Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 7205–7209, September 1990
Botany

Identification of potato nuclear proteins binding to the distal promoter region of the proteinase inhibitor II gene

(wound-inducible gene/tissue-specific expression/DNA-binding proteins/*Solanum tuberosum*)

JOSE J. SÁNCHEZ-SERRANO, HUGO PEÑA-CORTÉS, LOTHAR WILLMITZER, AND SALOMÉ PRAT

Institit für Genbiologische Forschung Berlin GmbH, Ihnestrasse 63, 1000 Berlin 33, Federal Republic of Germany

Communicated by Jeff Schell, June 15, 1990

ABSTRACT **Potato nuclear proteins specifically bind to a DNA sequence at the most 5′ distal region of the promoter of a potato proteinase inhibitor II gene. Binding studies using the electrophoretic mobility-shift assay showed the appearance of two protein–DNA complexes in the presence of both tuber and leaf nuclear protein extracts. Mechanical wounding of the leaves had no effect on the amount of specific protein–DNA complexes formed. DNase I protection analysis and binding to synthetic oligonucleotides identified the sequence 5′-GAGGGTATTTTCGTAA-3′ as the target for the noncooperative binding of two potato nuclear proteins to the upstream element. Methylation interference experiments showed that guanine nucleotides separated by one turn of the DNA helix were in close contact with the proteins. The binding ability of a series of mutated synthetic oligonucleotides further defined the sequence requirements for protein binding, which appeared to contact one side of the DNA helix.**

The potato proteinase inhibitor II (PIn-II) is a small protein with trypsin/chymotrypsin inhibitory activity. It is encoded by a small multigene family that is constitutively expressed in the tubers and young floral buds of a normal greenhouse-grown plant. Mechanical wounding of the plant, however, triggers its transcription in the other aerial organs as well. This activation is not restricted to the damaged organs, but rather a more systemic response occurs, with PIn-II mRNA accumulating in nonwounded organs distant from the wound site (1). Evidence indicating that the plant hormone abscisic acid plays a pivotal role in this systemic induction has recently been obtained (2).

A PIn-II gene has been isolated from potato and its promoter has been used to drive the expression of the bacterial β-glucuronidase reporter gene in transgenic plants (3). The transgenic potato plants exhibit constitutive β-glucuronidase activity in tubers and, upon wounding, the reporter gene is expressed in the leaves as well. These results indicate that the promoter region of a single member of the potato PIn-II family is able to respond to both the environmental and the developmental signals that modulate its expression pattern. The same chimeric gene has been transferred to tobacco plants, where it is expressed only after wounding of the leaves (3, 4).

The correct wound-induced expression pattern of the gene in the heterologous tobacco environment was exploited for a deletion analysis of the potato PIn-II promoter region, in order to search for the cis-acting sequences involved in the wound response. One element required for maximal expression of the fused reporter gene (chloramphenicol acetyltransferase) was located at the most 5′ upstream region of the 1.3-kilobase (kb) PIn-II promoter. Truncated promoters with this upstream element deleted had little or no activity that could be recovered by addition of the strong enhancer of the cauliflower mosaic virus (CaMV) 35S promoter (5).

To gain further insight into the complex regulation of the PIn-II gene in potato, we have scanned the promoter region for sequences that could be the target for nuclear proteins, and compared the DNA-binding characteristics of proteins isolated from tissues where the gene is expressed (tubers and wounded leaves) with those of proteins from tissues where it remains silent (leaves of nonwounded plants).

MATERIALS AND METHODS

Plant Material and RNA Isolation and Analysis. Potato plants were grown under greenhouse conditions (16°C day/10°C night, 50–80% relative humidity, and 14-hr-light photoperiod). Wounding of the plants and isolation and analysis of RNA were performed as described (2).

PIn-II Promoter Subcloning. DNA manipulation was performed according to standard techniques (6). A *Hin*dIII–*Sca* I restriction fragment ranging from −1297 to +33 relative to the transcription start site of the PIn-II gene (3) was digested with the appropriate restriction enzymes, and the resulting fragments, which spanned the whole promoter region, were cloned in a pUC19 vector. The identity and orientation of the cloned restriction fragments was checked by sequencing, thereby extending the already published sequence of this PIn-II promoter to position −1297.

