The Plant Cell, Vol. 6, 287–301, February 1994

PF1: An A-T Hook-Containing DNA Binding Protein from Rice That Interacts with a Functionally Defined d(AT)-Rich Element in the Oat Phytochrome A3 Gene Promoter

Jorge Nieto-Sotelo,[1] Audrey Ichida,[2] and Peter H. Quail [3]

University of California-Berkeley/United States Department of Agriculture, Plant Gene Expression Center, 800 Buchanan Street, Albany, California 94710

Phytochrome-imposed down-regulation of the expression of its own phytochrome A gene (*PHYA*) is one of the fastest light-induced effects on transcription reported in plants to date. Functional analysis of the oat *PHYA3* promoter in a transfection assay has revealed two positive elements, PE1 and PE3, that function synergistically to support high levels of transcription in the absence of light. We have isolated a rice cDNA clone (pR4) encoding a DNA binding protein that binds to the AT-rich PE1 element. We tested the selectivity of the pR4-encoded DNA binding activity using linker substitution mutations of PE1 that are known to disrupt positive expression supported by the *PHYA3* promoter in vivo. Binding to these linker substitution mutants was one to two orders of magnitude less than to the native PE1 element. Because this is the behavior expected of positive factor 1 (PF1), the presumptive nuclear transcription factor that acts in *trans* at the PE1 element in vivo, the data support the conclusion that the protein encoded by pR4 is in fact rice PF1. The PF1 polypeptide encoded by pR4 is 213 amino acids long and contains four repeats of the A-T hook DNA binding motif found in high-mobility group I-Y (HMG I-Y) proteins. In addition, PF1 contains an 11-amino acid-long hydrophobic region characteristic of HMG I proteins, its N-terminal region shows strong similarities to a pea H1 histone sequence and a short peptide sequence from wheat HMGa, and it shows a high degree of similarity along its entire length to the HMG Y-like protein encoded by a soybean cDNA, SB16. In vitro footprinting and quantitative gel shift analyses showed that PF1 binds preferentially to the PE1 element but also at lower affinity to two other AT-rich regions upstream of PE1. This feature is consistent with the binding characteristics of HMG I-Y proteins that are known to bind to most runs of six or more AT base pairs. Taken together, the properties of PF1 suggest that it belongs to a newly described family of nuclear proteins containing both histone H1 domains and A-T hook DNA binding domains.

INTRODUCTION

Plants have several photoreceptors, such as the blue light receptor, the UV receptor, and phytochrome, that regulate growth and development (Kendrick and Kronenberg, 1986). Phytochrome is by far the best characterized of these receptors. Phytochrome is a family of photoreceptors designated as phytochromes A, B, C, D, and E (Quail, 1991), each consisting of a protein with a covalently linked chromophore. This holoprotein has the characteristic property of existing in two photointerconvertible forms: Pr, which absorbs maximally in the red region (666 nm) and Pfr, which absorbs maximally in the far-red region (730 nm) of the spectrum. After absorption of a photon, Pr is converted to Pfr and vice versa. This property of photoconvertibility allows the molecule to perform its regulatory function. Phytochrome is synthesized de novo as Pr, and photoconversion to Pfr can induce dramatic changes in the patterns of gene expression (Gilmartin et al., 1990; Thompson and White, 1991). These changes eventually result in new patterns of development, such as seed germination, deetiolation, stem elongation, and flowering. Immediate conversion back to the Pr form prevents the induction of many of these responses (Kendrick and Kronenberg, 1986).

One of the most rapid light-mediated effects on transcription reported to date is the down-regulation by phytochrome of the transcription of its own phytochrome A gene (*PHYA*) (Lissemore and Quail, 1988). This phenomenon has been used as a model to understand how phytochrome controls gene expression. The development of a transient expression assay in rice seedlings using oat *PHYA3* gene promoter–chloramphenicol acetyltransferase fusion constructs has made possible the functional analysis of this promoter (Bruce et al., 1989). Using this approach, it has been demonstrated that three positive elements, designated PE1 (−367 to −346 bp), PE2 (−635

[1] Current address: Instituto de Biotecnología, Universidad Nacional Autónoma de México, apdo. postal 510-3, Cuernavaca, Mor. 62271, México.

[2] Current address: Department of Biology, University of California at San Diego, La Jolla, CA 92093-0116.

[3] To whom correspondence should be addressed.

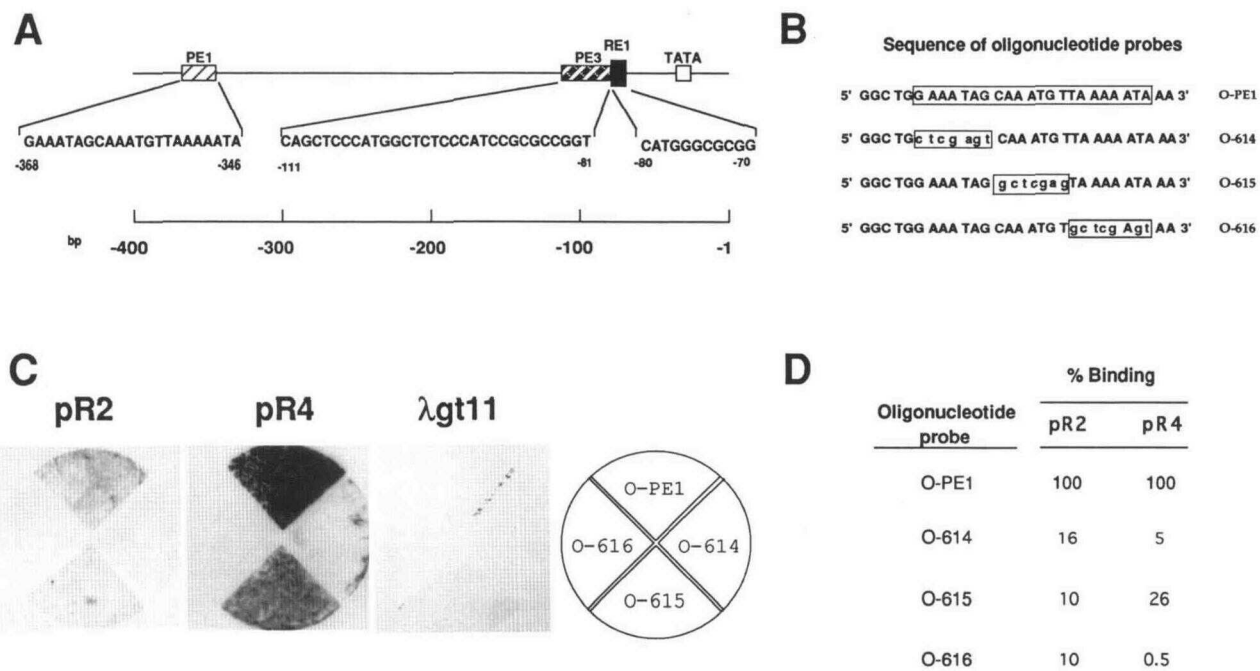
to −489 bp), and PE3 (−111 to −81 bp), are involved in high-level expression in cells that have received either no light or a terminal far-red light pulse (low-Pfr cells) (Bruce and Quail, 1990). PE1 and PE2 are functionally redundant, and PE3 interacts synergistically with either PE2 or PE1 for activity. In addition, one negative element, designated RE1, has been shown to repress transcription in cells that have received a terminal red light pulse (high-Pfr cells) (Bruce et al., 1991). In the rice *PHYA* promoter, which behaves similarly to the oat *PHYA3* promoter, no element with sequence similarity to the oat PE1 motif is present. Instead, an ensemble of GT box motifs has been shown to be involved in positive regulation of this promoter (Dehesh et al., 1992).

