MOLECULAR AND CELLULAR BIOLOGY, Nov. 1992, p. 4809–4816
0270-7306/92/114809-08$02.00/0

Vol. 12, No. 11

Identification of a Novel Dimer Stabilization Region in a Plant bZIP Transcription Activator

FUMIAKI KATAGIRI,† KATJA SEIPEL,‡ AND NAM-HAI CHUA*

Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021

Received 13 April 1992/Returned for modification 30 June 1992/Accepted 31 July 1992

We have carried out deletion analyses of a tobacco transcription activator, TGA1a, in order to define its functional domains. TGA1a belongs to the basic-region–leucine zipper (bZIP) class of DNA-binding proteins. Like other proteins of this class, it binds to its target DNA as a dimer, and its bZIP domain is necessary and sufficient for specific DNA binding. A mutant polypeptide containing the bZIP domain alone, however, shows a lower DNA-binding affinity than the full-length TGA1a. The C-terminal portion of TGA1a, which is essential for the higher DNA-binding affinity, contains a polypeptide region that can stabilize dimeric forms of the protein. This polypeptide region is designated the dimer stabilization (DS) region. Under our in vitro conditions, TGA1a derivatives with the DS region and those without the region do not form a detectable mixed dimer. This result indicates that in addition to the leucine zipper, the DS region can serve as another determinant of the dimerization specificity of TGA1a. In fact, the DS region, when fused to another bZIP protein, C/EBP, can inhibit dimer formation between the fusion protein and native C/EBP, whereas each of these can form homodimers. Such a portable determinant of dimerization specificity has potential application in studies of DNA-binding proteins as well as in biotechnology.

Many eukaryotic DNA-binding proteins are composed of discrete domains of defined functions (e.g., the DNA-binding domain and the transactivation domain [11, 22]). The functional properties of such domains are, in many cases, portable, i.e., their properties are retained when they are fused to another protein. This observation has encouraged the use of molecular dissection of DNA-binding proteins to define polypeptide regions that mediate specific functions.

The basic-region–leucine zipper (bZIP) (18) domain of DNA-binding proteins is a well-defined DNA binding domain structure (29). This domain dimerizes by the formation of a coiled-coil structure in the leucine zipper region (25), and the dimerization is thought to be essential for DNA binding (6, 19, 27, 28). Generally, a bZIP protein can form dimers with a distinct subset of bZIP proteins. Not only can heterodimer formation provide the opportunity to generate DNA-binding proteins of novel specificities and functions, but nonspecific heterodimer formation can also disrupt specific gene regulation mediated by these proteins. Therefore, it is of interest to identify the structural features that determine the dimerization specificities of bZIP proteins.

TGA1a (12) is a tobacco DNA-binding protein that specifically binds to the activation sequence 1 (*as*-1) element of the cauliflower mosaic virus 35S promoter (15). This protein can function as an *as*-1-specific transcription activator in vitro (13, 30) and in vivo (23a). Although TGA1a has a bZIP domain structure, it has not been shown that this domain in fact functions in dimerization and DNA binding. The possibility that other polypeptide regions of the protein may modulate its dimerization and/or DNA-binding activity also exists.

To address these questions, we have constructed deletion derivatives of TGA1a and analyzed their DNA-binding activities by gel retardation assays. Our results demonstrate that the bZIP domain of TGA1a is necessary and sufficient for sequence-specific DNA binding. Furthermore, we have identified another functional region in TGA1a, designated the dimer stabilization (DS) region, that stabilizes the dimeric form of this bZIP protein. The DS region, as well as the leucine zipper domain, functions as a determinant of dimerization specificity.

MATERIALS AND METHODS

Plasmid construction. The plasmid pKT1a (13), which contains the entire coding region of TGA1a (1,119 bp encoding 373 amino acid residues), was used as the starting construct to generate deletion mutants of TGA1a. The following restriction sites within the coding region were used as breakpoints to generate the mutants (numbers in parentheses are the nucleotide numbers of the 5′ ends of the cutting site for the enzymes): *Hpa*II (238), *Stu*I (326), *Nhe*I (431 and 458), *Afl*II (533), *Sac*I (762), and *Bgl*II (946). After pKT1a was digested with one of these restriction enzymes, the staggered ends were, if needed, either filled in with Klenow enzyme (in the case of 5′ overhang) or removed with T4 DNA polymerase (in the case of 3′ overhang). A *Cla*I linker was then ligated to the flushed ends. The following procedures were used to generate the C-terminal, N-terminal, or internal deletion mutants.

(i) C-terminal deletion mutants. DNA was digested with *Eco*RI and *Cla*I, and the fragments corresponding to the N-terminal regions of TGA1a were purified by agarose gel electrophoresis and cloned into the *Eco*RI-*Cla*I site of pKE1 (13). The C termini of the deletion mutants were fused to one of the stop codons that are present in three reading frames in pKE1. This method was used to construct the plasmids encoding the C-terminal deletion mutants, ΔC108, ΔC144,

* Corresponding author.
† Present address: Department of Genetics, Harvard Medical School, and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114.
‡ Present address: Institute of Molecular Biology II, University of Zurich, CH-8057 Zurich, Switzerland.

ΔC178, ΔC252, and ΔC316 (each number refers to the last amino acid residue in the deletion mutant).

(ii) N-terminal deletion mutants. DNA was digested with *Cla*I and *Xho*I, and the fragments corresponding to the C-terminal regions of TGA1a were purified by agarose gel electrophoresis and cloned into the *Cla*I-*Xho*I site of pKE1. After isolation of these plasmids, to obtain in-frame fusions of the initiator methionine codon in pKE1 and the deleted coding region of TGA1a, if necessary, the plasmids were redigested with *Cla*I and the staggered ends were either filled in with Klenow enzyme or removed with mung bean nuclease and then religated. This procedure was used to construct the N-terminal deletion mutants, ΔN80, ΔN110, ΔN154, ΔN178, ΔN255, and ΔN316 (each number refers to the position of the N-terminal amino acid residue in the deletion mutant).