Nuclear Protein Isolation. Some modifications were introduced in the procedure we described for the isolation of nuclear proteins (7). Filtration of the nuclei through a 25-μm-pore sieve cloth was omitted. After nuclear lysis and removal of nuclear debris, protein extracts were eventually passed through a Whatman DE52 column at 0.4 M NH_4^+. The remaining DNA was removed by 1 hr of ultracentrifugation at 100,000 × *g* in a Beckman 45Ti rotor and proteins were subsequently precipitated with $(NH_4)_2SO_4$ (0.37 g/ml), collected at 100,000 × *g* for 15 min, and resuspended in a minimal volume of dialysis buffer.

Electrophoretic Mobility-Shift Assay, DNase I Footprinting, and Methylation Interference. These assays were performed essentially as described (7). Double-stranded poly(dI-dC) was used as nonspecific competitor for the binding reactions (1 μg/μg of nuclear protein). Sonicated salmon sperm DNA was also used as indicated.

Oligodeoxynucleotide Synthesis and Cloning. The oligonucleotides were converted to double-stranded forms by polymerization with the Klenow enzyme, using 5′-GCGCTGT-3′ as primer for the oligonucleotides wt and m1–m10, and 5′-GCGCTCA-3′ for the oligonucleotide m11. The various oligonucleotides were radioactively labeled and were eluted from 6% polyacrylamide gels after electrophoresis. Oligonucleotides aI and maI were previously cloned in pUC19 and end-labeled with the Klenow enzyme.

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

Abbreviations: CaMV, cauliflower mosaic virus; PIn-II, proteinase inhibitor II.

RESULTS

A Prominent Complex Is Formed Between Potato Nuclear Proteins and the Most Distal 5′ Region of the PIn-II Promoter. A 1.3 kb PIn-II promoter confers tuber-constitutive and wound-inducible expression on a fused β-glucuronidase gene in transgenic potato plants. The putative protein–DNA interactions at the promoter region were analyzed by the gel mobility-shift assay. To directly compare DNA-binding activity with the transcriptional activity of the gene, nuclear protein extracts were prepared from tubers and wounded leaves, both of which support active transcription of the PIn-II gene. By using extracts from nuclei isolated at various times after wounding, we hoped to avoid missing any factors with short half-lives that might be produced upon wounding. Extracts from nonwounded leaves served as negative controls or, alternatively, might unveil the presence of repressor-like factors.

Various fragments that together spanned the whole length of the PIn-II promoter region were separately incubated with nuclear protein extracts, and the protein–DNA complexes were separated from the free DNA fragments by gel electrophoresis under nondenaturing conditions (Fig. 1). A prominent protein–DNA interaction was detected on the most 5′ upstream fragment (positions −1282 to −1093), leading to the appearance of two retarded bands. This DNA-binding activity was present in both tubers and leaves, remaining constant upon wounding. The appearance of bands with a retarded mobility was abolished by heating the extracts at 65°C for 15 min or by proteinase K digestion, but not by RNase A digestion, indicating the involvement of proteins in the formation of the complexes (data not shown). Additional protein–DNA complexes were formed with some of the more downstream fragments (fragments II, IV, V, and VI) when the concentration of nonspecific competitor DNA was lowered (Fig. 2 and data not shown). Fragment VI included the putative TATA box, but the detected binding activity was most likely due to protein interaction with a different sequence, since the complex was still formed in the presence of another fragment containing the TATA motif, namely, a truncated CaMV 35S promoter (−90Δ35S; ref. 7). The amount of complex formed was similar in tubers and in leaves, where the activity showed no significant variation upon wounding (not shown). The nuclear protein(s) involved