To characterize protein–DNA interactions involved in the expression of the rice *PHYA* gene, we have previously isolated and characterized a cDNA clone encoding a protein, designated GT-2, that binds in sequence-specific fashion to the GT box ensemble (Dehesh et al., 1990, 1992). In this work, we report the isolation and characterization of the DNA binding activity of cDNA-encoded rice proteins that bind to the PE1 element of the oat *PHYA3* gene promoter.

RESULTS

Isolation of Rice cDNA Clones Encoding Proteins That Bind to the AT-Rich DNA Element PE1

A λgt11 cDNA expression library prepared from poly(A)$^+$ RNA isolated from etiolated rice seedlings was screened with a double-stranded synthetic oligonucleotide (O-PE1) containing the sequence of the PE1 region of the oat *phyA3* promoter, as shown in Figures 1A and 1B. Screening a total of 750,000 recombinant phage yielded two clones expressing proteins that bound selectively to the ^{32}P-labeled PE1 oligonucleotides



Oligonucleotide probe	% Binding pR2	% Binding pR4
O-PE1	100	100
O-614	16	5
O-615	10	26
O-616	10	0.5

Figure 1. Filter Binding Assay of Recombinant pR2 and pR4 to Synthetic Oligonucleotides.

(A) Structural organization of the oat *PHYA3* gene promoter indicating the locations and sequences of the functionally defined *cis*-acting elements PE1, PE3, and RE1 (Bruce et al., 1991).

(B) Sequences of oligonucleotide probes. Probe O-PE1 contains the sequence of PE1. Probes O-614, O-615, and O-616 have the sequences of linker substitution mutants of the PE1 region that disrupt positive expression of a *PHYA3* gene promoter–reporter chimeric construct in gene transfer experiments (Bruce et al., 1991). Boxes delimit the PE1 region in O-PE1 and the base substitutions in O-614, O-615, and O-616. Lowercase letters indicate changes made in the sequence of the PE1 region in O-614, O-615, and O-616.

(C) Autoradiograms of the filter binding assay made to test the binding of pR2 and pR4 clones and a λgt11 nonrecombinant phage to synthetic oligonucleotides whose sequences are indicated in **(B)**. *E. coli* Y1090 cells were infected with pR2 or pR4 phages or a λgt11 nonrecombinant phage, plated on Petri dishes, and grown to confluence. Plaque-lift filters were prepared from each plate, cut into four quarters, and further processed by filter binding assay, as described in Methods. The diagram at the right shows the probes utilized for hybridization to each quarter filter. Specific activities (10^8 cpm/μg of DNA) of probes were O-PE1, 1.6; O-614, 1.7; O-615, 1.9; O-616, 2.1.

(D) Quantitation of the amount of radiolabeled probe bound to filters containing recombinant proteins expressed by pR2 and pR4 clones. Sequences of oligonucleotide probes O-PE1, O-614, O-615, and O-616 utilized in this assay are shown in **(B)**. Quantitation of the binding activity is expressed as the percentage of the amount of radioactivity bound to each filter divided by the amount of radioactivity bound to the O-PE1 filter.

after tertiary screening. The two positive clones were designated pR2 and pR4. To confirm the binding selectivity of the positive clones, a filter binding assay was performed with the O-PE1 oligonucleotide and with oligonucleotides carrying linker substitution mutations in PE1, designated O-614, O-615, and O-616, the sequences of which are shown in Figure 1B. Because these linker substitution mutations caused the loss of expression of a reporter gene in a rice seedling transient expression assay (Bruce et al., 1991), we reasoned that the binding to DNA of factors involved in transcription of the oat *PHYA3* gene in rice seedlings should be affected as well by these mutations.

Figures 1C and 1D show that binding of O-614, O-615, and O-616 oligonucleotides to the recombinant protein expressed by pR4 was between 0.5% and 26% of the level of binding of O-PE1. This result demonstrates that pR4 encodes a DNA binding protein, designated R4, with the characteristics expected of positive factor 1 (PF1), the nuclear factor that presumably interacts with the PE1 region in vivo to facilitate enhanced transcription. Binding of all mutant oligonucleotides to R2, the pR2-encoded recombinant protein, ranged between 10% and 16% of O-PE1. However, the absolute binding activity of O-PE1 to the recombinant protein encoded by pR2 was one order of magnitude lower than to the one encoded by pR4 (data not shown). Furthermore, sequence analyses of pR2 and pR4 demonstrated that they were unrelated and that the pR4-encoded protein was a cognate of pO2, an oat cDNA independently isolated for its binding activity toward O-PE1 (J. Nieto-Sotelo and P.H. Quail, unpublished data). Therefore, we concentrated our attention on the pR4 clone in subsequent studies.