(iii) Internal deletion mutants. Internal deletion mutants, Δ108-154, Δ144-178, and Δ178-255, were constructed by cloning *Cla*I-*Xho*I fragments containing the appropriate N-terminal deleted sequences into the *Cla*I-*Xho*I sites of the appropriate C-terminal deletion constructs (e.g., ΔN154 and ΔC108 were used for construction of Δ108-154). After isolation of these plasmids, if necessary, the reading frames at the *Cla*I site were adjusted as described for the N-terminal deletions. pKT1a was digested with *Nhe*I and then religated to obtain Δ144-154.

The 50-bp *Eco*RI-*Aat*II fragment of ΔN80 purified by polyacrylamide gel electrophoresis (PAGE) was exchanged with the 290-bp *Eco*RI-*Aat*II fragments of the C-terminal deletion mutants and the internal deletion mutants to generate ΔN80ΔC108, ΔN80ΔC144, ΔN80ΔC178, ΔN80ΔC252, ΔN80ΔC316, ΔN80Δ108-154, ΔN80Δ144-154, ΔN80Δ144-178, and ΔN80Δ178-255.

For in vitro translation of the TGA1a derivatives, the plasmids were digested with *Hin*dIII and *Eco*RI, and the *Hin*dIII-*Eco*RI fragment was exchanged with the synthesized DNA fragment shown below. This fragment contains a sequence mimicking a consensus sequence surrounding eukaryotic translation initiation sites (14) (the initiator methionine codon is indicated).

```
                Met
5' AGCTTGCCGCCACCATG     3'
3'      ACGGCGGTGGTACTTAA 5'
```

The 1.7-kb *Bam*HI-*Hin*dIII fragment that contains the entire coding region of C/EBP was cloned from pMSV-C/EBP-wt (4) into the *Bam*HI-*Hin*dIII site of pBluescript II KS+ (Stratagene) to obtain pC/EBP. The fragment containing the bZIP domain of C/EBP was prepared from pC/EBP by polymerase chain reaction by using the following primers: primer 1, 5'. . . GACCTGCAGGGCAAGGCCAAGAAGT CG. . .3'; primer 2, 5'. . . GACCAAGGAGCTAGCAGGC AGCTG. . . 3'. These primers were designed to amplify the sequence between nucleotides 946 and 1162, according to Landschulz et al. (17), to add three extra codons at the 5' end of primer 1, including a *Pst*I site, and to create an *Nhe*I site in the middle of the sequence corresponding to primer 2. Plasmid pKT1a, containing Kozak's consensus sequence (see above) (pKT1aK), was digested with *Eco*RI, and the overhang was removed with mung bean nuclease. The linearized pKT1aK was then digested with *Nhe*I. The fragment amplified with primers 1 and 2 was digested with *Nhe*I and ligated into this *Eco*RI (flushed)-*Nhe*I (the second *Nhe*I site in Fig. 1) site of pKT1aK to obtain pSC/EBP.DS. pSC/EBP.DS contains the initiation codon of pKT1a, the coding region of the bZIP domain of C/EBP (amino acid residues 272 through 346, according to Landschulz et al. [17]), and the C-terminal region of TGA1a (amino acid residues 154 through 373), all in the same reading frame. The 650-bp *Nco*I-*Pst*I fragment (corresponding to amino acid residues 1 through 217) of pC/EBP was cloned into the *Nco*I (created at the initiation codon in pSC/EBP.DS during cloning)-*Pst*I site of pSC/EBP.DS to obtain pC/EBP.DS. The 230-bp *Hin*dIII-*Nhe*I fragment of pSC/EBP.DS (the coding region of the C/EBP bZIP domain) was cloned into the *Hin*dIII-*Nhe*I site of the construct for ΔC144 to obtain pSC/EBP. pC/EBP and pSC/EBP were used for expression of C/EBP and its small derivative (SC/EBP), respectively. pC/EBP.DS and pSC/EBP.DS were used for expression of the long and short versions of the C/EBP bZIP-DS fusion proteins (C/EBP.DS and SC/EBP.DS), respectively.

Binding-site sequences. The HW oligonucleotide (13) was used as the TGA1a-binding site, *hex-1* (12, 16, 21). For the perfect palindrome site, P1 (5'. . . TCTGACGTCAGA. . . 3'), which has a higher TGA1a-binding affinity (25a) than *hex-1*, the following oligonucleotide was used.

```
5'. . . AATTCGATAGTGGGATTCTGACGTCAGAATCCCACTATCG    . . .3'
3'. . .    GCTATCACCCTAAGACTGCAGTCTTAGGGTGATAGCTTAA. . .5'
```

For the C/EBP binding site, a 100-bp *Xho*I-*Cla*I fragment containing the sequence 5'. . .CGATTGCGCAATC. . .3' was used (20a). All binding sites were labeled by Klenow fill-in reaction and used as probes in gel retardation assays.

DNA-binding assays. All plasmid constructs for the expression of DNA-binding proteins except C/EBP were linearized by *Bam*HI digestion (the *Bam*HI site is located 3' to the coding region) and transcribed with T7 RNA polymerase in vitro by using an RNA transcription kit (Stratagene). The C/EBP construct was linearized by *Hin*dIII digestion instead. The synthesized RNA species were translated in vitro with rabbit reticulocyte lysate (Promega) according to the supplier's instructions, and the synthesized polypeptides were labeled with [^{35}S]methionine. After translation, the mixture was tested for DNA-binding activity of the derivatives by gel retardation assay. The binding reaction (15 μl) contained 5 μl of the translation mixture, 0.6 μg of bovine serum albumin, 4.5 μg of poly(dI-dC), and the indicated DNA-binding probe in buffer B-0.04 (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid]-KOH [pH 7.5], 1 mM EDTA, 10% glycerol, 0.5 mM dithiothreitol, 40 mM KCl), unless otherwise indicated. The mixture was incubated at room temperature for 30 min and then analyzed by 5% PAGE. After electrophoresis, the gel was dried on DE81 paper (Whatman) and subjected to autoradiography. The in vitro translation products (0.5-μl aliquots) were also analyzed by sodium dodecyl sulfate (SDS)–15% PAGE to verify their sizes and amounts. The SDS-polyacrylamide gel was treated with Enlightening (NEN) according to the supplier's instructions before being subjected to autoradiography.