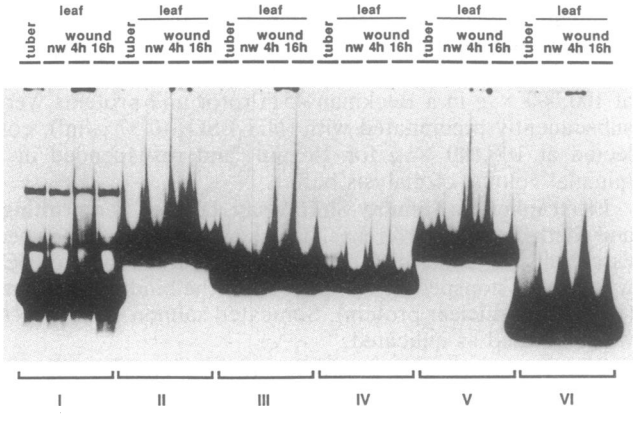

FIG. 1. The various PIn-II promoter fragments (identified by roman numerals at the bottom) were radioactively labeled and incubated with nuclear protein extracts (5 μg) from tubers or from leaves, either nonwounded (nw) or 4 hr and 16 hr after a single wounding, as indicated above the gel slots. One microgram of salmon sperm DNA was included in each incubation mixture. Fragment I, positions −1282 to −1093 relative to the transcription start of the PIn-II mRNA; fragment II, −1092 to −808; fragment III, −807 to −625; fragment IV, −624 to −406; fragment V, −405 to −104; fragment VI, −103 to +33.

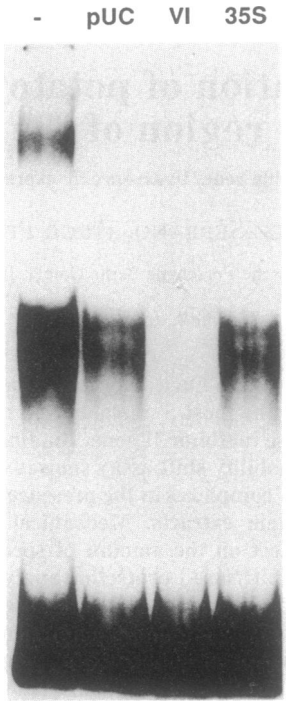

FIG. 2. Fragment VI DNA, including the putative TATA box, was radioactively end-labeled and incubated with a leaf nuclear protein extract. The protein–DNA complexes were separated from the free DNA in a native polyacrylamide gel. For each lane, 0.5 ng of the radioactive fragment was incubated with 5 μg of protein extract and a 200-fold molar excess of nonlabeled competitor DNA as indicated above each slot: −, no competitor; pUC, isolated *Eco*RI/*Hin*dIII-digested pUC19 polylinker fragment; VI, homologous fragment VI DNA; 35S, truncated CaMV 35S promoter (−90Δ35S). The second, slower migrating band shown in the control lane (no competitor) was formed as well with most of the labeled fragments used and disappeared when low amounts of native DNA (100-fold weight excess of salmon sperm DNA) were included in the incubation mixture together with poly(dI-dC) as nonspecific competitor.

appeared to bind preferentially to A+T-rich DNA tracts, since formation of the labeled complex was prevented by an excess of unlabeled DNA fragments including A+T-rich stretches.

The Upstream-Element-Binding Nuclear Proteins Interact with Nucleotides That Are Separated by One Turn of the DNA Helix. We decided to concentrate our work on the protein(s) binding to the upstream element of the PIn-II promoter because a region encompassing its binding site was shown by promoter deletion analysis to be required for maximal expression of the gene (5).

In the first series of experiments, radioactively labeled fragment I DNA was incubated with tuber nuclear protein extract and subjected to mild DNase I digestion. The protein–DNA complexes were subsequently separated and the DNA was isolated separately from each complex. The oligonucleotide mixture was resolved in a denaturing sequencing gel. Weak DNase I protection was detected from nucleotides −1274 to −1255 on the upper strand and between nucleotides −1275 and −1251 on the lower strand (Fig. 3). These "footprints" were not present in the unbound DNA isolated from the gel and thus reflected the specific interaction of one or more proteins with the underlying sequence. Both the faster and the slower moving complexes produced the same DNase I footprint. An identical footprint was obtained when leaf extracts were used (data not shown).