Characterization of the pR4-Encoded PF1 Protein

The 1-kb pR4 cDNA insert was subcloned into the expression vector pPO-9. Total extracts from isopropyl-β-D-thiogalactopyranoside (IPTG)-induced or noninduced *Escherichia coli* cells transformed with the recombinant plasmid in both orientations were made. Gel blot analysis of these extracts, shown in Figure 2, demonstrates that the subclone containing the cDNA insert in the sense orientation produced a protein with an apparent molecular mass of 36 kD on SDS-PAGE and that this protein bound to the O-PE1 DNA probe. No binding to the O-PE1 probe was observed with either extracts obtained from IPTG-induced bacteria expressing the pR4 cDNA in antisense orientation or from uninduced bacteria expressing the pR4 cDNA in sense orientation.

RNA gel blot analysis of poly(A)$^+$ RNA isolated from rice seedlings using the pR4 cDNA insert as a probe detected an mRNA of 1 kb, suggesting that the cDNA insert is close to full length (data not shown). To assess this possibility directly, we raised polyclonal antibodies against the PF1 polypeptide encoded by pR4. These polyclonal antibodies recognized two proteins in rice nuclear extracts with apparent molecular masses of 34 and 26 kD. As shown in Figure 3A, the larger polypeptide recognized by these antibodies was

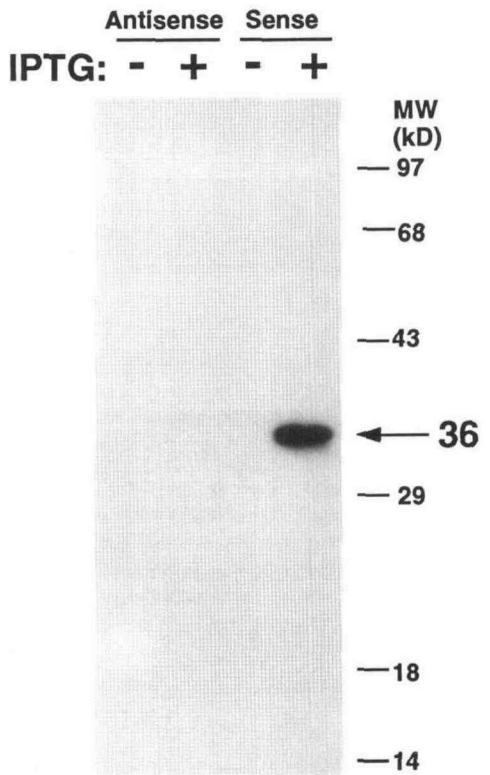


Figure 2. Protein Gel Blot Analysis of Recombinant PF1 Encoded by the pR4 cDNA.

Total protein extracts from IPTG-induced or noninduced cultures of sense or antisense orientation constructs in vector pPO-9 transformed in *E. coli* BL21(DE3) cells were separated by SDS-PAGE and transferred onto nitrocellulose. Filters were denatured/renatured and incubated with a radiolabeled O-PE1 concatenated probe (see Figure 1B), as described in Methods. The arrow shows the position of the protein bound to O-PE1. Molecular mass markers are given at right in kilodaltons.

indistinguishable in size from in vitro–synthesized recombinant PF1. Figure 3B shows an autoradiogram of the ^{35}S-labeled in vitro–synthesized PF1 whose signal superimposes that of the PF1 detected by immunoblotting. The nature of the 26-kD polypeptide in nuclear extracts is uncertain, but it may represent (1) a proteolytic product of PF1 caused during extract preparation, (2) another protein with epitopes recognized by anti-PF1 antibodies, or (3) a mature PF1 polypeptide that has been processed in vivo from a 34-kD form.

The fact that nuclear and in vitro–synthesized PF1 comigrate as 34-kD polypeptides indicates that pR4 encodes a full-length PF1. The size of both nuclear and in vitro–synthesized recombinant PF1 (Figure 3) is smaller by 2 kD than recombinant PF1 expressed from vector pPO-9 (Figure 2). This difference in molecular mass is explained by the fact that the PF1 recombinant protein expressed in the pPO-9 vector (Figure 2) has 20 extra amino acids at the N terminus compared to the in vitro–synthesized product (Figure 3). These 20 extra amino acids are encoded by sequences between the ATG codon and the EcoRI

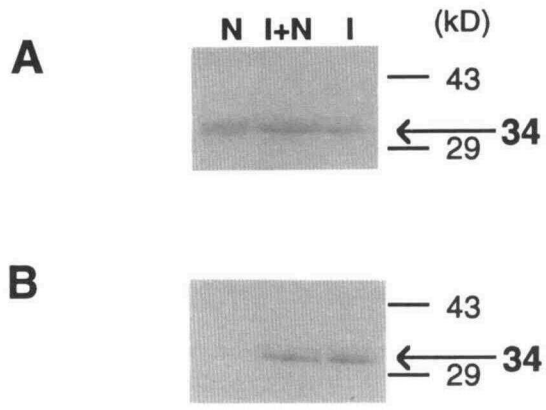


Figure 3. Recombinant PF1 Produced by in Vitro Transcription–Translation and a Rice Nuclear Protein Comigrate in SDS-PAGE.

(A) Immunoblot analysis of protein extracts with a polyclonal antibody raised against purified PF1.

(B) Autoradiogram of the same filter showing ^{35}S-labeled in vitro–translated products using pR4 cDNA cloned in pBluescript SK+ plasmid and T7 RNA polymerase in a coupled in vitro transcription–translation reaction (see Methods).

N, rice nuclear protein extract (25 μg); I+N, mixed in vitro translation products and rice nuclear extract; I, in vitro–translated PF1 (10 μL standard reaction). Arrow shows the positions of comigrating protein bands. Molecular mass markers are given at right in kilodaltons.

cloning site in pPO-9, by the EcoRI adaptor used for subcloning pR4 cDNA, and by the 34 bp upstream of the first ATG codon of pR4 cDNA.

Characterization of Recombinant PF1 DNA Binding Activity

A gel mobility shift assay was performed in 6% polyacrylamide gels with protein extracts from bacteria expressing recombinant PF1 and the 300-bp oat *PHYA3* promoter fragment containing the PE1 region shown in Figure 4 (P-PE1 [positions −415 to −116]). As shown in Figure 5, the assay revealed a single protein–DNA complex (Figure 5, lanes 1, 2, and 3) that was absent in reactions containing protein extracts from bacteria expressing pR4 cDNA in antisense orientation (data not shown). The binding of PF1 to P-PE1 was also compared in this assay to the binding to 300-bp fragments of the same region of the promoter containing the same linker substitution mutations within PE1 as those shown in Figure 1B. Figure 5 shows that complex formation between recombinant PF1 and P-PE1 (lanes 1 to 3) requires lower protein concentration than complex formation with any of the promoter mutations (lanes 4 to 12). These results suggest that binding of PF1 to this 300-bp promoter fragment takes place mostly within the PE1 region, although some interaction may still occur outside PE1 at high protein concentration, because some complex formation is seen with P-615 and P-616 in the lanes containing 5 μg of protein extract (Figure 5, lanes 9 and 12).

