GT-2: In Vivo Transcriptional Activation Activity and Definition of Novel Twin DNA Binding Domains with Reciprocal Target Sequence Selectivity

Min Ni, Katayoon Dehesh, James M. Tepperman, and Peter H. Quail

Department of Plant Biology, University of California, Berkeley, California 94720, and Plant Gene Expression Center, U.S. Department of Agriculture–Agricultural Research Service, Albany, California 94710

GT-2 is a novel DNA binding protein that interacts with a triplet of functionally defined, positively acting GT-box motifs (GT1-bx, GT2-bx, and GT3-bx) in the rice phytochrome A gene (PHYA) promoter. Data from a transient transfection assay used here show that recombinant GT-2 enhanced transcription from both homologous and heterologous GT-box-containing promoters, thereby indicating that this protein can function as a transcriptional activator in vivo. Previously, we have shown that GT-2 contains separate DNA binding determinants in its N- and C-terminal halves, with binding site preferences for the GT3-bx and GT2-bx promoter motifs, respectively. Here, we demonstrate that the minimal DNA binding domains reside within dual 90-amino acid polypeptide segments encompassing duplicated sequences, termed trihelix regions, in each half of the molecule, plus 15 additional immediately adjacent amino acids downstream. These minimal binding domains retained considerable target sequence selectivity for the different GT-box motifs, but this selectivity was enhanced by a separate polypeptide segment farther downstream on the C-terminal side of each trihelix region. Therefore, the data indicate that the twin DNA binding domains of GT-2 each consist of a general GT-box recognition core with intrinsic differential binding activity toward closely related target motifs and a modifier sequence conferring higher resolution reciprocal selectivity between these motifs.

INTRODUCTION

The light-induced changes in gene expression that drive photomorphogenesis in plants are directed by a set of regulatory photoreceptors of which the red/far-red light responsive phytochrome (phy) family is the best characterized (Gilmartin et al., 1990; Quail, 1991, 1994a, 1994b; Furuya, 1993; Batschauer et al., 1994; Furuya and Song, 1994; Kendrick and Kronenberg, 1994; Millar et al., 1994; Tobin and Kehoe, 1994; Quail et al., 1995; Terzaghi and Cashmore, 1995). In Arabidopsis, the phy apoprotein is encoded by five genes designated PHYA, PHYB, PHYC, PHYD, and PHYE (Sharrock and Quail, 1989; Clack et al., 1994), with sequences related to these in a variety of other plant species (Schneider-Poetsch et al., 1994; Mathews et al., 1995; Pratt, 1995).

In monocots, the PHYA gene exhibits strong and rapid negative autoregulation when phytochrome A (phyA) is photoactivated (Quail, 1994a). This autoregulation has been shown to occur at the transcriptional level within 2 to 5 min of light signal perception and in the absence of new protein synthesis (Lissemore and Quail, 1988). The data indicate therefore that the necessary components of the signal transduction chain linking the photoreceptor to this responsive gene exist in the cell before signal perception. Because of these features, the PHYA gene has been used as a model system for investigating the mechanisms of light-induced transcriptional regulation in plants (Quail, 1991, 1994a).

The use of a series of deletion and linker-scanning derivatives has defined three functionally active DNA sequence elements in the minimal promoters of PHYA genes from oats and rice (Bruce and Quail, 1990; Dehesh et al., 1990, 1992; Bruce et al., 1991). Each promoter contains two constitutively active, synergistically positive elements that drive high-level transcription in the absence of photoactivated phy and a single negative element that imposes transcriptional repression on the gene in response to phy photoactivation. The upstream positive element in the rice PHYA promoter consists of a triplet of closely related motifs, designated GT-boxes (GT1-bx, GT2-bx, and GT3-bx) (Dehesh et al., 1990, 1992; Quail, 1994a). Recombinant clones encoding a protein designated GT-2 that binds in sequence-specific fashion to the GT-box ensemble have been obtained from rice (Dehesh et al., 1990, 1992). This protein is a unique factor with no sequence similarity to...
Figure 1. The Recombinant Rice GT-2 Protein Expressed in Tobacco Protoplasts Stimulates Transcription in a Binding Site-Dependent Manner.

(A) Schematic diagrams of the GT-2 activator plasmids and corresponding GUS reporter templates used in the transient assays. For an activator, expression of the full-length GT-2 protein (GT-2) was driven by the CaMV 35S promoter with its own enhancer (35SP:GT-2). A control activator plasmid without the GT-2 insertion was also used (35SP). For reporters, the GUS gene was driven by either the rice PHYA promoter (PHYAP) truncated at position −188 (PP-188) or the CaMV 35S promoter (35S) truncated at position −59 (35S-59), with (GTB:PP-188 and GTB:35S-59) or without (PP-188 and 35S-59) the GT-box binding sites (GTBS, a fragment from positions −247 to −189 from the rice PHYA promoter) upstream. Arrowheads indicate transcription start sites. NOS3' represents the polyadenylation signal of the nopaline synthase gene.
other classes of nuclear factors currently in the data bases, except for the Arabidopsis GT-2 homolog (Kuhn et al., 1993) and a related tobacco and Arabidopsis factor called GT-1 (Gilmartin et al., 1992; Perisic and Lam, 1992; Hiratsuka et al., 1994; Lam, 1995).

The rice GT-2 protein contains a series of acidic and proline/glutamine-rich domains implicated in transcriptional activation in other factors (Johnson and McKnight, 1989; Mitchell and Tjian, 1989; Gill and Tjian, 1992; Gill et al., 1994; Tjian and Maniatis, 1994), but no direct evidence of such activity has been presented. The most striking feature, however, is a 75- amino acid sequence that is internally duplicated, once in each of the N- and C-terminal halves of the polypeptide (Dehesh et al., 1992). These duplicated sequences have been termed trihelix motifs based on secondary structure predictions indicating the potential for the formation of three \( \alpha \)-helices separated by two short loops. Each trihelix region contains a classical bipartite nuclear localization sequence (NLS) motif, and each NLS has been shown recently to function independently in nuclear targeting in vivo (Dehesh et al., 1995). In situ immunocytochemical techniques also have shown that GT-2 is localized to the nucleus in rice seedlings, and this is consistent with a role in transcriptional regulation (Dehesh et al., 1995).

In vitro DNA binding studies have shown that rice GT-2 binds preferentially to the GT2-bx and GT3-bx motifs in the GT-box triplet, with only low apparent affinity for the GT1-bx motif (Dehesh et al., 1992). This behavior is in contrast with that of the GT-1 protein that has been shown to bind preferentially to the GT1-bx motif (Gilmartin et al., 1992; Perisic and Lam, 1992; Hiratsuka et al., 1994; Lam, 1995). Moreover, in striking fashion, the two halves of the GT-2 polypeptide were found to bind independently, with high target sequence selectivity, such that the N-terminal half binds with high selectivity to the GT3-bx motif, whereas the C-terminal half binds with high selectivity to the GT2-bx motif. The duplicated nature of the trihelix region of GT-2 and the close similarity in the GT-box sequences led to the suggestion that the trihelix motif may be involved in target sequence recognition (Dehesh et al., 1992).

