deletion mutants and tested them for interaction with Sd in solution (Fig. 4B). We assayed for Sd-Vg interaction in solution by co-immunoprecipitation (Co-IP). We co-expressed 35 S labeled Myc-tagged Sd with the Vg deletion mutants and immunoprecipitated Sd using an anti-Myc antibody (Fig. 4B). The same buffer conditions and protein concentrations were used for the EMSA and Co-IP assays. We found that only the SID is required for interaction with Sd in solution because all Vg deletion mutants, except for the deletion in SID itself, are still co-immunoprecipitated with Sd (Fig. 4B). In fact, the SID by itself was able to bind to Sd (Fig. 4B, lane 22). This demonstrates that the SID is necessary and sufficient to mediate Vg-Sd interaction in solution.

In contrast to Sd binding in solution, the SID by itself, however, was not able to substitute for full-length Vg and to mediate the formation of a complex with Sd on the DNA template (Fig. 3C). In fact, Sd interaction inhibited Sd from binding to A-sites (not shown). Therefore, domains in addition to the SID are required for complex formation on DNA. Indeed, deletion of the green domain or of the N-terminus up to the SID reduced DNA complex formation, indicating that the green domain is important for this activity (Fig. 4A, lanes 5, 6). Similarly, deletion of the region C-terminal to the SID reduced complex formation on DNA (Fig. 4A, lane 7). Deletion of the green domain or of the C-terminal deletion completely abolished complex formation (Fig. 4A, lanes 8, 9). These results indicate that the green domain and C-terminal residues act redundantly and/or cooperate in complex-forming activity.

We also created four deletion mutants to test whether the non-conserved residues between the green domain and the SID have essential functions. Internal deletions of residues 73-176 and 73-274 had essentially no effect on Vg activity and mediated the formation of a complex with Sd on DNA in Fig. 3B, arrowheads). We now define the extent of the SID by the region that is conserved between insect Vg and Fdu (Fig. 3B). In addition to the SID, Fdu and Vg share two other domains: the N-terminal 66 amino acids (green domain in Fig. 3A,B) and a domain rich in histidine and alanine residues C-terminal to the SID (orange domain in Fig. 3A,B). Therefore, the newly identified Fdu protein is more similar to Vg than is the Tdu protein.

The design of this experiment was to test for the formation of Vg-Sd complex on DNA. All three domains show conservation between insect Vg and vertebrate Fdu.

Vg and Sd form a heterodimer in solution but a heterotetramer on target DNA

The observation that the Vg-Sd complex did not bind to the 1xGT probe was surprising given that it bound to the 2xGT probe, which contains two tandemly arranged 1xGT binding sites. This raises the question of why the Vg-Sd complex requires doublet sites but does not recognize single binding sites. One possibility is that DNA binding by the Vg-Sd complex requires more than one Sd and Vg molecule in the complex. For example, a single molecule of Vg could bridge two Sd molecules to enable DNA binding or it may be that Vg-Sd dimers interact to increase the affinity or stability of the complex on DNA.

To determine how many Vg molecules are present in the respective complexes, we examined the mobility of complexes formed when two Vg molecules of different sizes were present. The design of this experiment was to test for the formation of the heterotetramer on target DNA.