In a 10% polyacrylamide gel system, two protein–DNA complexes, one major (B1) and one minor (B2), were resolved with a bacterial extract containing recombinant PF1 and an end-labeled P-PE1 probe, as shown in Figure 6 (lane 1). The molecular basis for these two complexes has not been determined, but one possibility is that the stoichiometry of the two complexes may be different, with B2 perhaps having a lower molar ratio of protein to DNA than the B1 complex. Competition experiments using end-labeled P-PE1 and recombinant PF1 showed that both protein–DNA complexes are strongly reduced in the presence of an excess of nonlabeled oligotrimers of PE1 (OT-PE1) (Figure 6, lanes 2 and 3). By contrast, oligotrimers of the linker substitution mutations of PE1, OT-614, OT-615, and OT-616, did not block formation of either of the two protein–DNA complexes (Figure 6, lanes 4 to 9). Curiously, OT-614 and OT-615 enhanced protein–DNA complex formation, especially the B2 complex (lanes 4 to 7). In contrast, OT-616 did not seem to either block or enhance protein–DNA complex formation (lanes 8 and 9). These results indicate that PF1 has a higher affinity toward P-PE1 than to any of the mutant oligotrimers. At present, we do not understand the reason for the enhancing effect of OT-614 and OT-615.

In vitro footprinting using DNase I was performed to determine the sequence selectivity of recombinant PF1 binding to a 259-bp DNA fragment containing sequences of the oat *PHYA3* promoter region between −415 and −307 bp (Figure 4). Figure 7 shows that recombinant PF1 protects three distinct segments within this promoter fragment in the region −396 to −347 (lanes 4 and 5). Part of this protected region (−347 to −358 and −363 to −368) overlaps with PE1. Two other regions of AT-rich DNA (−374 to −385 and −387 to −396) just

Figure 4. Nucleotide Sequence of the Oat *PHYA3* Gene Promoter in the −415 to −116 Region.

Box represents the PE1 region (see Figure 1A) (Bruce et al., 1991). Underlined bases indicate regions protected by recombinant PF1 in the in vitro footprinting experiment; the results are shown in Figure 7. This sequence was obtained by Hershey et al. (1987).

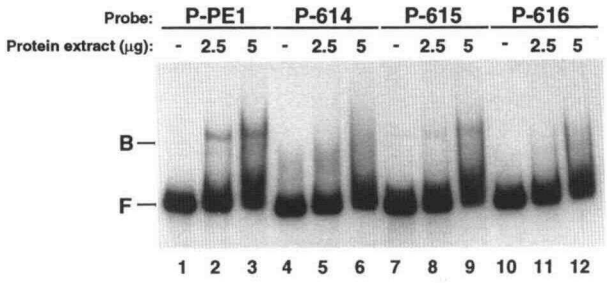


Figure 5. Binding of Recombinant PF1 to a Promoter Fragment Containing PE1 or Linker Substitution Mutants of PE1.

Mobility shift analysis was performed with or without extracts from bacteria overexpressing recombinant PF1 protein and equal amounts of cpm of end-labeled 300-bp probes containing the oat *PHYA3* gene region between −415 bp and −116 bp (P-PE1), as shown in Figure 4, or 300-bp probes of the same promoter region containing linker substitutions within PE1 (P-614, P-615, and P-616), as shown in Figure 1B. Lanes 1, 4, 7, and 10, no protein extract; lanes 2, 5, 8, and 11, 2.5 μg of protein extract; lanes 3, 6, 9 and 12, 5 μg of protein extract. Probes were labeled to approximately the same specific activities (cpm/fmole of DNA): P-PE1, 506; P-614, 505; P-615, 393; P-616, 617. Samples were resolved on a 6% polyacrylamide-1 × Tris-borate-EDTA gel. F, free probe; B, protein bound to DNA complex.

upstream of the PE1 region (Figure 4) are also protected. Protection was not seen with protein extracts prepared from bacteria expressing pR4 in antisense orientation (Figure 7, lanes 7 and 8). Together, the data in Figures 1C, 5, 6, and 7 indicate that PF1 can bind to AT-rich DNA sequences other than PE1 but that the factor has highest affinity for PE1 of the AT-rich motifs in the *PHYA3* promoter.

Sequence Analysis of PF1

The pR4 cDNA insert was sequenced by the dideoxynucleotide chain termination method. Figure 8 shows the nucleotide and predicted amino acid sequence of PF1. The cDNA has a length of 953 bp, and the largest open reading frame is in frame with the λgt11 β-galactosidase gene. Assuming the first in-frame ATG after the β-galactosidase sequence at nucleotide 35 to be the initiator codon, this open reading frame spans 639 nucleotides in length, followed by a stop codon and 311 nucleotides of 3′ noncoding region not showing a poly(A) tail. The amino acid sequence derived from the largest open reading frame contains 213 amino acids with a calculated molecular mass of 21.7 kD. This size differs from that of the apparent molecular mass of 34 kD determined by SDS-PAGE of both the rice nuclear protein and the PF1 recombinant protein made in vitro. The reason for this discrepancy is not clear, but it may be due in part to the structural composition of the protein. PF1 is 24% alanine and 17% proline with a calculated pI of 10.7 due to the abundance of arginine and lysine residues.

Inspection of the predicted protein sequence indicates the absence of certain of the better known DNA binding motifs like the zinc finger, basic-leucine zipper, helix-turn-helix, or other motifs characteristic of eukaryotic transcription factors. Also noticeable is the absence of acidic activation domains arranged in amphipathic α-helices as in GAL4 protein and other transcription factors (Ptashne, 1988). A dot plot analysis of the protein sequence compared to itself revealed four internal repeats of nine amino acids in length present in the C-terminal half of the protein (data not shown). The consensus sequence for these four repeats is K/PRG/ARGRPPK (Figure 8).