Recent deletion analysis studies with tobacco and Arabidopsis GT-1 have shown that the corresponding single trihelix region in each of these proteins is necessary for GT1-bx recognition (Hiratsuka et al., 1994; Lam, 1995). Lam (1995) has shown further that an additional 26 amino acids to the C-terminal side of the trihelix region of tobacco GT-1 are necessary for GT1-bx binding and that this additional segment is predicted to form a fourth \( \alpha \)-helix. Similar detailed mapping of the dual GT-2 DNA binding domains has not been reported.

In this study, we used a protoplast-based, transient transfection assay to determine whether recombinant rice GT-2 can indeed function as a transcriptional activator in vivo. In addition, we performed high-resolution deletion mapping and sequence substitution mutagenesis of GT-2 to define and characterize the minimal DNA binding domains in each of the two halves of the GT-2 polypeptide.

**RESULTS**

Recombinant Rice GT-2 Protein Activates Transcription in Tobacco Protoplasts

Protoplast-based transient assays were used to test whether the rice GT-2 protein could activate expression from its target sites in plant cells. Figures 1A to 1C show the respective activator and reporter constructs used. The full-length GT-2 protein sequence was introduced into tobacco cells under the control of the cauliflower mosaic virus (CaMV) 3SS promoter with its native enhancer (35SP-GT-2). A plasmid carrying only the 3SS promoter and nopaline synthase (nos) polyadenylation sites was used as a negative control (35SP). Two types of basal promoter-reporter constructs were made: one containing the rice GT-2 functional domains.
PHYA promoter truncated at position -188 (GT-box binding sites are upstream of -188 in the native promoter), and the other containing the TATA sequence from the CaMV 35S promoter (from nucleotide -59 relative to the transcription start site), each fused to the β-glucuronidase (GUS) reporter gene (Figure 1A). Constructs containing GT-2 binding sites were generated by fusing a sequence containing the GT-box motifs from nucleotides -247 to -188 from the rice PHYA promoter upstream of the -188 rice PHYA promoter and the -59 CaMV 35S TATA sequence.

An internal control plasmid expressing the luciferase (LUC) gene driven by the CaMV 35S promoter also was included in each transformation as a control for the efficiency of DNA transfer (Figure 1C). Three plasmids, one activator, one reporter, and one internal control along with carrier DNA were simultaneously electroporated into tobacco protoplasts, and GUS and LUC activities in the extracts were assayed. Figures 1D and 1E show the data expressed as a ratio of GUS to LUC activity.

The rice -188 PHYA promoter–GUS reporter construct (PP-188) directed a low level of expression in the tobacco protoplasts when cotransformed with either the control activator plasmid (35SP) or the plasmid expressing GT-2 (35SP:GT-2) (Figure 1D). In general, all the rice PHYA promoter–derived reporters (Figure 1D) gave relatively low activity compared with CaMV 35S promoter–derived reporters (Figure 1E). Cotransformation of the rice -247 PHYA promoter–GUS reporter construct (bearing native GT-2 binding sites upstream) (GTB:PP-188) and control activator plasmid (35SP) did not substantially increase the detectable basal activity (Figure 1D). This result indicates that the GT-2 binding sites are not targets for endogenous homologous transcriptional activators in these tobacco protoplasts. Expression was, however, stimulated an average of 17.4-fold from the rice -247 PHYA promoter when cotransformed with the activator plasmid expressing GT-2 (35SP:GT-2), relative to cotransformation with the control activator plasmid (35SP) (Figure 1D). The highest level of stimulation by GT-2 that we observed was 39-fold. GT-2 had no effect on expression from the reporter plasmid lacking the GT-2 binding site (PP-188) (Figure 1E). This result demonstrates that rice GT-2 is capable of activating gene expression in a binding site–dependent manner in the context of the native PHYA promoter in plant protoplasts.

To determine whether GT-2 could activate transcription from its DNA target sites in a heterologous promoter, CaMV 35S minimal promoter–derived reporters were used (Figure 1A). In contrast to the rice PHYA promoter–based reporters (Figure 1D), expression from the constructs carrying the -59 CaMV 35S core promoter alone (35S-59) was relatively strong, thus giving a relatively high background (Figure 1E). Nevertheless, the rice GT-2 protein (35SP:GT-2) clearly stimulated transcription over that supported by the control activator plasmid (35SP) when GT-2 binding sites were present in the reporter (GTB:35S-59) (Figure 1E). This stimulation averaged 8.5-fold. This activation also was binding site–dependent because GT-2 (35SP:GT-2) failed to activate transcription from the -59 CaMV 35S promoter lacking the GT-2 binding sites (35S-59) (Figure 1E).

As a positive control system, activators expressing yeast galactose 4 gene (GAL4) derivatives under the control of the CaMV 35S promoter also were made (Figure 1B). GABD, the 147 amino acids in the N-terminal portion of the 881-residue GAL4 protein, binds its target sites but fails to activate transcription in plant cells (Ma et al., 1988). In contrast, GABD+AD, containing the GAL4 activating region (residues 768 to 881) fused to the DNA binding portion, increases transcription significantly. The reporters bore the GUS gene fused to either the rice -194 PHYA promoter or the CaMV 35S -59 minimal promoter, each with GAL4 binding sites integrated upstream. As shown in Figures 1D and 1E, the fusion protein containing the GAL4 DNA binding domain plus activation domain (GABD+AD) augmented the transcription from the reporters containing GAL4 binding sites, an average of 11.5-fold (GAB:PP-194) and 17.5-fold (GAB:35S-59), respectively, over that of the protein containing only the GAL4 DNA binding domain (GABD) (Figures 1D and 1E). The observed effect was dependent on DNA binding, because transcription of a reporter template without GAL4 binding sites was not stimulated in the presence of the activator protein (data not shown).

### Polypeptide Segments Involved in Discrimination between Closely Related GT-Box Target Sites

Previous work has shown that DNA binding determinants in the N-terminal half of GT-2 (from amino acids 1 to 261) preferentially recognize the GT3-bx motif, whereas C-terminal DNA binding determinants (between amino acids 404 and 737) preferentially recognize the GT2-bx motif (Dehesh et al., 1992). As a first step toward identifying sequences within the two halves of the protein that bind to the GT boxes and that determine this binding selectivity, we made deletion mutants of the GT-2 protein. In this analysis, deletion mutant GT-2 polypeptides were synthesized by in vitro transcription/translation and assayed for their ability to bind the GT boxes by using electrophoretic mobility shift assays (EMSA).

Figure 2 shows the results of using a series of N- and C-terminal deletion mutants within each half of the GT-2 molecule. The deletion subclones and their corresponding synthetic polypeptides are shown in Figures 2A and 2B. DNA binding activity was evaluated using monomers of the defined GT-box probes and the synthetic protein products at the same molar ratio for all products (Figures 2C and 2D). Both short and long film exposures were used for autoradiographic detection of DNA–protein complexes to enable semiquantitative comparisons of the multiple peptide–probe combinations examined in these experiments.