---

**Fig. 6.** Vg-Sd complex formation on DNA does not require bases outside of the Sd-binding sites. (A) Top strands of the series of truncated templates. TT and AA bases at the ends of the probes were added in order to label the probes with 32P α-dATP. The extent of the Sd binding sites (shaded regions) was defined by systematic mutational analysis (Butler and Ordahl, 1999). (B) EMSA using TNT produced Vg and Sd and the probes shown in A. The 31 and 25 mers bind strongly to Sd and the Vg-Sd complex (lanes 1-7). The 21 mer is weakly bound by Sd (lane 8), while the Vg-Sd complex forms nearly as efficiently as on the longer templates. The 18 mer is not bound by Sd or Vg-Sd. Labeling is as is in Fig. 1.
heteromeric complexes of intermediate size that would indicate the presence of two (or more) Vg molecules in the complexes. We made use of two internal deletion mutants that showed nearly normal activity in the DNA-binding and Sd interaction assays (Fig. 4). The Vg deletion Δ(73-274) formed complexes with Sd on DNA that migrate faster on EMSA gels than complexes containing full-length Vg, owing to the smaller size of the mutant protein (Fig. 5A, lanes 1,3). Because we wanted to use the same proteins for the EMSA and the Co-IPs described below, we used an HA-tagged Vg protein (VgHA), which gave the same results as native Vg (not shown). When full length VgHA was co-expressed with VgΔ(73-274) and Sd, complexes of intermediate mobility formed (Fig. 5A, lane 2 arrowhead). Complexes migrating at the position of the complexes formed with VgΔ(73-274) were still present (open arrowhead). However, little (if any) complexes with the Vg deletion were observed, which may be due to competition by VgHA, which binds with higher affinity than Δ(73-274) (Fig. 5A, compare lane 1 with lane 3). We interpret the intermediate complexes as hetero-complexes formed between Sd, a mixture of VgHA and VgΔ(73-274), and DNA. Results using the VgΔ(73-176) deletion were identical (not shown). Thus, more than one Vg molecule is present in the shifted complexes. More precisely, as one extra complex of intermediate size appeared, it most probably contains one of each of the two Vg forms and is a heterotetramer.

We next examined whether or not this heterotetrameric complex formed in solution independently of DNA binding. We inserted an HA-tag into the middle of the non-conserved and dispensable region of Vg at a position predicted to be exposed to the surface. We first evaluated the activity of the VgHA mutant and a possible negative effect of anti-HA antibody binding on DNA complex formation in EMSA supershifts. Addition of anti-HA antibody supershifted complexes from TNT reactions producing VgHA. Importantly, the anti-HA antibody also supershifted the intermediate complex containing the full-length and shortened forms of Vg (Fig. 5A, lane 5 arrowhead), demonstrating that the anti-HA antibody does not disrupt the formation of heterotetrameric complexes on DNA. Immunoprecipitation of VgHA in the absence of DNA template co-precipitated Sd as efficiently as Vg was co-precipitated by SdΔREV (compare Fig. 5B, lanes 1 and 2 with Fig. 4, lanes 3 and 4). Significantly, immunoprecipitation of VgHA from a TNT reaction containing Sd, VgHA and one of the Vg deletions Δ(73-176) or Δ(73-274) co-precipitated Sd but not the other Vg protein species (Fig. 5B, lanes 3, 4). Thus, the Vg-Sd complex is a heterodimer in solution but a heterotetramer on DNA.

Vg-Sd complex formation on target DNA does not require bases outside the Sd-binding site

The observation that the Vg-Sd complex has increased affinity for B-sites compared with the Sd protein alone raises the possibility that Vg makes DNA contacts outside the region that is contacted by Sd, which could enlarge the DNA interaction surface and thereby increase the affinity of the complex. It has not been possible to produce sufficient quantities of active Vg-Sd complexes for chemical interference and DNaseI footprinting assays that could localize exactly the region contacted by the Vg-Sd complex. Bacterially produced Sd and Vg are insoluble and do not form active complexes upon renaturation, either refolded together or refolded separately and then mixed together. As an alternative to a chemical interference assay, we designed a series of 2×GT probes with increasingly truncated ends to test whether bases outside of the Sd binding sites are required for Vg-Sd complex formation (Fig. 6A). We chose the 2×GT probe because this template allows observation of Sd and Vg-Sd binding. Our rationale was that, if Vg contacts DNA outside the Sd-binding sites, then the Vg-Sd complex may not form on shorter probes that are nevertheless bound by Sd alone. However, if Vg does not contact DNA, the minimal template length requirement should be similar for Vg-Sd and for Sd binding.