A search of the GenBank and EMBL data bases revealed a 47% similarity (30% identity) to histone H1 from pea (Gantt and Key, 1987) when the whole PF1 sequence was used for comparison. Interestingly, none of the similarities resides within any of the four internal repeats of PF1. Instead they seem to be distributed along the entire length of the protein with a peak of conservation in the N-terminal end (data not shown). When a comparison to the data base was made with the sequence of the two identical nanopeptide repeats 1 and 4 of PF1 (amino acid residues 98 to 106 and 192 to 200), a 100% match to this nanopeptide sequence was obtained with two asparagine-rich antigens of unknown function from *Plasmodium falciparum* (L. Schreiber, U. Deutsche, T. Storck, and D. Mueller-Hill, unpublished data, in GenBank and EMBL accession numbers X17485 and X17488). In addition, a match of eight of nine identical amino acids was found (1) with the sequence from

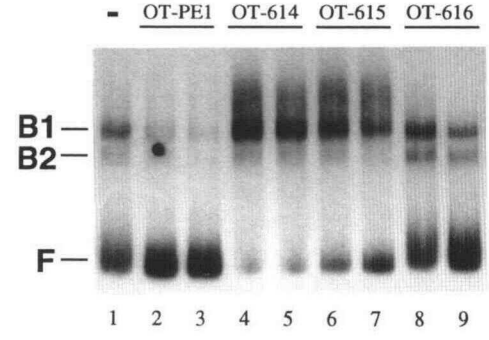


Figure 6. Competitive Binding of Recombinant PF1 to a Labeled PE1-Containing Promoter Fragment and Unlabeled Wild-Type and Mutant Oligotrimers of PE1.

Unlabeled oligotrimers of PE1 (OT-PE1) (lanes 2 and 3), O-614 (OT-614) (lanes 4 and 5), O-615 (OT-615) (lanes 6 and 7), and O-616 (OT-616) (lanes 8 and 9) were added at 350 × molar excess (lanes 2, 4, 6, and 8) or 700 × molar excess (lanes 3, 5, 7, and 9) in addition to an end-labeled 300-bp oat *PHYA3* promoter fragment (−415 to −116 bp) (P-PE1), as shown in Figure 4, and 2 μg of protein extract from bacteria overexpressing PF1. Lane 1 (−), control plus labeled 300-bp promoter fragment and 2 μg of protein extract, without unlabeled oligotrimers. Oligomonomer sequences corresponding to oligotrimers used in this experiment are shown in Figure 1B. Samples were resolved on a 10% polyacrylamide-1 × Tris-borate-EDTA gel. F, free probe; B1 and B2, complexes between protein and DNA.

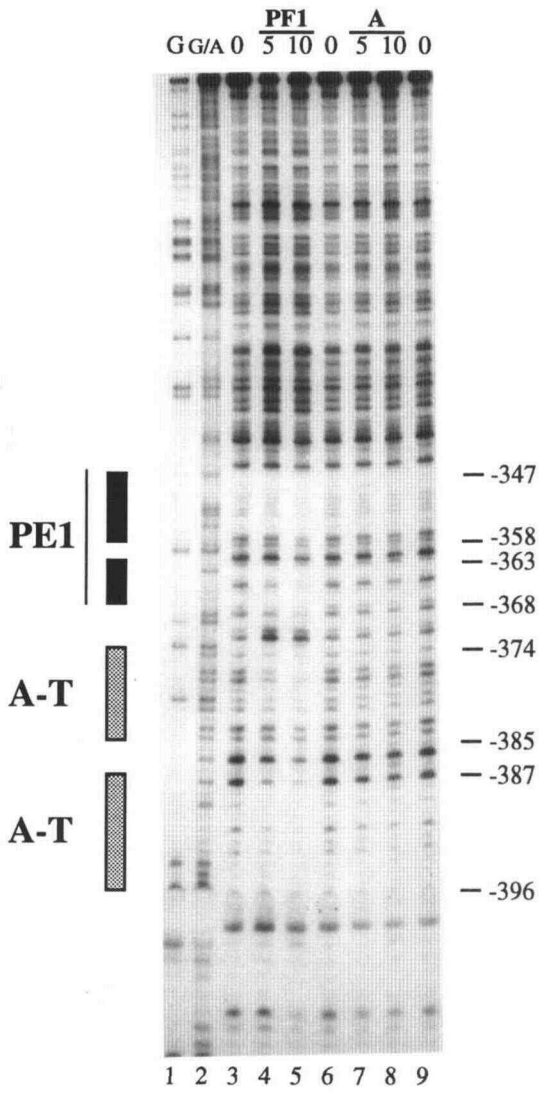


Figure 7. DNAse I Protection Assay Using Recombinant PF1 Overproduced in Bacteria and a Promoter Fragment Containing the PE1 Region.

A DNase I protection assay was performed by using a 259-bp DNA fragment (containing the region of the oat *PHYA3* gene promoter at positions −415 to −307 bp [see Figure 4] plus pBluescript SK+ sequences between SalI and PvuII sites) labeled at the BamHI site. Labeled fragment was incubated with DNase I in the absence of bacterial extract or in the presence of bacterial extract obtained from cells overproducing recombinant PF1 or overexpressing pR4 cDNA in the antisense orientation, as described in Methods. Lanes 1 and 2 (G, G/A), G and G+A Maxam and Gilbert sequencing reactions; lanes 3, 6, and 9, reactions with no bacterial extract added (0); lanes 4 and 5, reactions with 5 and 10 μg of bacterial extract obtained from cells overproducing PF1, respectively (PF1); lanes 7 and 8, reactions with 5 and 10 μg of bacterial extract obtained from cells overexpressing pR4 cDNA in the antisense orientation, respectively (A).

Saccharomyces cerevisiae encoded by the SNF2 protein gene (Laurent et al., 1991) that shows functional interdependence with SNF5 and SNF6 proteins in transcriptional activation, (2) with a developmentally regulated mRNA of *Dictyostelium discoideum* (Shaw et al., 1989), and (3) with the mammalian high-mobility group I-Y (HMG I and HMG Y) proteins (Eckner and Birnstiel, 1989; Johnson et al., 1989).