Figure 2C shows the results for the N-terminal domain of GT-2. When the GT3-bx was used as a probe, deletion of 78 amino acids from the N-terminal end of the construct p1-314 had little effect on complex-forming capacity, indicating that...
these amino acids are not critical for DNA–protein interaction (at right in Figure 2C, p79-314). C-terminal deletions of construct p79-314 to residue 238 resulted in a slight apparent increase in complex-forming capacity with GT3-bx (at right in Figure 2C, p79-238), whereas further deletion to amino acid 184 rendered the complex-forming activity weak (at right in Figure 2C, p79-184). Deletion into the trihelix region completely abolished the detectable GT3-bx binding activity even in the context of an intact N terminus (at right in Figure 2C, p1-155).

When GT1-bx and GT2-bx motifs were used as probes, the two peptides, p79-261 and p79-238, exhibited some complex-forming capacity with these motifs (at left and center in Figure 2C) but at considerably lower levels than with the GT3-bx probe (at right in Figure 2C). These data indicate that these two shorter peptide segments contain information involved in discrimination between the three DNA target sites. Fivefold longer exposure of the film did not reveal significant binding of the shorter or longer peptide derivatives to these probes (data not shown). These results suggested that this absence of detectable binding for the two longer peptides (p1-314 and p79-314) with the GT1-bx and GT2-bx probes was not due simply to a lower general complex-forming capacity, as might have been implied from the pattern of interaction of these peptides with the GT3-bx probe (at right in Figure 2C). These data suggested therefore that the amino acid sequence from residues 261 to 314 is involved in enhancing differentiation between the different GT-box motifs by the N-terminal domain.

To examine this possibility more quantitatively, we performed additional binding assays at constant protein concentration and varying probe concentration. Figure 3A shows the deletion derivatives tested. Increasing concentrations of the GT3-bx probe resulted in increasing amounts of retarded complex formation for both p79-238 and p79-314 peptide segments up to saturation at ~2 nM, with little difference between the binding curves of the two derivatives (Figures 3C and 3D). Half-maximal binding occurred at ~2 nM GT3-bx. In contrast, increasing concentrations of the GT2-bx probe did not lead to binding saturation for either derivative in the range tested, and p79-238 formed much higher levels of retarded complex than did p79-314 at all probe concentrations (Figures 3B and 3D). Slightly less than half-maximal binding occurred at 8 nM GT2-bx probe (the highest concentration tested) for p79-238, whereas binding was barely detectable at this concentration for the p79-314 derivative (Figure 3D).

These data lead to several conclusions. First, because retarded complex formation saturated at the same level for p79-238 and p79-314 with the GT3-bx, the data indicate that these two preparations contained equivalent amounts of "binding-competent" peptide product (Figure 3D). Second, because half-maximal binding occurred at the same probe concentration for these two derivatives, the data indicate that p79-238 and p79-314 have indistinguishable apparent affinities for the same GT3-bx probe (Figure 3D). Third, the binding curves for p79-238 and p79-314 with the GT2-bx probe confirm not only that both peptide derivatives have lower apparent affinities for the GT2-bx than for the GT3-bx motif but that p79-314 has a much lower apparent affinity for the GT2-bx than does p79-238 (Figure 3D). This observation confirms the conclusion that sequences C-terminal of residue 238 in the N-terminal domain strongly enhance the capacity of this domain for selective binding among the GT-box motifs.

Figure 2D shows the data obtained for initial mapping of the C-terminal DNA binding domain of GT-2. As reported earlier (Dehesh et al., 1990, 1992), the C-terminal half of the GT-2 protein (p404-737) has the highest capacity for complex formation with the GT2-bx and a lower capacity for complex formation with the GT1-bx and GT3-bx probes (Figure 2D). Although removal of 55 amino acids from the N-terminal end of p404-737 caused an apparent partial reduction in complex formation with the GT2-bx probe in the data presented in Figure 2D (central panel, p459-737), it did not strongly alter this complex formation in other repeats of this experiment (data not shown). In contrast, deletion from the C-terminal end of the p459-737 construct to amino acids 646 and 611 resulted in an unexpected apparent increase in complex formation with the GT2-bx (p459-646 and p459-611). Removal of the next 35 amino acids completely abolished detectable DNA binding (at center in Figure 2D, p459-576), even at a fivefold longer exposure time (data not shown).

When GT1-bx and GT3-bx probes were tested for binding to the GT-2 deletion derivatives, a quantitatively different pattern from that with the GT2-bx was observed. Where complexes were detected with the GT1-bx and GT3-bx probes (that is, for peptides p459-646 and p459-611), the levels were lower than for the GT2-bx probe (Figure 2D). These data indicate that these two shorter peptide segments contain information involved in discrimination between the three DNA target sequences. Polypeptides longer than deletion point 646 did not have detectable complex-forming activity with GT1-bx and GT3-bx probes under these conditions, even with a fivefold longer autoradiographic exposure time (data not shown).

These results suggested that this absence of detectable binding for the three longer peptides (p404-737, p459-737, and p459-681) with the GT1-bx and GT3-bx probes was not due simply to the apparently lower general complex-forming capacity observed for these peptides with the GT2-bx probe (at center in Figure 2D). These data suggested therefore that the sequence between amino acid 646 and 681 contains information for enhancement of selective binding among the three closely related GT-box motifs by the C-terminal half of GT-2, because removal of the sequence led to an apparent decrease of this selectivity.

To examine this possibility further, we performed quantitative binding assays at constant protein concentration and increasing DNA probe concentration (Figure 3). Increasing concentrations of the GT2-bx probe resulted in increasing amounts of retarded complex for both p459-737 and p459-611 peptide segments up to saturation at ~2.5 nM, with little difference between the binding curves of the two derivatives (Figures 3E and 3G). Half-maximal binding occurred at ~0.6 nM GT2-bx.
Figure 2. Mapping of Polypeptide Segments in the N- and C-Terminal Domains of GT-2 Involved in Differential Binding to the Closely Related GT-Box Target Sites.
In contrast, increasing concentrations of the GT3-bx probe did
not result in binding saturation for either peptide segment up
to the maximum of 8 nM tested, and p459-611 formed much
greater amounts of retarded complex than did p459-737 at all
probe concentrations (Figures 3F and 3G). Half-maximal bind-
ing occurred at ~2.5 nM GT3-bx for p459-611, whereas binding
was barely detectable for p459-737 at an 8-nM probe concentra-
tion and did not approach the half-maximal value in the range
tested (Figure 3G).