We found that the Vg-Sd complex still formed on a probe that was shortened sufficiently to reduce Sd binding (Fig. 6B). The 31 and 25 mer probes bound to Sd and Vg-Sd with high affinity (Fig. 6, lanes 1-7). The shorter 21 mer had strongly reduced Sd binding (Fig. 6, lane 8), but Vg-Sd complex formation was nevertheless nearly as efficient on this template as on the longer probes (Fig. 6, lanes 3, 6, 9). Neither the Vg-Sd complex nor Sd alone bound to the shorter 18 mer. Thus, Vg interacts with Sd and increases its binding to a template that is too short for efficient Sd binding. We conclude that Vg-Sd complex formation does not require bases outside the region required by Sd, and thus that Vg does not contact DNA, at least not outside of the Sd-binding sites.

DISCUSSION

We have analyzed the DNA-binding properties of Vg-Sd complexes, the full-length Sd protein, and of the TEA domain of Sd with respect to a number of binding sites, particularly the functional sites identified in native cis-regulatory elements. We have made four key findings. First, we found that the Sd protein has a more restricted DNA-binding specificity than its isolated TEA domain. Second, we showed that the Vg-Sd complex binds well to sites in native cis-regulatory elements to which Sd alone does not bind well. Third, we found that two domains of the Vg protein are required for Vg-Sd complex formation on DNA that are not required for Vg binding to Sd in solution. And fourth, that this complex is a heterotetramer on DNA while apparently a heterodimer in solution. Below we present a mechanistic model for the control of Vg-Sd DNA target selectivity that considers these findings.

Vg binding switches the DNA target selectivity of Sd

We propose a model in which Vg binding to Sd switches the DNA target selectivity of Sd (Fig. 7). We found that the Sd protein alone binds to sites with a particular composition, termed A-sites, which exist singly or as doublets. In the latter case, Sd may bind cooperatively if the two sites are arranged in tandem. When Vg is also present, Vg and Sd interact and form a dimer in solution (Fig. 5B). This complex has two distinct properties. First, the Vg-Sd dimer has a greatly reduced affinity for A-sites (Fig. 1). Vg may either induce a conformational change in Sd that inhibits the TEA domain from interacting with DNA, or Vg could directly mask the TEA domain. Second, the dimer forms a higher order complex on a different set of binding sites, termed B-sites (Fig. 1). These two activities of Vg are distinguished by their structural requirements. While the SID domain of Vg is sufficient to
inhibit Sd DNA-binding to A-sites, additional domains N- and C-terminal to the SID are required for complex formation on B-sites (Fig. 4). Importantly, B-sites are poorly bound by Sd in the absence of Vg. Thus, Vg binding to Sd inhibits binding to A-sites while enabling binding to B-sites, that is, Vg switches the DNA-binding preference from A-sites to B-sites.

How does Vg binding affect the target selection of Sd? Two, not necessarily mutually exclusive models, may be postulated. First, Vg may influence Sd through global effects on Sd DNA binding. That is, Vg may act to reduce the DNA binding affinity of Sd to any target DNA, while also enhancing cooperativity of neighboring Vg-Sd complexes on DNA. We found that Vg and Sd form dimers in solution and that these dimers do not bind single A-sites. We have never observed any complexes of Sd and Vg on DNA migrating at a position, indicating Vg-Sd dimers bound to DNA on either A- or B-sites (the Vg-Sd complexes bound weakly to 1×GT have the same mobility as the ones bound to B-sites, and thus also consist of heterotetramers). However, in spite of the negative effect of Vg on DNA binding, two Vg-Sd dimers bound strongly to doublet B-sites. Apparently, strong cooperative interactions between two Vg-Sd dimers allow binding to B-sites. The N- and C-terminal protein domains of Vg that are required in addition to the SID for complex formation on DNA may be required for these interactions, which could involve Vg-Sd and/or Vg-Vg interactions between the two dimers on DNA.