A more recent search of the GenBank and EMBL data bases revealed a 100% match to the soybean HMG Y-a and HMG Y-b sequences in this region (Laux et al., 1991). This amino acid repeat, present in all HMG I and HMG Y proteins, constitutes their DNA binding domain, and because it binds to dA-dT–rich DNA, it has been termed the A-T hook (Reeves and Nissen, 1990). Figure 9 shows a comparison of the amino acid sequences of rice PF1 reported here to the human and soybean HMG I and HMG Y proteins. The two soybean HMG Y proteins show higher identity to PF1 (54% identity; 66% similarity) than do their mammalian counterparts to PF1 (27% identity; 37% similarity). The soybean HMG Y proteins and rice PF1 have additional N- and C-terminal sequences absent in the mammalian proteins. The alignment also shows three small gaps of 10, 9, and 14 amino acids in length in the soybean HMG Ys relative to the rice PF1 sequence (amino acids 6 to 15, 109 to 117, and 166 to 179 of PF1 sequence, respectively).

```
1    TTCTCTCCCCCTGTTCACACCACGCCTCACTCCCATGGCGACCGAGGAGGACGCGTCGAC   60
1                                 M  A  T  E  E  D  A  S  T          9

61   CATGGCGGCGGCCGAGGCGGACCCCAAGCCCGCCGCCACCCCGTCTTACCCGGAGATGAT  120
10   M  A  A  A  E  A  D  P  K  P  A  A  T  P  S  Y  P  E  M  I     29

121  ACTGGCGGCGATCGAGGCGCTGGACGACAGGAACGGGTCGAACAAGACGGCGATCTCGCA  180
30   L  A  A  A  I  E  A  L  D  D  R  N  G  S  N  K  T  A  I  S  Q  49

181  GCACATCGAGGGGAAGTACGAGGGGCTCCTCCCCCCCGCGCACCCGTCGCTCCTCACAGC  240
50   H  I  E  G  K  Y  E  G  L  L  P  P  A  H  P  S  L  L  T  A     69

241  GCACCTCGCCCGCATGAAGCAGACCGGCGAGCTCGCCTTCTCCAAGAACAACTACTTCCG  300
70   H  L  A  R  M  K  Q  T  G  E  L  A  F  S  K  N  N  Y  F  R     89

301  CGGCGACGACCCGTCCCTCCCGCCCAAGCGCGGCCGCGGCAGGCCGCCCAAGCCCAAGGA  360
90   G  D  D  P  S  L  P  P  K  R  G  R  G  R  P  P  K  P  K  D     109

361  CGCCGCCGCCGCCGCCGCTGCACCGGCGCCCGCGCCCGCGGCGCCGGCGGCCTCCTCGCC  420
110  A  A  A  A  A  A  A  P  A  P  A  P  A  A  P  A  A  S  S  P     129

421  GCGCCCCCGTGGCCGCCCGCCTAAGCCCAAGGACCCCCTCGCCGAGGCGGTCGCGAAGGC  480
130  R  P  R  G  R  P  P  K  P  K  D  P  L  A  E  A  V  A  K  A     149

481  CACCAGTGGCATGCCCCGCGCGCGCGGCCGCCCGCCCAAGAAGGCCAAGGTCGAACAGGA  540
150  T  S  G  M  P  R  A  R  G  R  P  P  K  K  A  K  V  E  Q  E     169

541  AGACCCCATTGGCGCTCCCGCCGCAGCTTCGGCTCCCGCCGCGGCGGCTGAGGCCGCGCC  600
170  D  P  I  G  A  P  A  A  A  S  A  P  A  A  A  A  E  A  A  P     189

601  GCCTGTCAAGCGCGGCCGTGGGCGCCCGCCCAAGGTGAGACCCGCCGCGCCGGTCGGGGA  660
190  P  V  K  R  G  R  G  R  P  P  K  V  R  P  A  A  P  V  G  E     209

661  GCCAGCCGCTGCTTGATTAGCTCCGATTCTCCGCATCGCCGCGATGGCGACGATGGTAAG  720
210  P  A  A  A  *                                                  213

721  CTCCTTTAGGTTGTAACGTAGTCTTCAGATTATGTCTCGTACTGATTGGAGGCGCGGCTC  780

781  GGGTTGCTGCTTATTGTAAGTTAATCGGATGGTGGGAGGTTTAGGTGTGCTGGATGGATC  840

841  ATCCGTTGCTGAGAGTTGTGATTGAATTCGTGTTATCGCGTGCTGATTATTATGGTAATG  900

901  TTGGTGACTTTCGTGTCGTTTGTGTTAATTCTAGATTGGTGGCTGCATTTAGT  953
```

Figure 8. Nucleotide and Derived Amino Acid Sequences of Rice cDNA Clone pR4 Encoding PF1.

Amino acids are numbered starting with the first methionine in frame with the λgt11 β-galactosidase gene and are indicated in the standard one-letter code below the nucleotide sequence. Boxes indicate nonapeptide repeat with similarity to the A-T hook domain. This sequence has been submitted to the EMBL, GenBank, and DDBJ Nucleotide Sequence data bases as accession number L24390.