For mixing experiments, the RNAs were translated separately, and 2.5 μl of each translation reaction was added to the DNA-binding reaction mixture without the DNA-binding probe and poly(dI-dC). After a 30-min incubation at room temperature, the DNA-binding probe and poly(dI-dC) were added to initiate the binding reaction.

Apparent K_d values in DNA binding. Assuming a simple equilibrium in the interaction between a DNA-binding protein and its target DNA, the dissociation constant (K_d) can be represented as follows: K_d = [DNA][protein]/[protein ·

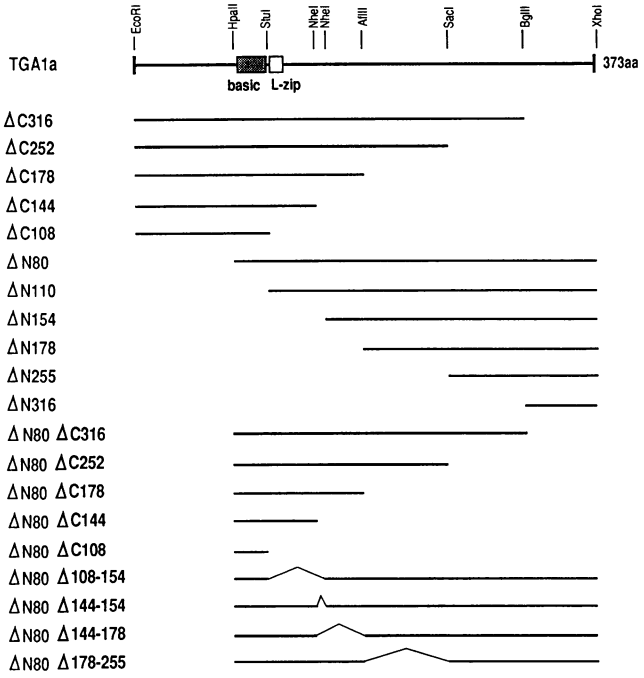

FIG. 1. Deletion derivatives of TGA1a. Top, schematic diagram of TGA1a which contains 373 amino acid (aa) residues. The basic domain and the leucine zipper (L-zip) are boxed. The deletion mutants are shown below the diagram, with N and C representing N-terminal and C-terminal deletions, respectively. The numbers following N and C refer to the positions of the terminal amino acid residues in the derivatives. Internal deletions are indicated by the positions of the terminal amino acid residues connected by a hyphen. The derivatives contain a small number of additional amino acid residues at the deleted termini, because of plasmid construction (see Materials and Methods). Restriction sites used for mutant constructions are indicated. For further details, see Materials and Methods.

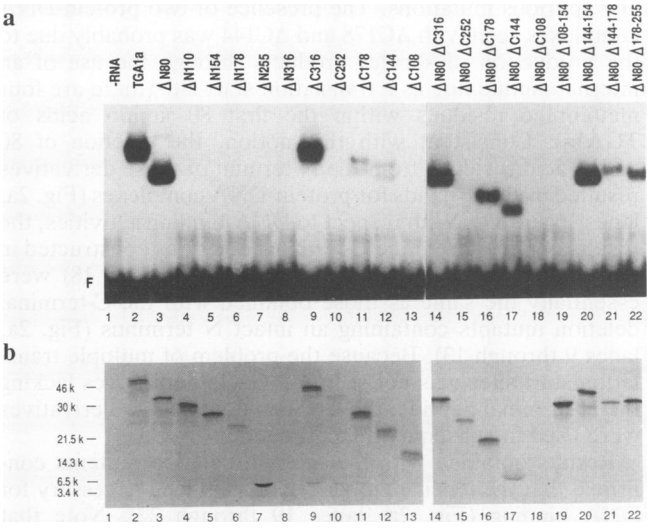

FIG. 2. DNA-binding activities of TGA1a deletion derivatives. (a) DNA-binding activities of the deletion derivatives were tested by gel retardation assays with *hex*-1 as a binding probe (1.3×10^{-8} M). Lane 1, no-RNA control; lane 2, full-length TGA1a; lanes 3 through 8, N-terminal deletion mutants; lanes 9 through 13, C-terminal deletion mutants; lanes 14 through 18, progressive C-terminal deletions of ΔN80; lanes 19 through 22, internal deletion mutants of ΔN80. See Fig. 1 for details of the derivatives. F, free probe. (b) Analysis by SDS-PAGE of in vitro translation products shown in panel a. Lane numbers correspond to the lane numbers in panel a. Migration positions of M_r markers are shown on the left.

DNA], where [protein], [DNA], and [protein · DNA] are concentrations of the DNA-binding protein, its target DNA, and the protein-DNA complex, respectively.

When an excess amount of DNA compared with the amount of protein is used (and therefore [DNA]>>[protein · DNA]), the concentration of DNA at equilibrium ([DNA]) can be approximated by the concentration of total DNA ([total DNA]). In practice, these conditions can be ensured by monitoring whether the amount of free DNA far exceeds that of the protein-DNA complex. When a certain concentration of DNA ($[DNA]_{1/2}$) gives half-saturation binding, i.e., [protein] = [protein · DNA], then K_d = $[DNA]_{1/2}$ = $[total\ DNA]_{1/2}$. Thus, K_d values can be determined simply by the titration of DNA in binding assays. Because we do not know whether this simple equilibrium indeed occurs in our system, K_d values determined this way are referred to as apparent K_d values.