To further establish which nucleotides within the recognition sequence are in direct contact with the protein, the fragment I DNA was partially methylated and incubated with

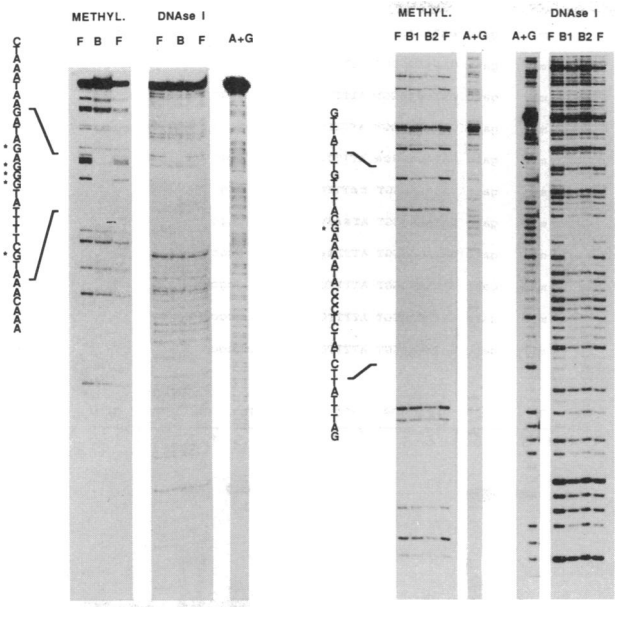

FIG. 3. Isolated fragment I DNA was radioactively labeled in the upper strand (*Left*) or lower strand (*Right*) and incubated with tuber nuclear proteins. The methylated guanine residues that interfered with protein binding were determined by comparing the oligonucleotides produced by piperidine cleavage of the free DNA (lanes F) and the protein-bound DNA (lanes B) eluted from a native gel. These residues are indicated with stars on the sequence given, from −1282 to −1250 in the upper strand (from top to bottom, *Left*). The region protected from the DNase I digestion is indicated on that sequence as well. The oligonucleotides methylated at G^{-1259} in the lower strand (indicated with a star) were reproducibly found to be slightly more abundant in the free DNA fraction. The DNase I footprint is indicated in the lower-strand sequence shown, which goes from −1246 to −1282 (top to bottom, *Right*). In this set of experiments, the DNAs from the slower and faster migrating protein–DNA complexes were isolated separately (lanes B1 and B2, respectively). Products of the Maxam–Gilbert A+G sequencing reaction (8) were included as molecular size references.

tuber nuclear extract. As for the DNase I protection experiments, the DNA was isolated from both the faster and the slower migrating complexes and cleaved at the methylated guanine residues. The resulting oligonucleotides were separated in a sequencing gel; the free DNA, isolated and cleaved in the same manner was used as a control. Fig. 3 shows the result of such an experiment; the fragments corresponding to the methylation of the guanine residues at positions −1270, −1268, −1267, −1266, and −1258 were underrepresented in the bound DNA as compared to the free DNA, thus indicating that methylation at these positions prevented binding of the protein. Remarkably, the methylated guanine nucleotides at positions −1268 and −1258, both of them interfering with protein binding, are exactly one DNA helix turn apart. Formation of the protein–DNA complex was not prevented by methylation of any of the guanine nucleotides on the lower strand. Interestingly, oligonucleotides methylated at G^{-1259} on the lower strand were only slightly underrepresented in the bound DNA fraction, whereas those methylated at the next nucleotide, G^{-1258}, on the upper strand were totally absent from it.