Alternatively, Vg interaction may specifically enhance binding to doublet B-sites. We favor this model because we found that Vg-Sd had a similar affinity for several B-sites such as those in cut and 2×GT, even though 2×GT is a much better Sd binding site. The affinities of Sd for these sites therefore do not translate directly into the relative affinities observed for Sd-Vg binding, as would be expected if Vg only enhanced cooperativity. In addition, we found that the TEA domain binds several A- and B-sites with high affinity, but that full-length Sd has a strong preference for A-sites over B-sites. Thus, in the absence of any co-factor, Sd is in a conformation in which a domain of Sd separate from the TEA domain inhibits the TEA domain from binding to B-sites specifically (Fig. 1). In vitro, Vg interaction appears to be able to alleviate this inhibition because Vg-Sd complexes bind strongly to B-sites. This alleviation only occurs when complexes form on doublet sites, as Vg-Sd complexes do not bind to DNA as a dimers. We suggest that some sort of conformational change is associated with binding to doublet B-sites (Fig. 7). Our model is supported by the finding that the region of Sd that binds to the SID of Vg is homologous to a region of the vertebrate TEF-1 that negatively affects DNA binding (Hwang et al., 1993; Simmonds et al., 1998). This model is analogous in part to the role of Exd overcoming the inhibitory effect of the YKWM motif in the Labial Hox protein (Chan et al., 1996).

We have argued here that Sd and the Vg-Sd complex differentiate between A- and B-sites. What then are the distinguishing features of these sites? The sequences of the A- and B-sites are quite diverse and their alignment does not reveal different consensus sequence motifs. However, Sd clearly prefers binding to A-sites, and the inability of Sd to bind strongly to B-sites, such as that in the cut element, must therefore be due to the sequence of the template site. Vg-Sd complexes bind with high affinity to only two sites when arranged in tandem, and do not form on single A- or B-sites.

Thus, Sd discriminates between A- and B-sites based on sequence, while the binding of Vg-Sd complex depends both on sequence and the arrangement of the sites. We have identified two sites (DSRF and 2×GT) that have A- as well as B-site properties, so these properties are not mutually exclusive. However, many sites exist that are bound well by Sd or Vg-Sd, but not by both. Most of the essential sites for Vg-Sd regulation in vivo have mainly B-site character and are bound poorly by Sd. The identification of the exact sequence requirements that distinguish native essential Sd sites from the known Vg-Sd target sites will require some knowledge of Sd-regulated target genes in other tissues (see below).

**Vg as a determinant of the specificity of Sd action in vivo**

Vg binding and its effect on the DNA target selectivity of Sd plays a major role in distinguishing the biological specificity of Sd action in the developing wing from Sd function in other tissues. Sd is required for the development of tissues other than the wing, for example, the eye and the PNS, where it is not co-expressed with Vg (Campbell et al., 1992; Inamdar et al., 1993). Based on our results, we postulate that Sd selects a different set of target genes there, at least in part because its DNA-binding specificity is different in the absence of Vg.

No direct target genes for Sd in these other tissues have been identified. However, many target genes for the vertebrate Sd homolog TEF-1 are known (reviewed by Jacquemin and...
G/

TAT) is found once about every 2 kb, on average. However, regions of virtually any gene. This also holds true for Sd. The potential binding sites are predicted to occur in regulatory consensus sequences are often short and degenerate, so that binding sites is obviously important; however, binding site protein function studies. Searching genomic sequences for considered for such predictions, including the presence or goal is the ability to correctly predict the in vivo target genes decipher genetic regulatory sequences involved in development networks on a genomic scale. factor family member between different tissues. 

1 in muscle cells. The association of TEF-1 with Max may sequence specifically, it increases the target selectivity of TEF-
differentiation (Gupta et al., 1997). Because Max contacts DNA

1993). The gene

17234-17240.

172

436

21

127

15

17

3560.

358

439

15

438

118

719-729.

R245-R247.


REFERENCES


Many thanks to Jim Williams for the mosquito Vg sequence. We thank Kirsten Guss for sharing information on the Sd-binding sites in the cat, knt and sal cis-regulatory elements before publication. We also thank Vivian Bardwell for the TNT vectors; Vivian Bardwell, Meinrad Busslinger, Thomas Czerny, Allen Laughon, Richard Mann and M. Alessandra Vigano for helpful discussions; Allen Laughon and Craig Nelson for comments on the manuscript; and Jamie Carroll for help with its preparation. This work was supported by the Howard Hughes Medical Institute (S. B. C.).
The target selectivity of the Vg-Sd protein complex 3305