RESULTS

The bZIP domain is necessary and sufficient for sequence-specific DNA binding. A series of deletion derivatives of TGA1a (Fig. 1) produced by in vitro transcription and translation were assayed by gel retardation by using the TGA1a-binding-site probe, *hex*-1 (12, 16, 21). Figure 2a shows that the bZIP domain of TGA1a is necessary for DNA binding (lanes 1 through 13). In this assay, full-length TGA1a (Fig. 2a, lane 2) produced a distinct protein-DNA complex compared with a control mixture containing no RNA (Fig. 2a, lane 1). The sequence specificity of this binding was confirmed by competition with an excess of the unlabeled probe and with a mutant version of *hex*-1, which is unable to bind TGA1a (16) (data not shown). Among the six N-terminal deletion derivatives tested (Fig. 2a, lanes 3 through 8), only ΔN80, which contains the entire bZIP domain, was able to bind the DNA probe (Fig. 2a, lane 3). Deletion of the basic region (amino acid residues 80 through 110) was sufficient to block specific DNA-binding activity (Fig. 2a, lane 4). DNA-binding activity was similarly abolished when the C-terminal deletion included the leucine zipper region (cf. ΔC144 and ΔC108 [Fig. 2a, lanes 12 and 13, respectively].

Analyses of progressive C-terminal deletion derivatives, however, showed the following unexpected results. (i) Whereas the initial 57-amino-acid deletion from the C terminus did not affect DNA-binding activity (ΔC316 [Fig. 2a, lane 9]), an additional 64-amino-acid deletion severely reduced the activity (ΔC252 [Fig. 2a, lane 10]). (ii) DNA-binding activity could be recovered by further C-terminal deletions extending to amino acids 178 and 144; however, these derivatives produced two complexes, and the intensities of the bands were weaker than that of the band produced by the full-length TGA1a (ΔC178 and ΔC144 [Fig. 2a, lanes 11 and 12, respectively]). SDS-PAGE analyses showed that translation products of the expected sizes were obtained. Moreover, the amounts of all products were comparable (Fig. 2b), thus excluding the trivial possibility that TGA1a derivatives that showed weak DNA-binding activities or none at all were produced in low amounts. We have no good explanation as to why ΔC252 was inactive in DNA binding; sequence determination verified that its bZIP domain does not contain

any spurious mutations. The presence of two protein-DNA complexes seen with ΔC178 and ΔC144 was probably due to the production of shorter products through the use of an internal methionine as a translation start site (there are four methionine residues within the first 80 amino acids of TGA1a). Consistent with this notion, the deletion of 80 amino acid residues from the N termini of these derivatives resulted in single bands for protein-DNA complexes (Fig. 2a, lanes 16 and 17). With respect to DNA-binding activities, the results obtained from the C-terminal deletions constructed in the context of ΔN80 (Fig. 2a, lanes 14 through 18) were essentially the same as those obtained with the C-terminal deletion mutants containing an intact N terminus (Fig. 2a, lanes 9 through 13). Because the problem of multiple translation start sites was not seen in TGA1a derivatives lacking the N-terminal peptide, the ΔN80 mutant and its derivatives were used in subsequent experiments.

Results obtained with the internal deletion mutants confirmed that the leucine zipper region is indeed necessary for DNA binding (Fig. 2a, lanes 19 through 22). Note that ΔN80Δ144-178 and ΔN80Δ178-255 showed lower DNA-binding activities than ΔN80Δ144-154. ΔN80ΔC144, which does not contain many amino acid residues outside the bZIP domain, showed clear DNA-binding activity (Fig. 2a, lane 17). The sequence specificity of this binding was confirmed by competition experiments using an excess of unlabeled probe and a mutant version of *hex*-1 (data not shown). All mutants which contain the bZIP region, except ΔC252 and ΔN80ΔC252, showed DNA-binding activity. Therefore, we conclude that the bZIP domain of TGA1a is not only necessary but also sufficient for binding to target DNA.

A C-terminal region is involved in high-affinity DNA binding. The DNA-binding affinities of TGA1a mutants appeared to be low when the polypeptide region between amino acids 154 and 316 or a portion of this region was deleted (Fig. 2a). Although we have verified by SDS-PAGE that comparable amounts of mutant polypeptides were produced in vitro, we could not rule out the possibility that a subpopulation of the products were inactive in DNA binding, e.g., because of inappropriate folding of the polypeptides. To exclude this possibility, we titrated the amounts of active DNA-binding proteins in the in vitro translation mixtures for TGA1a, ΔN80, and ΔN80ΔC144. With *hex*-1 as a binding probe (Fig. 3a), we found that saturated binding was not obtained even at the highest concentration of DNA tested (2.5×10^{-8} M). Therefore, we used a high-affinity binding site for TGA1a, P1 (25a) (Fig. 3b), as a binding probe. For TGA1a and ΔN80 (Fig. 3b, lanes 2 through 5 and 6 through 9, respectively), the P1 DNA concentration range used was high enough to reach saturation binding. For both proteins, the DNA concentration for half-saturation ($[DNA]_{1/2}$ and apparent K_d; see Materials and Methods) is less than 1.6×10^{-9} M. Because the lanes for samples containing 3.9×10^{-10} M DNA probe (Fig. 3b, lanes 5 and 9) do not satisfy the condition [DNA]>>[DNA · protein], we cannot tell whether the $[DNA]_{1/2}$ is higher or lower than 3.9×10^{-10} M. In contrast, saturated binding with ΔN80ΔC144 was not obtained, even at the highest concentration of DNA tested (2.5×10^{-8} M P1), indicating that the apparent K_d value for this small derivative has a lower limit of 6.3×10^{-9} M. Taken together, these results suggest that the affinities of binding of TGA1a and ΔN80 to the P1 probe are apparently at least four times higher than that of ΔN80ΔC144. At the highest DNA concentration, the intensity of the shifted band for ΔN80ΔC144 (Fig. 3b, lane 10) was comparable to those for TGA1a and ΔN80 (Fig. 3b, lanes 2 and 6, respectively), indicating that