These data lead to several conclusions similar to those de-
derived for the N-terminal domain. First, the indistinguishable
saturation binding curves for p459-611 and p459-737 with the
GT2-bx probe indicate that these two preparations contain
equivalent amounts of binding-competent peptide product with
similar apparent affinities for the same GT2-bx probe (Figure
3G). The reason for the apparent difference in complex for-
mation between p459-611 and p459-737 observed earlier (Figure 2D) but not here is unclear but could reflect the higher
protein concentrations and lower probe-to-protein ratios used
in the former experiments. Second, the binding curves for p459-611 and p459-737 with the GT3-bx probe confirm not only
that both derivatives have lower apparent affinities for the GT3-bx
than for the GT2-bx motif but that p459-737 has a much lower
apparent affinity for the GT3-bx than does p459-611 (Figure
3G). This observation confirms the conclusion that sequences
C-terminal of residue 611 strongly enhance the capacity of the
C-terminal domain to discriminate among the GT-box motifs.

The combined data in Figures 2 and 3 suggest that the infor-
mation for enhanced binding selectivity resides between
residues 261 and 314 in the N-terminal domain and between
646 and 681 in the C-terminal domain. Interestingly, both of
these sequences overlap with a stretch of acidic amino acids
(Figure 2A), but there is little if any overall sequence similarity
between the two regions (Dehesh et al., 1992).

Reduction in Binding Selectivity by Short GT-2
Derivatives Is Not Due to Nonspecific
DNA Binding

Given the increase in complex-forming capacity associated
with the decrease in sequence selectivity by the deletion-
derived shorter peptide segments (Figures 3D and 3G), we
asked whether these peptide segments still bind only to GT-
box family motifs or more broadly to random DNA sequences.
To approach this question, two short polypeptide segments
were selected from each half of the GT-2 protein (Figure 4A)
and synthesized in vitro. Both wild-type and mutated GT boxes
were used as probes in the EMSA (Figure 4C). The mutated
GT boxes were made such that the paired G nucleotides were
changed to paired C nucleotides because the paired G nucleo-
tides in the GT boxes have been shown to be crucial to the
binding of the GT-2 protein (Dehesh et al., 1990, 1992). In ad-
in, these substitutions do not alter the overall nucleotide
composition of the double-stranded DNA probes. For each wild-
type probe, one peptide segment from the N-terminal domain
(p79-238) and another from the C-terminal domain (p459-646)
were chosen as positive controls for effective DNA binding.
All of the polypeptide segments failed to bind detectably to
any of the mutated GT-box probes in which the GG nucleo-
tides were changed to CC (Figure 4B), even for extended
autoradiographic exposure times (data not shown). Therefore,
the data indicate that sequence-selective recognition of GT-
box family members relative to other DNA motifs is embed-
ded in the core DNA binding regions of the GT-2 protein.

Figure 2. (continued).

(A) Schematic diagram of GT-2 protein structure (top) and a series of N- and C-terminal deletions within each half of the molecule (bottom). Proline-
and glutamine-rich regions (P/Q), basic regions (Basic), and acidic regions (Acidic) are shown above, and previously predicted trihelix regions
(helix-loop-helix-loop-helix [HLHLH] and corresponding stippled columns) are shown below the GT-2 protein structure. Numbers indicate amino
acid positions.

(B) In vitro-synthesized GT-2 deletion derivatives. 35S-methionine-labeled proteins were synthesized by in vitro transcription/translation, sepa-
rated by electrophoresis on 16% SDS-polyacrylamide gels, and visualized by autoradiography. The relative positions of prestained molecular
mass markers are indicated in kilodaltons. The templates for in vitro transcription/translation were pPO8 plasmids carrying the subregions begin-
ning and ending at the residue positions indicated in (A). At left is the N-terminal domain; at right is the C-terminal domain.

(C) Gel mobility shift assay of the interaction of N-terminal GT-2 deletion derivatives with GT-1 box (left), GT-2 box (center), and GT-3 box (right)
DNA target sequences.

(D) Gel mobility shift assay of the interaction of C-terminal domain GT-2 deletion derivatives with GT-1 box (left), GT-2 box (center), and GT-3 box (right) DNA target sequences.

(E) GT-box probes used in gel mobility shift assays. Core sequences are boxed. DNA probes containing a single copy of the relevant GT-box
target sequence were used in the gel shift assays as indicated. One femtomole of each of the in vitro-synthesized polypeptides was incubated
with 0.05 ng (0.25 nM) of the 3' end-labeled monomer GT-box probe, as indicated, for 30 min at room temperature and electrophoresed on 8%
polyacrylamide gels in 0.25 x Tris-borate-EDTA buffer.
Figure 3. Quantitative DNA Binding Analysis Demonstrates Enhanced Reciprocal Target Sequence Selectivity by Extended DNA Binding Domains.

A diagram of GT-2 is shown in (A). (B) to (G) show the amount of DNA-protein complex formed at a constant protein concentration and increasing DNA probe concentration by using gel mobility shift assays.
Definition of Minimal DNA Binding Domains within the N-Terminal and C-Terminal Halves of GT-2

To define the boundaries of the minimal DNA binding domains of GT-2 more precisely, we analyzed the binding properties of an additional series of truncated proteins by using the EMSA. Figure 5A shows the coordinates of this series. The N-terminal boundary of the N-terminal DNA binding domain had been tentatively set at residue 79 in the above-mentioned studies (Figure 2C). Removal of the next 19 amino acids had no effect on complex-forming capacity with the GT3-bx (Figure 5B, p98-238), whereas deletion of another 16 residues into the helix 1 region abolished this activity (p114-314). Thus, at this level of resolution, these deletion mutants place residue 98 at or near the N-terminal boundary of this DNA binding region.

Based on this information, deletions from the C-terminal end of this region were produced, as indicated in Figure 5A. Strong complex-forming capacity with wild-type GT3-bx, but not the mutated mGT3-bx, remained down to residue 184 (data not shown), verifying the absence of nondiscriminatory binding of these peptide segments to DNA. Therefore, the minimal N-terminal region of GT-2 that confers GT-box recognition resides in a stretch of 90 amino acids (98 to 187).

The initial deletion analysis of the C-terminal domain of GT-2 above was based on a 5’ boundary at amino acid position 459 (Figure 2). Removal of the next 30 residues from the N-terminal end of this region to amino acid 489 did not strongly alter the complex-forming capacity with GT2-bx (Figures 5C and 6). In contrast, removal of an additional 14 residues from the N-terminal end of this region to amino acid 503, thus deleting part of the trihelix region, eliminated binding (Figure 2D). Thereafter, a series of C-terminal derivatives with residue 489 at the N terminus were generated, as depicted in Figure 5A. This series of closely spaced deletions was made near residue 576 based on the absence of binding activity in p459-576 (Figure 2). The addition of the next three or six amino acids C-terminal of residue 576 restored a considerable degree of complex-forming capacity with the GT2-bx probe (Figure 5C, p489-579 and p489-582). The addition of the next three glycine residues (p489-585) resulted in a yet greater complex-forming capacity, comparable with the level of the control polypeptide p489-611. None of the peptides bound detectably to the mutated mGT2-bx probe (Figure 5C), even at long autoradiographic exposure times (data not shown), verifying the absence of nondiscriminatory binding to DNA. Thus, the minimal C-terminal region of GT-2 that confers GT-box recognition resides in a stretch of 90 amino acids (489 to 579).