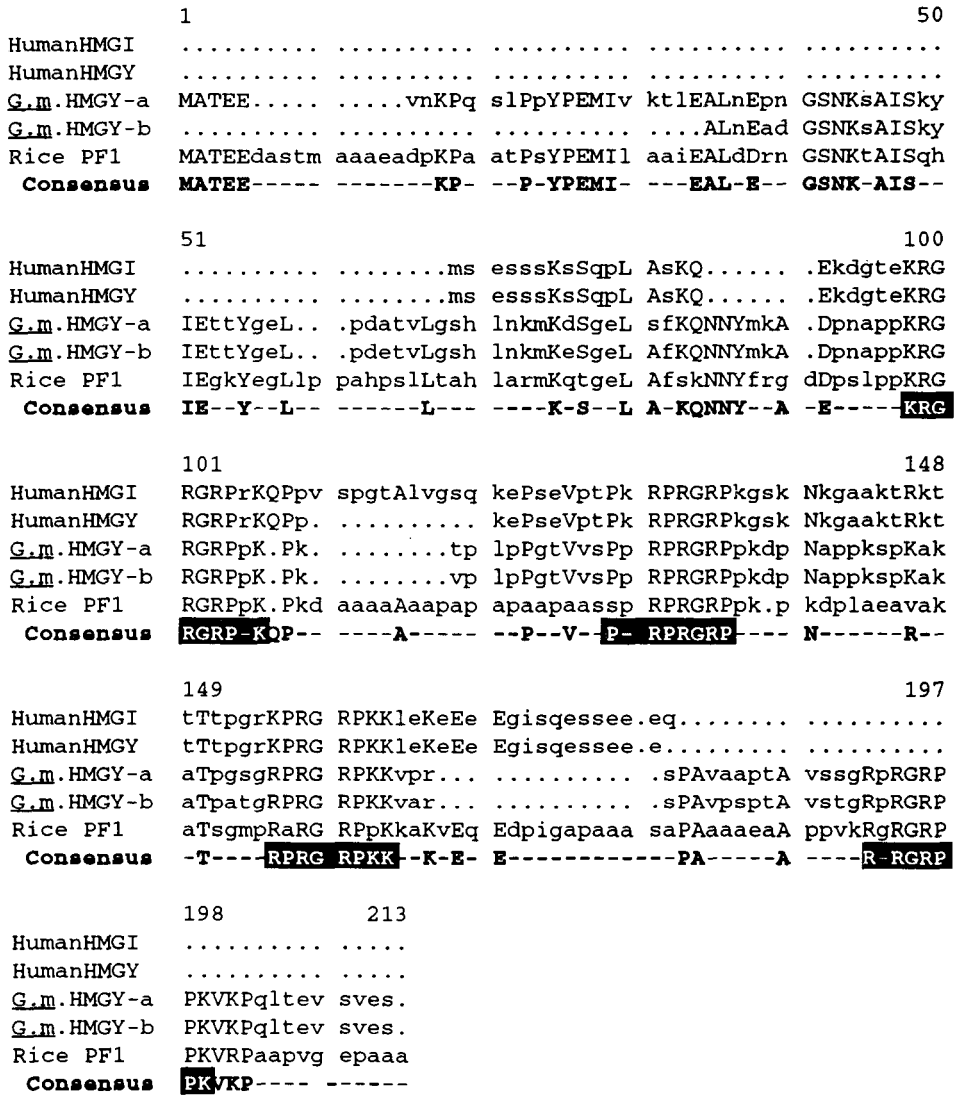


Figure 9. Comparison of Rice PF1 Deduced Amino Acid Sequence to Human and Soybean HMG I and HMG Y Sequences.

The numbers refer to the amino acid positions of rice PF1. Amino acids that match the consensus sequence are indicated in capital letters. Boxes represent the positions of the A-T hook repeats. G.m. is the soybean HMG Y-a or HMG Y-b sequence.

The second gap is a short hydrophobic region characteristically present in mammalian HMG I proteins and absent in HMG Y. The presence of this hydrophobic region in PF1 suggests that it is more HMG I-like.

Rice PF1 also shows a high degree of similarity to a wheat HMGa peptide sequence obtained by Arwood and Spiker (1990) from purified protein. The sequence from amino acid 28 to 48 of rice PF1 shows 90% similarity (81% identity) to the sequenced 21-amino acid-long peptide of the wheat HMGa (Arwood and Spiker, 1990). This N-terminal sequence is found in the region of higher similarity to histone H1 and is more conserved between rice PF1 and wheat HMGa than between wheat HMGa and soybean HMG Y-a (76% similarity; 62% identity).

Two C-Terminal A-T Hooks Are Dispensable for High-Affinity Binding to DNA

Synthetic oligopeptides corresponding to a single A-T hook are capable of binding to DNA (Reeves and Nissen, 1990). The presence of multiple DNA binding domains in PF1 raises the question of how many A-T hooks are necessary for high-affinity binding of this protein to P-PE1. Mobility shift assays were