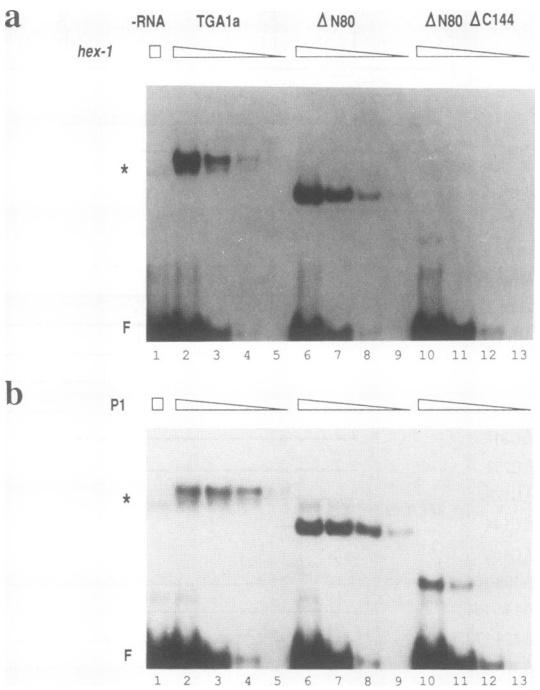

FIG. 3. TGA1a derivatives show differential DNA-binding affinities. (a) DNA-binding affinities of TGA1a, ΔN80, and ΔN80ΔC144 to *hex*-1. Concentrations of the *hex*-1 probe are as follows: lanes 1, 2, 6, and 10, 2.5×10^{-8} M; lanes 3, 7, and 11, 6.3×10^{-9} M; lanes 4, 8, and 12, 1.6×10^{-9} M; lanes 5, 9, and 13, 3.9×10^{-10} M. Note that in all three cases, 2.5×10^{-8} M *hex*-1 was not sufficient to give saturated binding. (b) Affinities of binding of TGA1a, ΔN80, and ΔN80ΔC144 to the P1 probe. P1 concentrations correspond to those of *hex*-1 in panel a. Note that 6.3×10^{-9} M P1 was sufficient to give saturated binding with TGA1a and ΔN80, whereas even 2.5×10^{-8} M was not a high enough concentration to saturate binding to ΔN80ΔC144. Each binding assay contained 2.5 μl of reticulocyte lysate after translation in both procedures. −RNA, no-RNA control; F, free probe; ★, nonspecific binding activity from reticulocyte lysate.

the number of ΔN80ΔC144 molecules active in DNA binding was at least comparable to those of TGA1a and ΔN80. Therefore, the apparent low DNA-binding affinity of ΔN80ΔC144 for *hex*-1 observed before (Fig. 2a) indeed reflects a change in the apparent K_d. If we assume that the other derivatives that show lower DNA-binding activities in Fig. 2a (ΔC178, ΔC144, ΔN80ΔC178, ΔN80Δ144-178, and ΔN80Δ178-255) also have higher apparent K_ds, the C-terminal region of TGA1a should contain a functional region that is important for high DNA-binding affinity. Because of its greater sensitivity, the P1 probe was used instead of *hex*-1 in the following experiments.

The C-terminal region contains a region that stabilizes the dimeric form. All bZIP proteins investigated thus far have been shown to bind to their target DNA sites as dimers (e.g., references 6, 9, 19, and 25–29). To see whether this was also true for TGA1a, the abilities of the derivatives of TGA1a with different lengths to form mixed dimers were tested by gel retardation assays (9) (Fig. 4). When RNAs encoding TGA1a and ΔN80 were cotranslated (Fig. 4, lane 6), the products gave an additional protein-DNA complex migrating between the complexes corresponding to separately translated TGA1a (lane 2) and ΔN80 (lane 3). This result indicates that TGA1a binds to DNA as an oligomer and most likely as

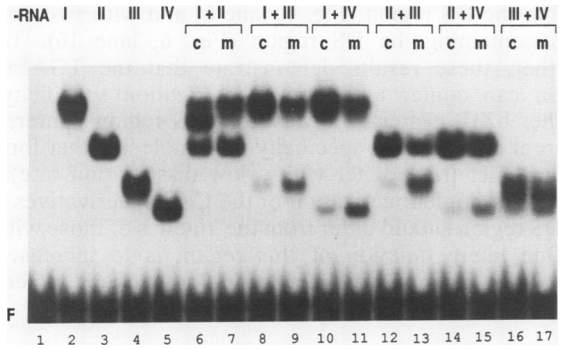

FIG. 4. Mixed-dimer formation by pairs of TGA1a derivatives. Heterodimer formation was tested with pairwise combinations of TGA1a (I), ΔN80 (II), ΔN80ΔC178 (III), and ΔN80ΔC144 (IV). The TGA1a derivatives were assayed for binding to the P1 probe (1.3×10^{-8} M) individually (lanes 2 through 5), with two of them cotranslated (c), and with two of them mixed after separate translations (m). Note that mixed-dimer formation is observed as an additional band with intermediate mobility (lanes 6, 16, and 17). −RNA, no-RNA control; F, free probe.

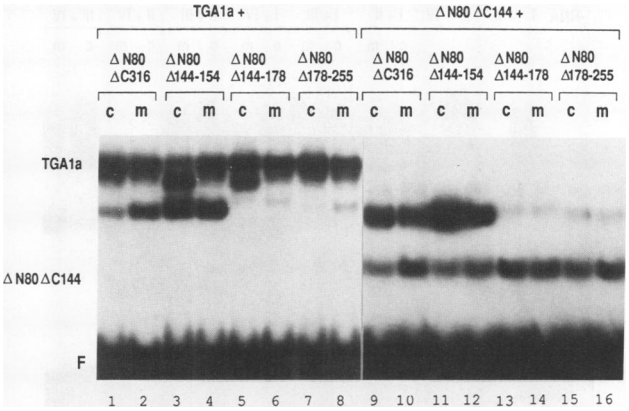

FIG. 5. Mixed-dimer formation with TGA1a and ΔN80ΔC144. The abilities of ΔN80ΔC316, ΔN80Δ144-154, ΔN80Δ144-178, and ΔN80Δ178-255 to form mixed dimers with TGA1a and ΔN80ΔC144 were tested. c, cotranslation; m, mixing after separate translations; TGA1a, position of TGA1a-DNA complex; ΔN80ΔC144, position of ΔN80ΔC144-DNA complex; F, free probe. Probe P1 (1.3×10^{-8} M) was used for the binding assay.

a dimer, like other bZIP proteins. The unique complex obtained with the cotranslation mixture likely corresponds to the mixed dimer formed between TGA1a and ΔN80, while the complexes observed for separately translated TGA1a and ΔN80 likely correspond to their homodimers.