Secondary structure predictions for the N-terminal minimal DNA binding region made using the Chou-Fasman (Chou and Fasman, 1978) or PHD neural network (Rost and Sander, 1993a, 1993b) procedures show that this region is predicted to form a helix-loop-helix-loop-helix-helix structure (Figure 7A). The prediction by the PHD procedure for the original trihelix region closely matches that of the previous prediction by the Chou-Fasman algorithm (Dehesh et al., 1992), as does the prediction of a fourth α-helix in the extended region. It is especially noteworthy that the C-terminal border of helix 4 matches precisely with the border of the functionally defined GT-box binding region (Figures 5 and 7).

The structural prediction obtained for the C-terminal minimal DNA binding region by using the PHD procedure mostly matched that of the previous Chou-Fasman prediction in the trihelix region (Dehesh et al., 1992), except for a disagreement regarding helix 2, which is predicted as a weak β-sheet by the PHD procedure (Figure 7B). In addition, the sequence after helix 3 is predicted as a loop followed by β-sheet by both algorithms (Figure 7B). It is noteworthy that the border of the β-sheet (residue 579) in the C-terminal domain precisely

Figure 3. (continued).

(A) Schematic diagram of the GT-2 protein structure (top) and deletions of N- and C-terminal domains of GT-2 tested (bottom). Numbers indicate the amino acid positions. Regions of the polypeptide are as described in the legend to Figure 2A.

(B) p79-314 and p79-238 with GT2-bx.

(C) p79-314 and p79-238 with GT3-bx.

(D) Phosphorlmager quantitation of the amount of retarded complex in (B) and (C).

(E) p459-737 and p459-611 with GT2-bx.

(F) p459-737 and p459-611 with GT3-bx.

(G) Quantitation by Phosphorlmager of the amount of retarded complex in (E) and (F).

Each binding assay contained a constant amount (0.2 fmol) of in vitro-produced polypeptide and increasing concentrations (indicated by triangular rams) of GT2-bx or GT3-bx probes, as shown in (B), (E), and (F), and are as follows: lanes 1 and 6, 0.1 x 10^{-9} M; lanes 2 and 7, 0.3 x 10^{-9} M; lanes 3 and 8, 0.9 x 10^{-9} M; lanes 4 and 9, 2.7 x 10^{-9} M; lanes 5 and 10, 8.1 x 10^{-9} M. The concentrations of the GT2-bx probe in (G) are as follows: lanes 1 and 6, 0.3 x 10^{-8} M; lanes 2 and 7, 0.9 x 10^{-9} M; lanes 3 and 8, 2.7 x 10^{-9} M; lanes 4 and 9, 8.1 x 10^{-9} M; lanes 5 and 10, 2.43 x 10^{-8} M. The same dried gels were subjected to autoradiographic analysis in (B), (C), (E), and (F). Quantitation of the retarded bands was done on the same dried gels by using a Phosphorlmager, and the results are shown in (D) and (G).
matches the border of the functionally defined GT-box binding activity (Figures 5 and 7). Thus, within the limits of accuracy afforded by these predictive programs, the two minimal DNA binding domains in the N- and C-terminal halves of GT-2 appear to differ structurally, as well as in amino acid sequence, at their C-terminal ends.

Proline Substitutions Abolish DNA Binding in the C-Terminal Domain

The necessity of helix 4 for activity of the N-terminal minimal DNA binding domain and of β-sheet 4 for activity of the C-terminal DNA binding domain has been demonstrated by deletion analysis (Figure 5). To test further the involvement of the putative helices in the original trihelix structure of the C-terminal domain for DNA binding, site-directed mutagenesis was used to change amino acid Leu-498 to Pro-498 in helix 1, Ser-523 to Pro-523 in helix 2, and Asn-544 to Pro-544 in helix 3 (Figures 6A and 7C). The selection of these amino acids was based on the observation that they are all strictly conserved between the N- and C-terminal DNA binding domains. These mutations all strongly impaired or abolished detectable protein–DNA interaction (Figure 6B).

The proline-substituted C-terminal motif also was subjected to secondary structure prediction analysis (Figure 7C). The
Figure 5. Fine Mapping of Twin Minimal DNA Binding Domains in the N- and C-Terminal Halves of GT-2.

(A) Schematic diagram of GT-2 protein structure (top) and various deletion derivatives in the N- and C-terminal domains of the GT-2 protein (bottom). Numbers indicate amino acid positions. Regions of the polypeptide are as given in the legend to Figure 2A.

(B) Gel mobility shift assay of the interaction of N-terminal domain GT-2 deletion derivatives synthesized in vitro with the monomeric GT3-bx and mutated GT3-bx (mGT3-bx) DNA target sequences. The GG-to-CC dinucleotide change in the core motif of the mutated GT-box probe is underlined. The templates for the p79-314 and p114-314 in vitro transcription/translation products were generated by digesting to completion, with the restriction enzymes indicated in Figure 2, with pPO9 vectors carrying truncated GT-2, either from amino acids 79 to 521 or from amino acids 114 to 737, respectively. The rest of the templates were generated by PCR, as described in Methods.

(C) Gel mobility shift assay of the interaction of C-terminal domain GT-2 deletion derivatives synthesized in vitro with the monomeric GT2-bx and mutated GT2-bx (mGT2-bx) DNA target sequences. The GG-to-CC dinucleotide change in the core motif of the mutated GT-box probe is underlined. All of the templates for in vitro transcription/translation were generated by PCR reaction, as described in Methods.

One femtomole of each of the in vitro-synthesized polypeptides was incubated with 0.08 ng (0.4 nM) of the 3' end-labeled monomeric GT-box probe, as indicated, for 30 min at room temperature and electrophoresed on 6% (B) or 8% (C) polyacrylamide gels in 0.25 x Tris-borate-EDTA buffer.
Figure 6. Proline Substitutions Disrupt the DNA Binding Activity of the C-Terminal Binding Domain of GT-2.

(A) Schematic diagram of the structure of the GT-2 protein (top) and amino acid sequences (bottom) of the previously predicted (Dehesh et al., 1992) N- and C-terminal trihelix helix-loop-helix-loop-helix (HLHLH) regions. Vertical bars indicate identical residues. Boxes indicate α-helices. The amino acid residues substituted with proline are indicated by reverse typeface. Regions of the polypeptide labeled PQ, Basic, and Acidic are given in the legend to Figure 2A.

(B) Gel mobility shift assay of the interaction of wild type (W) and mutated GT-2 fragments bearing proline substitutions (P) with the GT2-bx DNA target sequence. P1, P2, and P3 indicate the presence of the proline substitution in helix 1, helix 2, or helix 3, respectively. The template for in vitro transcription/translation of p459-646W was the pPO9 plasmid carrying this sequence, as described in the legend to Figure 2B. The remaining templates for in vitro transcription/translation of p489-646W and the three proline-substituted mutants were generated as PCR products by using the appropriate primers on plasmids carrying the wild-type or mutated sequences, as described in Methods. The monomeric GT2-bx probe is indicated at left.

change of Leu-498 to Pro-498 in helix 1 and Asn-544 to Pro-544 in helix 3 disrupted the helix structures predicted by both PHD and Chou–Fasman analyses. Similarly, the introduction of Pro-523 for Ser-523 in helix 2 disrupted the helix predicted by the Chou–Fasman algorithm and reduced the size of the β-sheet predicted by the PHD procedure by one amino acid (Figure 7C).