Two Nuclear Proteins Interact with a Single Target Sequence in a Noncooperative Manner. A 38-mer synthetic oligonucleotide corresponding to the sequence where the protein-DNA interaction occurred (positions −1282 to −1245) was cloned and used in subsequent experiments aimed at identifying the target for the nuclear protein binding. As a control, a second oligonucleotide was cloned, containing four base-pair exchanges (residue −1267 interchanged with residue −1264 and residue −1261 with residue −1258) that altered the sequence but maintained the overall G+C content. The oligonucleotide representing the wild-type sequences, but not the mutated one, was able to interact with tuber nuclear proteins, giving rise to two complexes of differing electrophoretic mobilities (Fig. 4). The formation of both complexes was progressively inhibited by including increasing amounts of the wild-type oligonucleotide, whereas addition of the mutated oligonucleotide did not affect the binding of the protein. It is also worth mentioning that protein titration by addition of unlabeled homologous competitor DNA resulted in the simultaneous disappearance of both the faster and the slower migrating complexes, keeping their ratio constant over the different concentrations. Along this line, the binding abilities of both the wild-type oligonucleotide and the original restriction fragment I were checked at various nuclear protein concentrations (Fig. 5). Both DNAs gave rise to two complexes of differing mobilities. The differences in migration of the complexes formed with the fragment as compared to the oligonucleotide might be accounted for by the difference in size of the DNA fragments.

Two Regions of Close Contact with the Nuclear Proteins Can Be Defined in the Target Site. To investigate in more detail the DNA sequence requirements for the protein binding to the PIn-II upstream element, the binding ability of a series of mutant oligonucleotides was tested. In each mutant, two transversions at contiguous positions were introduced. Two regions where mutations prevented protein binding were identified (nucleotides −1270 to −1263 and nucleotides −1258 to −1255) that were separated by a region where mutations did not significantly affect the binding ability of the oligonucleotide (Fig. 6). Interestingly, the mutant oligonucleotide m8 bound the nuclear proteins, indicating that close contact to C^{-1259} is not absolutely required for protein binding, in agreement with the guanine methylation interference data.

Every mutation that interferes with binding prevents the formation of both complexes equally; none of the mutated oligonucleotides tested gave rise to a single (faster or slower) complex. These results, together with the guanine methylation interference, are consistent with the involvement of two

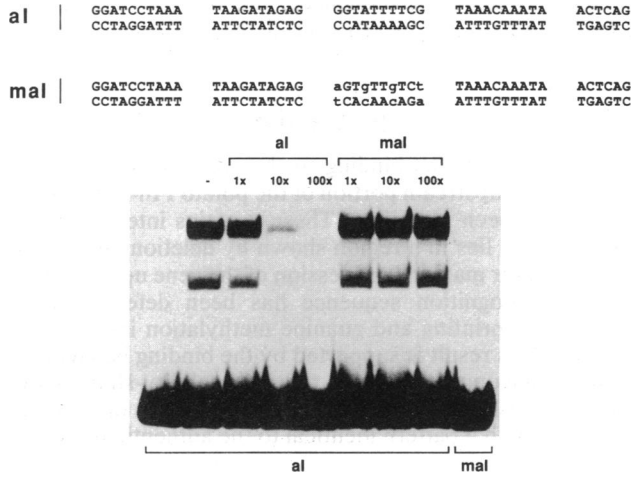

FIG. 4. (*Upper*) Sequences of the cloned oligonucleotides used for the binding experiments: aI, oligonucleotide representing the wild-type sequence; maI, mutant oligonucleotide. Lowercase letters indicate the mutated residues. (*Lower*) The aI oligonucleotide (0.5 ng, aI at bottom) was radioactively labeled and incubated with 5 μg of tuber nuclear protein extract in the presence of no competitor DNA (−) or increasing amounts (1×, 10×, or 100× molar excess) of nonradioactive wild-type (aI at top) or mutant (maI at top) oligonucleotide. As a control, the absence of protein binding to the mutant oligonucleotide is shown (maI at bottom).