During the course of our experiments, we noticed that groups of TGA1a derivatives exclude each other with respect to mixed-dimer formation. Figure 4 shows the results of all possible pairwise combinations for TGA1a, ΔN80, ΔN80ΔC178, and ΔN80ΔC144. When RNAs derived from these pairwise combinations of constructs were cotranslated, mixed-dimer formation was detected only for the combination of TGA1a and ΔN80 (Fig. 4, lane 6) and the combination of ΔN80ΔC178 and ΔN80ΔC144 (Fig. 4, lane 16) (the result showing mixed-dimer formation between ΔN80ΔC178 and ΔN80ΔC144 may not be very clear because of poor gel resolution, but we have confirmed this finding by separate experiments [data not shown]).

In order to examine stabilities of dimers, experiments in which two derivatives were mixed after translation, incubated for 30 min, and then used for the DNA-binding assay were performed. In these experiments, if the homodimers have significantly high dissociation rates, subunits of the homodimers formed during separate translations can be exchanged after mixing and mixed-dimer formation would be detected. Alternatively, if dimerization depends on the presence of a DNA target site (a low association rate in the absence of DNA), it could occur during the binding assay, and mixed-dimer formation can also be detected in this case. Note that in this kind of experiment, we cannot measure differences between the overall equilibrium constants of dimerization. Because in the case of TGA1a and ΔN80, mixed-dimer formation was not observed by mixing of separately translated products prior to the binding assay (Fig. 4, lane 7), these two proteins must form stable dimers. In contrast, both ΔN80ΔC178 and ΔN80ΔC144 form unstable dimers (Fig. 4, lane 17). The most likely explanation for these phenomena is that the C-terminal portion of TGA1a, deleted in ΔN80ΔC178 and ΔN80ΔC144, contains a region that stabilizes the dimeric form even in the absence of target DNA.

The hypothesis that the C-terminal region encodes a dimer stabilization function can explain why mixed-dimer formation was not detected in pairwise combinations of derivatives with and without this region. A derivative containing this region can form a stable homodimer even in the absence of target DNA. As a result, its monomer concentration would decrease rapidly in the reaction mixture. This process drastically reduces the chance for the derivative without the C-terminal region to form a mixed dimer. We designated this functional region (amino acid residues 144 through 373) the DS region. Note that the two derivatives without the DS region showed low DNA-binding affinities (Fig. 2a). The lower DNA-binding affinity might result from the lower efficiency of dimerization; alternatively, a domain involved in a higher DNA-binding affinity may function independently of dimerization, as in the case of the serum response factor (24).

We investigated other TGA1a mutants in an attempt to assess whether the following three phenomena are correlated with each other: (i) the ability to form a stable dimer, (ii) the exclusion of mixed-dimer formation, and (iii) the high DNA-binding affinity. ΔN80ΔC316 and ΔN80Δ144-154 showed high DNA-binding affinities while ΔN80Δ144-178 and ΔN80Δ178-255 showed low DNA-binding affinities (Fig. 2a). We tested whether these derivatives can form mixed dimers with either TGA1a or ΔN80ΔC144 and whether the mixed dimers are stable. Figure 5 shows that ΔN80Δ144-154 (lanes 3, 4, 11, and 12) and ΔN80Δ144-178 (lanes 5, 6, 13, and 14) formed stable mixed dimers only with TGA1a while ΔN80ΔC316 (lanes 1, 2, 9, and 10) and ΔN80Δ178-255 (lanes 7, 8, 15, and 16) did not form mixed dimers with either TGA1a or ΔN80ΔC144. The location of the DS region can be narrowed down to amino acid residues 178 through 373 because ΔN80Δ144-178 formed a stable mixed dimer with TGA1a. These results also indicate that the abilities to form mixed dimers with TGA1a and ΔN80ΔC144 are not correlated with high and low DNA-binding affinities, respectively. Furthermore, the finding that ΔN80ΔC316 and ΔN80Δ178-255 did not form mixed dimers with either TGA1a or ΔN80ΔC144 suggests that there is at least another group of derivatives that is defined by the ability to form mixed dimers within the group. It should be emphasized, however, that our interpretation of the results from this type of

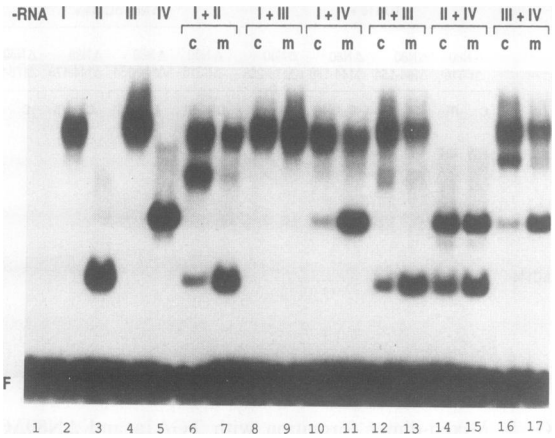

FIG. 6. Effect of the DS region on dimerization of C/EBP. Mixed-dimer formation by pairs of C/EBP derivatives containing (C/EBP.DS [III] and SC/EBP.DS [IV]) or not containing (C/EBP [I] and SC/EBP [II]) the DS region was tested. The C/EBP-binding site (0.5×10^{-9} M) was used for the assay. The derivatives were assayed individually (lanes 2 through 5), with two of them cotranslated (c), and with two of them mixed after separate translations (m). −RNA, no-RNA control; F, free probe.

experiment is based on the assumption that the deletion of a region of TGA1a does not affect the functions of other regions. Because this assumption may not necessarily be correct, we did not investigate the deletion derivatives further.