DISCUSSION

Transcriptional Activation

A large array of plant factors that bind in sequence-specific fashion to various DNA elements has been reported, and an increasing number of these factors have been cloned (Katagiri and Chua, 1992; Katagiri et al., 1992; Li et al., 1993; Ramachandran et al., 1994; Terzaghi and Cashmore, 1995). However, evidence that these factors are directly involved in transcriptional regulation of target genes is frequently circumstantial. The strongest such evidence is often that the recombinant factor binds to cis-acting promoter elements that have been defined by functional assay as being necessary and/or sufficient for expression in vivo but does not bind to mutated derivatives that are not active in the functional assay (Katagiri and Chua, 1992). Examples include members of the large superfamily of basic leucine zipper (bZIP) factors, such as the G-box binding factors (GBFs; Izawa et al., 1993; Li et al., 1993; Foster et al., 1994), the homeodomain-ZIP (HD-ZIP) factors (Schena et al., 1993; Sessa et al., 1993), GT-1 (Lam et al., 1989; Lam and Chua, 1990; Gilmartin et al., 1992; Persisic and Lam, 1992; Hiratsuka et al., 1994; Lam, 1995), the d(AT)-rich element binding factors ATBP-1 (Tjaden and Coruzzi, 1994).
and PF-1 (Nieto-Sotelo and Quail, 1994; Nieto-Sotelo et al., 1994a, 1994b), and factor 3AF1 (Lam et al., 1990). On the other hand, an increasing number of DNA binding proteins have been shown by transfection assays to be capable of activating transcription in vivo from promoters carrying specific target elements. Examples include the Myb-like Cl/Pl factors (Goff et al., 1991; Roth et al., 1991), the basic helix-loop-helix-containing R/B factors (Goff et al., 1990, 1992; Ludwig et al., 1990; Roth et al., 1991), and the bZIP factors Opaque2 (Lohmer et al., 1991; Schmidt et al., 1992; Ueda et al., 1992; Yunes et al., 1994), GBF1 (Schindler et al., 1992), TAF-1 (Oeda et al., 1991), PosF21 (Aeschbacher et al., 1991), TGA1a (Neuhaus et al., 1994), and RITA-1 (Izawa et al., 1994). In addition, TGA1a has been shown to function as a transcriptional activator, increasing the number of active preinitiation complexes, in cell-free transcription assays (Katagiri et al., 1990; Yamazaki et al., 1990).

The data from cotransfection assays presented here provide direct evidence that rice GT-2 can function as a transcriptional activator in vivo. We have previously shown that recombinant GT-2 binds to functionally defined, positively acting DNA elements in the rice PHYA gene promoter and not to mutated derivatives of these elements (Dehesh et al., 1990, 1992) and that the recombinant protein is immunochemically related to a protein of indistinguishable size and DNA binding properties in nuclear extracts (Dehesh et al., 1992). In addition, in situ immunocytochemical data indicate that all

**Figure 7.** Secondary Structure Predictions for the Functionally Defined Twin Minimal DNA Binding Domains in the N- and C-Terminal Halves of the GT-2 Protein.

(A) N-terminal domain.

(B) C-terminal domain.

(C) Proline-substituted C-terminal domain.

The two prediction methods used were the PHD procedure, as described in Methods, and the Chou–Fasman algorithm, as described by Dehesh et al. (1992). Residue coordinates are indicated over each segment. Predicted α-helices (Helix) or mutated α-helices (mHelix) are boxed, and predicted β-sheet (Sheet) or mutated β-sheet (mSheet) are indicated by dots under the sequence. Substituted prolines are indicated by reverse typeface.
detectable GT-2 is localized to the nucleus in rice seedlings, and transfection experiments have identified twin, functionally active, bipartite NLSs in the polypeptide, one in each trihelix region (Dehesh et al., 1995). The GT-2 protein is expressed constitutively in rice seedlings, and this is consistent with earlier promoter analysis indicating that the GT-box target sequences function as constitutively active, positive elements in the rice PHYA promoter (Dehesh et al., 1992, 1995). Taken together, the available data strongly support the conclusion that rice GT-2 functions as a constitutive transcriptional activator driving PHYA expression in rice. Further implementation of the functional assay used here should permit definition of the molecular domains responsible for this activity (Goff et al., 1991, 1992; Gill and Tjian, 1992; Luther et al., 1993; Gill et al., 1994; Tjian and Maniatis, 1994).

### DNA Binding Domains

Previously, we have shown that recombinant rice GT-2 contains two DNA binding activities, one in each half of the polypeptide, capable of discriminating between the closely related GT-box sequence motifs in the PHYA promoter (Dehesh et al., 1992). Here, we defined the minimal DNA binding domains in each half of GT-2 that facilitate GT-box family motif recognition and provide partial discrimination among the family members. In addition, we identified a separate polypeptide segment in each half that apparently confers the capacity for higher resolution discrimination among the closely related target sites. These data suggest therefore that the minimal binding domains contain intrinsic information that provides a considerable degree of sequence selectivity among GT-box family members but that the additional "modifier" sequence further increases this selectivity (Figure 8).

Based on extensive sequence similarity between the 75–amino acid trihelix segments in each half of GT-2, we proposed previously that these segments were likely to be involved in GT-box binding (Dehesh et al., 1992). The data presented here show that the minimal DNA binding domains do indeed encompass the trihelix regions but extend an additional 15 amino acids on the C-terminal side of each region. This additional contiguous sequence is predicted to form an extra α-helix in the N-terminal domain (suggesting a putative "tetrahelix" structure) and a β-sheet structure in the C-terminal domain (a putative "trihelix sheet"). A similar requirement for additional contiguous sequence on the C-terminal side of the originally defined trihelix region for successful DNA binding has been shown recently for the GT-1 factor from Arabidopsis (Hiratsuka et al., 1994) and tobacco (Lam, 1995).

Although the level of resolution in the deletion analyses thus far reported does not permit precise comparison, there is broad agreement as to the extent of the minimal binding domains among the different factors. The fourth helix identified here in the GT-2 N-terminal domain corresponds closely to that pointed out by Lam (1995) for the GT-1 minimal domain. On the other hand, there is little or no sequence similarity in this extended region between GT-1 and GT-2, or even between the two minimal DNA binding domains of GT-2, suggesting a possible contribution of the region to target site selectivity. The disruption of DNA binding competence by substitution of the helix-breaker proline in each of the three helices of the trihelix region here and in Arabidopsis GT-1 (Hiratsuka et al., 1994) suggests that maintenance of each of these putative helices is necessary for binding activity. Taken together, these data indicate that similar structural motifs formed by the tetrahelix (trihelix/sheet) domains of the GT-1 and GT-2 family of factors are responsible for sequence-specific contact with their respective DNA target sites.