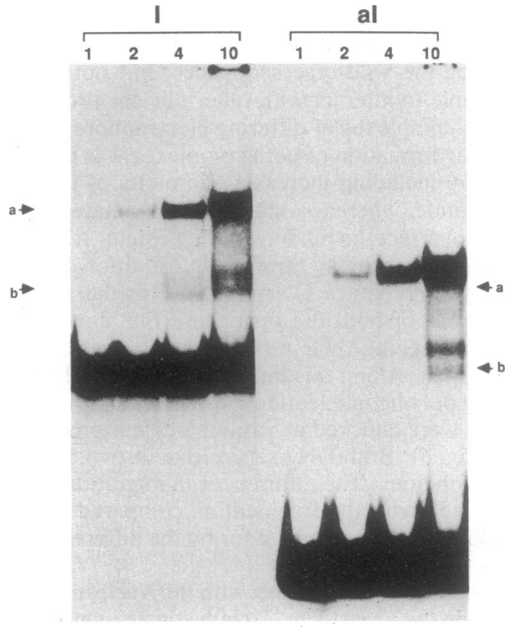

FIG. 5. The isolated fragment I DNA (*Left*) and the aI fragment (*Right*) were radioactively labeled and incubated with 1, 2, 4, or 10 μg of tuber nuclear protein extract, as indicated above the gel slots. Arrowheads indicate the equivalent slower (a) and faster (b) migrating protein–DNA complexes formed. The faster migrating complex can reproducibly be resolved in two separated bands in both cases.

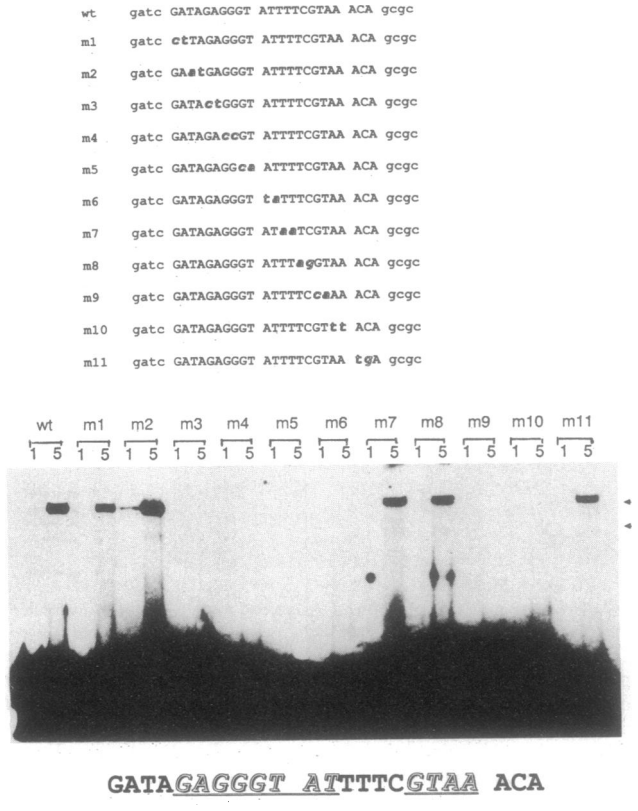

FIG. 6. (*Upper*) Sequences (5′ to 3′) of the oligonucleotides used for binding. Lowercase letters at each end correspond to the nucleotides used to provide common ends to all the oligonucleotides. wt, Wild-type sequence, from position −1274 to −1252 with respect to the transcription start site; m1–m11, oligonucleotides mutated at the two internal residues indicated with italic lowercase letters. (*Lower*) The radioactively labeled oligonucleotides were incubated with either 1 or 5 μg of tuber nuclear protein extract and the protein-bound DNA was separated from the free oligonucleotide in a native polyacrylamide gel. Arrows indicate the positions of the slower and faster (much weaker) protein–DNA complexes. The sequence below the autoradiogram summarizes the nucleotides that appear to establish close contacts with the nuclear proteins (underlined italics) and those where mutations do not affect their binding ability (bold letters).

proteins, either one of which can bind to the same target site on one side of the DNA helix, making contacts to residues one helix turn apart. To confirm this hypothesis, another mutant oligonucleotide was tested that had two nucleotides introduced in the A+T-rich stretch separating the two regions of close protein contact. The actual sequence of this A+T-rich stretch is seemingly unimportant, as slight changes introduced with the transversions are tolerated. The 2-nucleotide separation of the DNA regions in more direct contact with the protein prevented formation of either complex on the oligonucleotide (Fig. 7), thus indicating that the spacing between these sequences, and/or their relative side on the DNA helix, is crucial to permit the establishment of the proper protein–DNA interaction.