The DS region function is portable. The function of the DS region can be most convincingly demonstrated by a gain-of-function experiment. To this end, we chose to use the bZIP protein C/EBP (17) as an acceptor protein because it has been reported to form an unstable dimer (19). To examine whether fusion of the DS region to C/EBP can stabilize its dimeric form, we investigated mixed-dimer formation of all possible pairwise combinations for C/EBP, SC/EBP (a small version of C/EBP that contains the bZIP domain), C/EBP.DS (a fusion protein containing a large part of the C/EBP and the DS region [amino acid residues 154 through 373] of TGA1a), and SC/EBP.DS (a fusion protein of SC/EBP and the DS region) (see Materials and Methods for details of the constructs). However, we found that C/EBP can already form a stable dimer even in the absence of DNA under our experimental conditions (Fig. 6; compare lanes 6 and 7). A number of factors could explain the discrepancy between our results and those of Landschulz et al. (19). First, we used rabbit reticulocyte lysate to synthesize C/EBP in vitro, whereas Landschulz et al. (19) produced the protein in *Escherichia coli*. There could be differences in protein modification between the two systems. Second, the smaller derivatives of C/EBP used for mixed-dimer formation are different; ours (SC/EBP) contains amino acid residues 272 through 346, whereas their derivative contains amino acid residues 253 through 359. Third, different buffers were used for incubation of separately produced polypeptides. Notwithstanding the formation of a stable dimer by C/EBP, we found that fusion of the DS region to C/EBP strongly inhibited the formation of mixed dimer with C/EBP derivatives without the DS region (Fig. 6, lanes 6, 8, 10, 12, 14, and 16), although the result of one combination (Fig. 6, lane 8) was unclear because of similar migration rates of the protein-DNA complexes of the two constructs. Efficient mixed-dimer formation was detected only with two derivatives lacking the DS region (Fig. 6, lane 6) and with two derivatives containing the DS region (Fig. 6, lane 16). Taken together, these results demonstrate that the TGA1a DS region can confer a different dimerization specificity on another bZIP protein. Because the DS region conferred a different dimerization specificity on a molecule that forms a stable dimer (i.e., dimer with a low dissociation rate), the kinetics of homodimerization of the C/EBP derivatives with the DS region should differ from the kinetics of those without it. One likely function of this region is to increase the association rate during dimer formation between molecules containing it.

DISCUSSION

In this paper, we have analyzed the abilities of various deletion mutants of TGA1a to bind to target DNA in vitro. We found that the bZIP domain of TGA1a is sufficient for sequence-specific DNA binding, although high-affinity binding requires a polypeptide region located in the C-terminal portion downstream of the bZIP domain. In this C-terminal portion, TGA1a contains a DS region that stabilizes its dimeric form.

To detect dimer formation, we cotranslated RNAs encoding two TGA1a derivatives of different molecular sizes and analyzed their abilities to form a mixed dimer. The dimer was identified as a protein-DNA complex that migrates at an intermediate rate relative to the complexes for the two homodimers in gel retardation assays (9). Our study has demonstrated that mixed dimers are not formed if one of the two derivatives lacks the DS region. Therefore, when mixed-dimer formation is not detected in gel retardation assays, the results should be interpreted with caution. In addition, there are other circumstances under which a mixed dimer would not be detected. The mixed dimer may not have a DNA-binding affinity comparable to those of the homodimers. Alternatively, the migration rate of the DNA-protein complex for the mixed dimer may not be distinguishable from those of the complexes for the homodimers. To exclude these possibilities, an assay that detects dimers directly should be performed. Moreover, a direct assay of dimerization would allow us to address the question of whether TGA1a derivatives lacking the DS region (ΔN80ΔC178 and ΔN80ΔC144 in Fig. 4) can form a dimer in the absence of the target DNA. This question is important for distinguishing whether the observed effect of the DS region on dimer stabilization was due to a decrease in the dissociation rate or an increase in the association rate.

TGA1a has two determinants of dimerization specificity, the leucine zipper and the DS region. So far, this type of structural feature has not been described for other bZIP proteins. Among eukaryotic DNA-binding proteins, basic-region–helix-loop-helix–zipper (bHLH-Zip) proteins, including the Myc family (reviewed in reference 20), AP-4 (10), USF (8), TFE3 (1, 2), and Max (3), have been shown to possess more than one determinant of dimerization specificity within the molecule. Dimer formation of these bHLH-Zip proteins is regulated by the HLH region as well as the leucine zipper region (AP-4 contains a second leucine zipper region, which also affects dimer formation, in addition to the leucine zipper close to the HLH region [10]). Both HLH (23) and leucine zipper (18, 25, 29) are known to be dimerization motifs, and by analogy, the DS region of TGA1a could contain a dimerization motif(s). Although the primary sequence of the TGA1a DS region does not reveal that the region contains any structure typical of dimerization motifs,

two observations support this contention. First, results from the fusion of the DS region to C/EBP suggest that this region increases the association rate of the dimer. For an intramolecular region to increase the association rate, the simplest interpretation is that this region also has an affinity for a portion of the other subunit. Second, we found that the DS region has to be present in both subunits of the dimer in order for the dimer-stabilizing function to be expressed. These observations are consistent with the notion that the DS region contains a dimerization motif(s). To address this question, a direct dimerization assay independent of DNA-binding activity is required.

It should be emphasized that there is an important difference between the structural combination of HLH and leucine zipper and the combination of leucine zipper and DS region. In the case of bHLH-Zip, the distance between these two motifs is critical for dimerization specificity (1). In contrast, the distance between the leucine zipper and the DS region is not very critical because ΔN80Δ145-155 and ΔN80Δ145-178 form stable dimers with TGA1a (Fig. 5). Moreover, the DS region can be functionally transferred to a different bZIP protein. Thus, the DS region can function in a position-independent manner, whereas the leucine zipper of bHLH-Zip proteins appears somewhat structurally integrated with the HLH motif.

Certain selective advantages may be conferred by the combination of distinct determinants that together regulate dimerization of the molecule. Hu et al. (10) have pointed out that such a structural feature can confer a higher dimerization specificity on a DNA-binding protein. They also raised the possibility that modifications in the two leucine zipper motifs of AP-4 can lead to a change of its dimerization partner. Deletion of the two leucine zipper motifs allows AP-4 to form a heterodimer with other bHLH proteins with which the full-length AP-4 is unable to interact. Similarly, TGA1a might change its dimerization partner by modifications in the DS region. In addition to these possibilities, we can speculate on the possible advantage of this bipartite determinant of dimerization during the evolution of DNA-binding proteins. If the combination of these determinants of dimerization specificity in a DNA-binding protein was changed by DNA recombination or alternative splicing, the resulting new DNA-binding protein would not be able to form a heterodimer with its parental molecule. This process would increase, within a relatively short time, the number of DNA-binding proteins that can function independently in a single cell. These independent sets of DNA-binding proteins could diverge into proteins that mediate different regulatory functions. In this respect, we note that the bZIP domains of TGA1a and a TGA1a-related protein are encoded by exons separated from the other parts of their genes (5). Such a genomic organization would certainly facilitate the exchange of DNA-binding domains and DS regions by either alternative splicing or exon shuffling (7). The portability of the DS region function also supports this hypothesis.