### Target Sequence Selectivity Enhancement

The two additional "modifier" sequences within GT-2 that appear to enhance target sequence selectivity do not exhibit any striking sequence or structural similarities to each other, although both include acidic domains. Whether these polypeptide segments are directly involved in DNA contact or indirectly impose structural constraints that modulate the binding selectivity of the minimal DNA binding domains remains to be determined. It is uncertain whether GT-1 factors also have such a modifier segment, because all target–sequence selectivity tests reported thus far showing preference for the GT-1 box motif either were performed with the full-length protein or did not systematically examine deletion derivatives for binding selectivity toward a spectrum of GT boxes (Gilmartin et al., 1992; Hiratsuka et al., 1994; Lam, 1995). Evidence has been presented that GT-1 binds to DNA as a dimer (Hiratsuka et al., 1994) or a tetramer (Lam, 1995) via a multimerization domain near the C terminus, but no comparable information on GT-2 quaternary structure is available.

The rapid accumulation of sequence information on eukaryotic transcription factors has led to the recognition of several
broad classes of structural domains involved in DNA recognition (Johnson and McKnight, 1989; Mitchell and Tjian, 1989; Churchill and Travers, 1991; Harrison, 1991; Pabo and Sauer, 1992). The absence of striking sequence homology between extant proteins in the data bases and the GT-1 and GT-2 factors indicates that these plant proteins may constitute a novel class of transcription factor. On the other hand, the possibility of structural similarities to the DNA binding domains of factors such as the homeodomain proteins remains (Dehesh et al., 1992; Gehring et al., 1994a, 1994b). There is increasing interest in the molecular basis for the capacity of different members within a given family of DNA binding proteins to discriminate with high resolution between closely related DNA binding sites (Izawa et al., 1993; Sessa et al., 1993; Foster et al., 1994; Gehring et al., 1994a, 1994b; Ramachandran et al., 1994; Czerny and Busslinger, 1995; Kim et al., 1995). The GT-1 and GT-2 factors, and particularly the twin DNA binding domains of GT-2, offer an attractive model system with which to explore the mechanism involved in this discrimination between binding sites.

METHODS

Construction of Plasmids for Transactivation Analysis

For the galactose 4 (GAL4) positive control system, the ß-glucuronidase (GUS) coding region and the nopaline synthase (nos) terminator were cloned into pUC19 as a XbaI-EcoRI fragment from pBl101.2 (Clontech, Palo Alto, CA). The GUS gene was driven by a chimeric promoter with the GAL4 binding site fused upstream of either a truncated cauliflower mosaic virus (CaMV) 35S promoter or a truncated rice phytochrome A (PHYA) promoter. The truncated CaMV 35S promoter from positions -59 to +11 was amplified by polymerase chain reaction (PCR) with primers derived from pBl121 (Clontech); however, SaI and BamHI sites were incorporated at the 5' and 3' ends, respectively, for insertion into the homologous sites upstream of the GUS gene. The truncated rice PHYA promoter region from -184 to +111 (Dehes et al., 1992) was also amplified by PCR with primers incorporating XbaI and XmaI sites upstream of the CaMV 35S promoter. The truncated rice PHYA promoter from -188 to +111 was generated by PCR with incorporated BamHI (blunted) and SaI sites at the 5' and 3' ends and cloned into SaI (blunted) and XmaI sites upstream of the GUS gene. A 35S promoter truncated at position -59, as described above, and a rice PHYA promoter truncated at position -188 were used as control reporters. The rice PHYA promoter region from positions -186 to +111 was generated by PCR with introduced XbaI and SaI sites at the 5' and 3' ends for insertion into homologous sites upstream of the GUS gene.

The full-length GT-2 protein was expressed from the CaMV 35S promoter and tobacco etch virus translational enhancer leader sequence from RTL2-GUS (Carrington and Freed, 1990) to achieve high levels of expression in electroporated protoplasts. The GUS coding region was released as an NcoI-BamHI fragment from plasmid pRTL2-GUS. A PCR product (nucleotides 1 to 383 of GT-2) with introduced NcoI and BamHI sites at the 5' and 3' ends and an excised BamHI-SnaBI fragment (nucleotides 384 to 2241 of GT-2) were then cloned into NcoI and BamHI (blunted) sites downstream of the CaMV 35S promoter. A plasmid without the GT-2 insertion was used as a negative control.

For all electroporation experiments, an internal control plasmid containing the luciferase (LUC) gene driven by the CaMV 35S promoter also was included. A PCR product (nucleotides 1 to 126) with introduced NcoI and XbaI sites at the 5' and 3' ends and an excised XbaI-BamHI fragment encoding the remaining open reading frame were cloned into NcoI and BamHI sites of pRTL2 (Carrington and Freed, 1990).

Transactivation Analysis

Suspension cultures of Nicotiana tabacum (line XD) were maintained at 26°C in TXD medium (Howard et al., 1992). The preparation of protoplasts was done according to Howard et al. (1992), and the protoplasts were resuspended in electroporation buffer at a final density of 4 × 10^6/mL. After heat shock for 5 min at 44°C, 0.3 mL of protoplast suspension was mixed with DNA (7.5 µg of activator plasmid, 4 µg of reporter plasmid, 0.2 µg of internal control plasmid, and 15 µg of sheared calf thymus DNA) and incubated at room temperature for 10 min. The 0.15 mL of 24% (w/v) polyethylene glycol 8000 in TXD medium containing 0.4 M mannitol and 90 mM MgCl2 was added, and the mixture was incubated for 10 min at room temperature and then on ice for 10 min. Electroporation was done with a 0.2-cm cuvette (Bio-Rad) and a gene pulser (Bio-Rad) with three pulses at 10-sec intervals at 200 Ω, 25 µF, and 300 V. After recovering on ice for 10 min, the mixture was diluted with 8 mL of TXD medium containing 0.4 M mannitol (TXDM), and the pelleted protoplasts were resuspended in 1 mL of TXDM and cultured for 26 to 30 hr at 25°C in the dark. The harvested protoplasts were resuspended in 200 µL of extraction buffer (100 mM potassium...
Generation of Templates for in Vitro Transcription/Translation

For C-terminal deletion analysis of the N-terminal domain in Figures 2 to 4, the templates for in vitro transcription and translation were generated by digestion to completion of the pPO9 vectors carrying either full-length GT-2 or truncated GT-2 (amino acids 79 to 521 or amino acids 114 to 737, respectively) with the restriction enzymes indicated. For C-terminal deletion analysis of the C-terminal domain in Figures 2 to 4, the templates for in vitro transcription and translation were the pPO9 plasmids carrying the subregions beginning or ending at the deletion points, as indicated in the figures.