DISCUSSION

Two nuclear proteins binding to the same target DNA sequence in the upstream portion of the potato PIn-II promoter region have been identified. These proteins interact with a sequence that lies in a region shown by deletion analysis to be required for maximal expression of the gene upon wounding. The recognition sequence has been determined by DNase I footprinting and guanine methylation interference analyses. This result is supported by the binding behavior of synthetic oligonucleotides in the gel mobility-shift assay, which give rise to two protein–DNA complexes of differing mobilities, with a pattern identical to the authentic promoter fragment.

The differences in binding ability of a series of mutated oligonucleotides enabled us to draw a more refined picture of the protein–DNA interaction. The proteins bind to a site spanning two DNA helix turns, and they make close contacts with one side of the helix. Mutations affecting the binding ability of the oligonucleotide prevented the formation of the faster and the slower migrating protein–DNA complexes equally. This observation, together with the absence of any recognizable palindromic sequence, points to two different proteins binding independently to the target sequence. Consistent with this, no cooperative binding was apparent when the concentration of the nuclear proteins in the binding mixture was increased. However, we cannot exclude the possibility that the two complexes are due to the action of a protein and a proteolytic fragment of it that retains the DNA-binding domain. The data also do not allow us to exclude the possibility that the larger complex could be due to a second protein interacting with the one able to bind to DNA, but preliminary results from protein purification through gel filtration (not shown) suggest the presence of two proteins that are able to bind to DNA independently.

We found no similarity between the recognition sequence we have defined and any of the published targets for the binding of nuclear proteins. Although it contains a 7-nucleotide stretch, GGAGATA (nucleotides −1267 to −1273 of the PIn-II promoter), present in the so-called GATA motif implicated in the binding of activating-sequence factor 2 (ASF2) to the CaMV 35S promoter (nucleotides −85 to −91) (9), the characteristics of the respective interactions are quite different in regard to the guanine methylation interference analyses. This indicates that ASF2 and the factor(s) we describe are in fact different proteins.

The proteins binding to the upstream element of the PIn-II promoter are present both in tissues where the gene is actively expressed and in tissues where it is silent. This

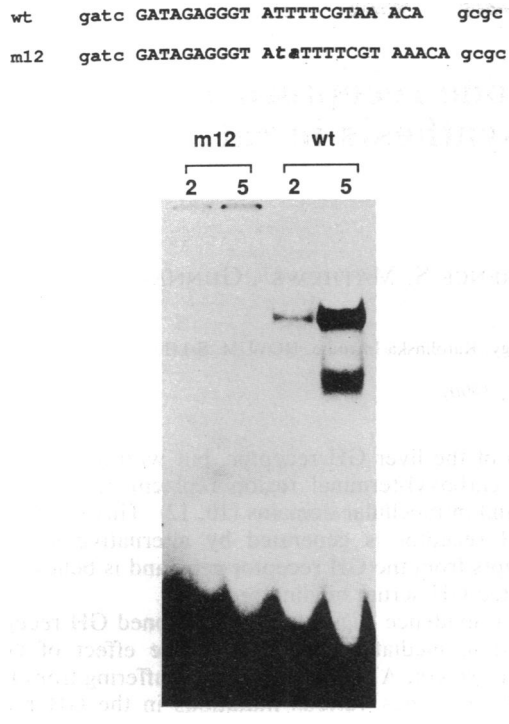

FIG. 7. The wild-type (wt) sequence corresponding to the proteinase inhibitor upstream element was altered by introducing two nucleotides (lowercase italics in the m12 sequence) to separate the regions contacting the nuclear proteins. Both oligonucleotides were radioactively labeled and incubated with either 2 or 5 μg of tuber nuclear protein extract, as indicated. The protein–DNA complexes were separated in a native polyacrylamide gel.

situation is consistent with the data obtained from the promoter deletion analysis, which revealed a constitutive enhancer function located in a 700-bp fragment in the 5′ upstream region of the PIn-II promoter, whose activity was dependent on wound-inducible elements placed further downstream (5).