Finally, the functional portability of the DS region suggests a potential use for this region in the investigation of DNA-binding proteins in vivo. When genes encoding engineered DNA-binding proteins are introduced into cells, their products could interfere with the functions of the endogenous counterparts through heterodimer formation. Also, if the purpose of introducing engineered DNA-binding proteins is to analyze their behavior in vivo (e.g., nuclear localization of the protein), interaction with the endogenous counterparts may complicate subsequent analysis. Such interference can be prevented by the addition of the DS region to the engineered proteins. Moreover, because DNA-binding proteins are potential targets for gene manipulation, the DS region can be used to create an engineered DNA-binding protein that does not interfere with endogenous DNA-binding proteins, thus providing a useful tool in biotechnology application.

ACKNOWLEDGMENTS

We thank A. D. Friedman and S. L. McKnight for the C/EBP cDNA clone and its binding site DNA and A. R. van der Krol and D. Ralston for critical reading of the manuscript.

This work was supported by a grant from Monsanto Co.

REFERENCES

1. **Beckmann, H., and T. Kadesch.** 1991. The leucine zipper of TFE3 dictates helix-loop-helix dimerization specificity. Genes Dev. **5**:1057–1066.
2. **Beckmann, H., L.-K. Su, and T. Kadesch.** 1990. TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer μE3 motif. Genes Dev. **4**:167–179.
3. **Blackwood, E. M., and R. N. Eisenman.** 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science **251**:1211–1217.
4. **Friedman, A. D., W. H. Landschulz, and S. L. McKnight.** 1989. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. Genes Dev. **3**:1314–1322.
5. **Fromm, H., F. Katagiri, and N.-H. Chua.** 1991. The tobacco transcription activator TGA1a binds to a sequence in the 5′ upstream region of a gene encoding a TGA1a-related protein. Mol. Gen. Genet. **229**:181–188.
6. **Gentz, R., F. J. Rauscher III, C. Abate, and T. Curran.** 1989. Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. Science **243**:1695–1699.
7. **Gilbert, W.** 1978. Why genes in pieces? Nature (London) **271**:501.
8. **Gregor, P. D., M. Sawadogo, and R. G. Roeder.** 1990. The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. Genes Dev. **4**:1730–1740.
9. **Hope, I. A., and K. Struhl.** 1987. GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. EMBO J. **6**:2781–2784.
10. **Hu, Y.-F., B. Luscher, A. Admon, N. Mermod, and R. Tjian.** 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes Dev. **4**:1741–1752.
11. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58**:799–839.
12. **Katagiri, F., E. Lam, and N.-H. Chua.** 1989. Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature (London) **340**:727–730.
13. **Katagiri, F., K. Yamazaki, M. Horikoshi, R. G. Roeder, and N.-H. Chua.** 1990. A plant DNA-binding protein increases the number of active preinitiation complexes in a human *in vitro* transcription system. Genes Dev. **4**:1899–1909.
14. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNA. Nucleic Acids Res. **15**:8125–8132.
15. **Lam, E., P. N. Benfey, P. M. Gilmartin, R.-X. Fang, and N.-H. Chua.** 1989. Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. Proc. Natl. Acad. Sci. USA **86**:7890–7894.
16. **Lam, E., F. Katagiri, and N.-H. Chua.** 1990. Plant nuclear factor ASF-1 binds to an essential region of the nopaline synthase promoter. J. Biol. Chem. **265**:9909–9913.
17. **Landschulz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
18. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new

class of DNA binding proteins. Science **240:**1759–1764.
19. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science **243:**1681–1688.
20. **Luscher, B., and R. N. Eisenman.** 1990. New light on Myc and Myb. Part I. Myc. Genes Dev. **4:**2025–2035.
20a. **McKnight, S. L.** Personal communication.
21. **Mikami, K., T. Tabata, T. Kawata, T. Nakayama, and M. Iwabuchi.** 1987. Nuclear protein(s) binding to the conserved DNA hexameric sequence postulated to regulate transcription of wheat histone genes. FEBS Lett. **223:**273–278.
22. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:**371–378.
23. **Murre, C., P. S. McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins. Cell **56:**777–783.
23a. **Neuhaus, G., G. Neuhaus-Url, I. Potrykus, F. Katagiri, K. Seipel, and N.-H. Chua.** Unpublished data.
24. **Norman, C., M. Runswick, R. Pollock, and R. Triesman.** 1988. Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the *c-fos* serum response element. Cell **55:**989–1003.
25. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1989. Evidence that the leucine zipper is a coiled coil. Science **243:**538–542.
25a. **Qin, X.-F.** Unpublished data.
26. **Tabata, T., T. Nakayama, K. Mikami, and M. Iwabuchi.** 1991. HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat. EMBO J. **10:**1459–1467.
27. **Talanian, R. V., C. J. McKnight, and P. S. Kim.** 1990. Sequence-specific DNA binding by a short peptide dimer. Science **249:**769–771.
28. **Turner, R., and R. Tjian.** 1989. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimer. Science **243:**1689–1694.
29. **Vinson, C. R., P. B. Sigler, and S. L. McKnight.** 1989. Scissors-grip model for DNA recognition by a family of zipper proteins. Science **246:**911–916.
30. **Yamazaki, K., F. Katagiri, H. Imaseki, and N.-H. Chua.** 1990. TGA1a, a tobacco DNA-binding protein increases the rate of initiation in a plant *in vitro* transcription system. Proc. Natl. Acad. Sci. USA **87:**7035–7039.