For mapping of minimal DNA binding domains and proline substitution analysis, some of the GT-2 subregions indicated in Figures 5 and 6 were generated by PCR on either wild-type or mutated templates, as described above, with oligonucleotide primers beginning or ending at the deletion endpoints indicated in the figures. In addition, the 5' primers also contained a modified T7 promoter (5'-GGAGGAGCTCTAGAGGTTCGATGG-3') upstream of the GT-2 sequence to facilitate subsequent transcription, as described by Lam (1996). All PCRs were conducted at 95°C for 1 min, 58°C for 45 sec, and 72°C for 1.5 min for 30 cycles.

In Vitro Transcription/Translation

One microgram of pPO9 plasmid or 2.0 µg of PCR-generated template was transcribed/translated in the presence of 35S-methionine by using the TNT T7-coupled reticulocyte lysate system (Promega). Three microliters of a 25-µL reaction was loaded on each lane of a 16% gel, as described below. The product was analyzed by SDS-PAGE followed by autoradiography, and the excised band was used to estimate the synthetic protein concentration. The counts per minute of the excised band were converted into picomoles by considering counting efficiency, radioactive decay, and the number of methionine residues present in the particular polypeptide.

Prediction of Secondary Structure

The secondary structure prediction analysis of GT-2 was performed using either the PHD (Rost and Sander, 1993a, 1993b) or the Chou-Fasman (Chou and Fasman, 1978) algorithm.

Construction of Plasmids for in Vitro Protein Synthesis

Nucleotide and amino acid coordinates in the descriptions of all constructions are as given in Dehesh et al. (1992). A PCR fragment (nucleotides 1 to 383 of GT-2) with incorporated EcoRI and BamHI sites at the 5' and 3' ends and an excised BamHI-EcoRI fragment (nucleotides 384 to 2957) were cloned into the EcoRI site of pPO9, a modified PET plasmid (Rottmann et al., 1991). To create a plasmid encoding truncated GT-2 from amino acids 79 to 521, an excised BamHI-BgIII fragment (nucleotides 383 to 1619) was cloned into the BamHI site of pPO9. An annealed pair of oligonucleotides (nucleotides 292 to 383) with incorporated BamHI and BgIII sites at the 5' and 3' ends was then cloned into the BamHI site of the above-mentioned plasmid. To create a plasmid encoding truncated GT-2 from amino acids 79 to 737, a PCR product (nucleotides 397 to 1619) with an introduced artificial ATG codon upstream of the GT-2 sequence and incorporated EcoRI and BgIII sites at the 5' and 3' ends and an excised BgIII-EcoRI fragment (nucleotide 1619 to 2957) were cloned into the EcoRI site of pPO9. Plasmids encoding the deletion mutants of the C-terminal domain were constructed by cloning PCR products into the EcoRI site of pPO9. The 5' primer in each case contained an EcoRI site and an artificial ATG codon upstream of the GT-2 sequence starting at amino acid 459. The 3' primers also contained EcoRI sites with deletion product end points, as indicated in Figures 2 to 6. The plasmid p404-737 is the original cDNA clone as described by Dehesh et al. (1990).

In Vitro Site-Directed Proline Substitution Mutagenesis

A PCR product corresponding to amino acids 459 to 646 with EcoRI sites at each end was inserted into pUC118 and used to generate single-stranded template DNA for oligonucleotide-directed mutagenesis (Kunkel, 1985; Vieira and Messing, 1987). Native T7 polymerase (United States Biochemical) was used for the extension reactions. Mutagenic primers complementary to the GT-2 open reading frame were synthesized to introduce proline substitutions at Leu-498 (P1), Ser-523 (P2), and Asn-544 (P3). The corresponding polypeptides were then synthesized in vitro from PCR products by using these plasmids as templates (see below).

Generation of Templates for in Vitro Transcription/Translation

For C-terminal deletion analysis of the N-terminal domain in Figures 2 to 4, the templates for in vitro transcription/translation were generated by digestion to completion of the pPO9 vectors carrying either full-length GT-2 or truncated GT-2 (amino acids 79 to 521 or amino acids 114 to 737, respectively) with the restriction enzymes indicated.

For C-terminal deletion analysis of the C-terminal domain in Figures 2 to 4, the templates for in vitro transcription and translation were the pPO9 plasmids carrying the subregions beginning or ending at the deletion points, as indicated in the figures.

For mapping of minimal DNA binding domains and proline substitution analysis, some of the GT-2 subregions indicated in Figures 5 and 6 were generated by PCR on either wild-type or mutated templates, as described above, with oligonucleotide primers beginning or ending at the deletion endpoints indicated in the figures. In addition, the 5' primers also contained a modified T7 promoter (5'-GGAGGAGCTCTAGAGGTTCGATGG-3') upstream of the GT-2 sequence to facilitate subsequent transcription, as described by Lam (1996). All PCRs were conducted at 95°C for 1 min, 58°C for 45 sec, and 72°C for 1.5 min for 30 cycles.

In Vitro Transcription/Translation

One microgram of pPO9 plasmid or 2.0 µg of PCR-generated template was transcribed/translated in the presence of 35S-methionine by using the TNT T7-coupled reticulocyte lysate system (Promega). Three microliters of a 25-µL reaction was loaded on each lane of a 16% gel, as described below. The product was analyzed by SDS-PAGE followed by autoradiography, and the excised band was used to estimate the synthetic protein concentration. The counts per minute of the excised band were converted into picomoles by considering counting efficiency, radioactive decay, and the number of methionine residues present in the particular polypeptide.

Gel Mobility Shift Assay

Monomeric wild-type and mutated GT-box probes were excised with XhoI and BamHI from pBluescript SK– plasmids carrying homotrimeric GT boxes with intervening restriction sites, which were made previously using synthetic oligonucleotides (Dehesh et al., 1992). Preparation of an end-labeled probe with the Klenow fragment of DNA polymerase I and standard gel retardation assays was performed according to Green et al. (1991) by using 0.05 to 0.1 ng probe (10,000 cpm) and 1 fmol of in vitro–synthesized protein. For quantitative assays of differential DNA binding affinity, a constant 0.2 fmol of in vitro transcribed/translated polypeptides and increasing GT-box probe concentrations were used, as described in Figure 3. After electrophoresis, the dried gels were subjected to autoradiography for visual presentation and PhosphorImager (Molecular Dynamics, Sunnyvale, CA) analysis for quantitation of the retarded bands.

Prediction of Secondary Structure

The secondary structure prediction analysis of GT-2 was performed using either the PHD (Rost and Sander, 1993a, 1993b) or the Chou-Fasman (Chou and Fasman, 1978) algorithm.
ACKNOWLEDGMENTS

We thank all members of our laboratory for their helpful criticism and suggestions throughout this work and Ron Wells for preparation and editing of the manuscript. This work was supported by grants from the National Science Foundation (No. MCB9220161) and U.S. Department of Agriculture—Agricultural Research Service Current Research Information Service (No. 5335-21000-006-00D).

Received January 11, 1996; accepted April 10, 1996.

REFERENCES


Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., 1058 The Plant Cell


GT-2: in vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity.
M Ni, K Dehesh, J M Tepperman and P H Quail
*Plant Cell* 1996;8;1041-1059
DOI 10.1105/tpc.8.6.1041

This information is current as of July 8, 2017