Several binding sites have also been detected that appear to interact with a protein with a high preference for A+T-rich DNA stretches. The sites are scattered throughout the PIn-II promoter, and no correlation has been found with the regions important for the wound inducibility of the gene as defined by the promoter deletion analysis (5). Several independent isolations of this binding activity have shown no reproducible fluctuation upon wounding. These data, together with the location of the binding sites, suggest that these sites play no role in the wound-mediated activation of the gene.

Recently, a wound-modulated DNA-binding activity has been described for the PIn-II promoter (10). The target site of this activity is located in fragment IV according to our nomenclature. We did not detect any protein binding to this fragment that displayed a reproducible difference upon wounding other than the above-mentioned proteins that bind to A+T-rich DNA, but subtle changes may have escaped our detection due to the size of the fragment used. Its putative target sequence (−156 to −165) lies in a region that has, however, been shown to be largely dispensable for the wound-inducible response of the gene (5), which might preclude it from being actually responsible for the activation of the gene upon wounding.

We detected no nuclear protein whose binding activity would precede the changes in the transcriptional activity of the gene. One can argue that these factors are required only at very precise moments, when the gene is switched on by tuberization stimuli or wound-mediating signals such as abscisic acid. These factors would only be necessary to trigger the tissue-specific, or environmentally regulated, expression of the gene by releasing it from a transcriptionally inactive conformation, perhaps due to the proteins interacting with the A+T-rich binding domains. In this case, the putative tissue-specific factors would only be present in a very short time gap and might result in concentrations below the level of detection. Alternatively, the DNA-binding ability of these putative induced factors might be absolutely dependent on a modification of the protein or on the presence of a cofactor affecting the DNA-binding domain of the protein, which could be lost upon nuclear isolation. Sequence-specific DNA binding might require the interactions of the transcription factors with proteins that are not eluted from the DNA at the salt concentrations used.

Are the same specific factors responsible for the wound-induced and tuber-constitutive expression of the PIn-II gene? If so, one can envision these factors, or any of the steps involved in the signal transduction, being constitutively active in tubers but inducible upon wounding in leaves. This issue has a large bearing on the procedures used to isolate these putative factors, the identities of which promise to shed new light on this exciting plant gene system.

The excellent technical assistance of S. Amati throughout this work is gratefully acknowledged. We thank R. Breitfeld, B. Küsgen, and C. Grams for maintaining the greenhouse-grown plants and M. Stratmann for the synthetic oligonucleotides. We thank especially J. Bergstein for the photographic work and S. Jackson, P. Morris, and C. Scheidereit for critical reading of the manuscript. H.P.-C. is a recipient of a Deutscher Akademischer Austauschdienst predoctoral fellowship. This work was supported by a grant from the Bundesministerium für Forschung and Technologie.

1. Sánchez-Serrano, J. J., Amati, S., Keil, M., Peña-Cortés, H., Prat, S., Recknagel, C. & Willmitzer, L. (1990) *Molecular and Cellular Biology of the Potato*, eds. Vayda, M. E. & Park, W. (CAB International, Wallingford, UK), in press.
2. Peña-Cortés, H., Sánchez-Serrano, J. J., Mertens, R., Willmitzer, L. & Prat, S. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 9851–9855.
3. Keil, M., Sánchez-Serrano, J. J. & Willmitzer, L. (1989) *EMBO J.* **8,** 1323–1330.
4. Sánchez-Serrano, J. J., Keil, M., O'Connor, A., Schell, J. & Willmitzer, L. (1987) *EMBO J.* **6,** 303–306.
5. Keil, M., Sánchez-Serrano, J. J., Schell, J. & Willmitzer, L. (1990) *Plant Cell* **2,** 61–70.
6. Ausubel, F., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G. & Struhl, K. (1987) *Current Protocols in Molecular Biology* (Wiley/Greene, Brooklyn, NY).
7. Prat, S., Willmitzer, L. & Sánchez-Serrano, J. J. (1989) *Mol. Gen. Genet.* **217,** 209–214.
8. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
9. Lam, E. & Chua, N. H. (1989) *Plant Cell* **1,** 1147–1156.
10. Palm, C. J., Costa, M. A., An, G. & Ryan, C. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 603–607.