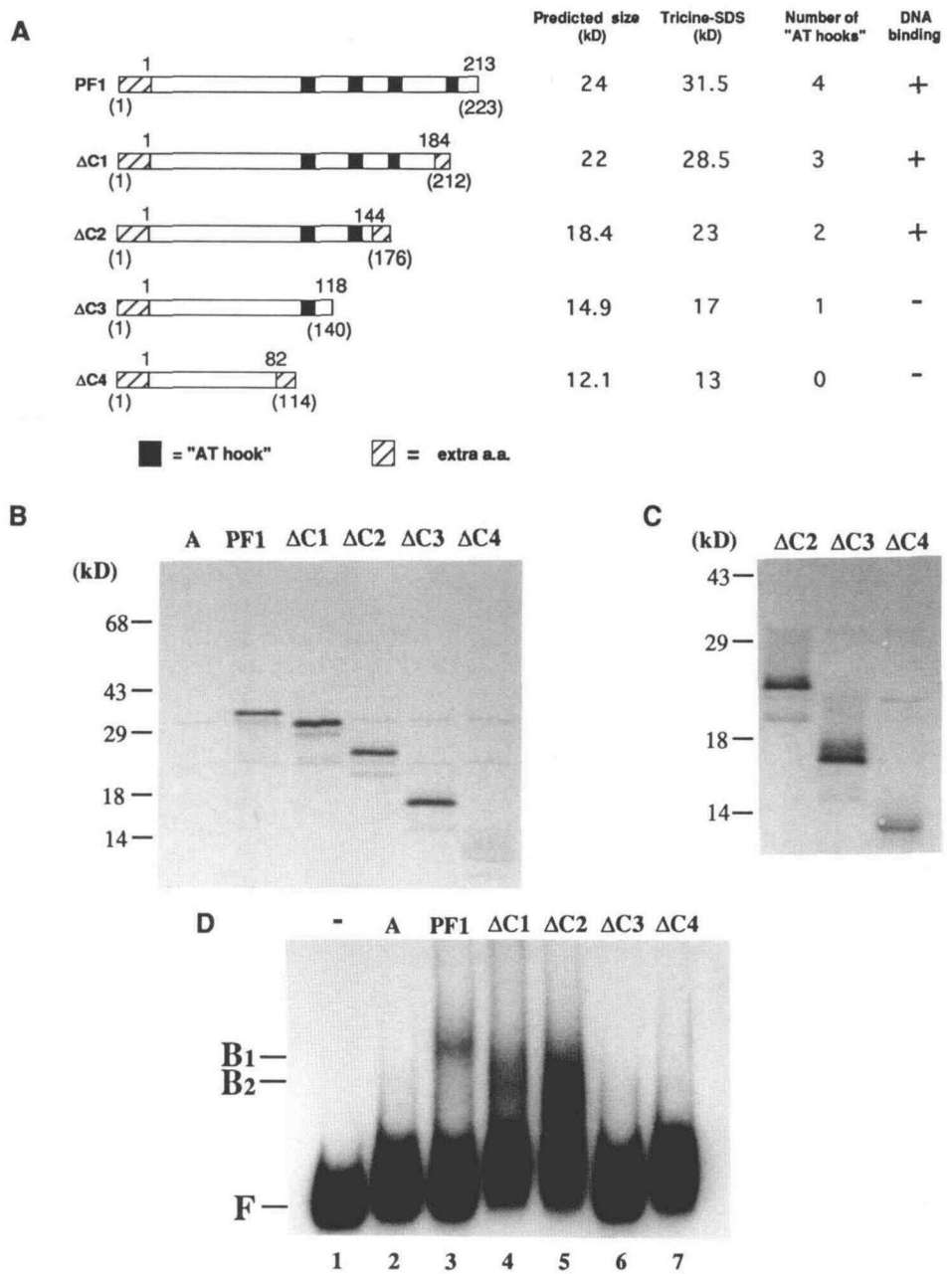


Figure 10. Binding of Recombinant PF1 and C-Terminal Deletion Derivatives of PF1 to a Promoter Fragment That Contains the PE1 Region.

(A) Schematic representations of polypeptides produced in vitro showing their predicted mass and their apparent molecular mass, as determined by the Tricine-SDS-PAGE system (Schägger and Von Jagow, 1987) shown in **(B)** and **(C)**; the positions and the number of putative A-T hook domains and their apparent DNA binding activity, as shown in **(D)**. Numbers above each protein diagram indicate the amino acid number of the PF1 sequence, as shown in Figure 8. Numbers in parentheses below indicate the amino acid position in protein encoded by each construct taking into account 20 extra amino acids at the N terminus and some extra amino acids at the C terminus due to the cloning procedure in vector pPO-9 (see Methods for details about extra sequences of each construct). Black boxes indicate A-T hooks; boxes with diagonal lines indicate extra amino acids at the N terminus or the C terminus; a.a., amino acids; molecular masses are given in kilodaltons.

(B) Aliquots of coupled in vitro transcription–translation reactions using as templates pR4 in antisense (A) or sense (PF1) orientations or C-terminal deletions ΔC1, ΔC2, ΔC3, and ΔC4. Reactions were analyzed by Tricine-SDS-PAGE to confirm and quantitate the presence of protein product. Molecular mass markers are given at left in kilodaltons.

(C) Aliquots of reactions transcribed–translated in vitro of constructs ΔC2, ΔC3, and ΔC4 were diluted fourfold relative to the experiment shown

performed with PF1 or C-terminal deletions of PF1 made in vitro using the 300-bp end-labeled P-PE1 fragment as target DNA. Four C-terminal deletions, ΔC1, ΔC2, ΔC3, and ΔC4, were tested. These deletions contain three, two, one, or no A-T hooks, respectively, as shown in Figure 10A.

Figures 10B and 10C show the in vitro transcription–translation products of reactions using as templates the antisense pR4 cDNA, the full-length pR4 cDNA in sense orientation, and the ΔC1, ΔC2, ΔC3, and ΔC4 deletions. As previously observed with the full-length PF1 made in vitro, the apparent molecular mass of all polypeptides in the Tricine-SDS gel system utilized in this experiment was larger than the predicted one based on sequence information (Figure 10A). It was observed that as deletions progressed toward the N terminus, the discrepancy between predicted and apparent molecular masses was smaller. This was in correlation with the gradual removal of both the A-T hooks and the proline- and alanine-rich regions found in the C-terminal half of the protein. The product of the ΔC4 deletion is better observed in Figure 10C after a fourfold dilution of the samples shown in Figure 10B. The ΔC4 protein comigrates with the most abundant protein in the rabbit reticulocyte lysate system, which apparently competes for binding sites during membrane transfer and therefore results in a diminished signal after transfer.

The results of the binding assay shown in Figure 10D indicated that PF1, ΔC1, and ΔC2 proteins retarded the mobility of the P-PE1 probe (lanes 3, 4, and 5). Therefore, deletion of the two C-terminal A-T hooks had no discernible effect on protein–DNA complex formation. Removal of three or more A-T hooks as in ΔC3 and ΔC4 resulted in loss of detectable protein–DNA complex formation in this assay (lanes 6 and 7). Superficially, these data suggest that removal of three or more A-T hooks from PF1 leads to loss of DNA binding activity. However, because we have no direct evidence that any ΔC3 or ΔC4 protein–DNA complexes would be resolved from the free probe under these assay conditions, the data do not unequivocally support this conclusion.

DISCUSSION

By screening an expression library from rice using a DNA probe with the sequence of the oat *PHYA3* promoter PE1 element, we have isolated two different cDNA clones, pR2 and pR4, encoding DNA binding proteins. Here, we have chosen to report on the characterization of clone pR4 because its absolute binding affinity toward the PE1 element is an order of magnitude higher than that of pR2. The data indicated the pR4 encodes a protein with DNA binding characteristics expected of PF1, the presumptive nuclear factor that enhances *PHYA3* gene transcription in vivo through binding to PE1. We therefore define the pR4-encoded protein as rice PF1.

The data supporting this conclusion were derived from several complementary types of experiments. First, we compared the binding to PE1 of synthetic oligonucleotides containing the O-PE1 sequence or mutations of PE1 (O-614, O-615, and O-616) that are known to disrupt expression of a reporter gene driven by the *PHYA3* promoter in a rice transient expression assay (Bruce et al., 1991). Binding of O-PE1 was one to two orders of magnitude higher than binding of the mutant oligonucleotides (Figure 1D).

Additional DNA binding assays (Figures 1C, 5, 6, and 7) showed that although PF1 can bind other AT-rich sequences upstream of PE1 in the oat *PHYA3* promoter, it binds to PE1 with higher affinity. Direct evidence that binding occurs outside of PE1 was obtained by in vitro footprinting of the oat *PHYA3* promoter. Protection occurred within the PE1 region (−347 to −358 and −363 to −368 bp) as well as in two other AT-rich regions just upstream of the PE1 region (positions −374 to −385 and −387 to −396) (Figure 7). Oat nuclear protein extracts used previously in in vitro–footprinting experiments also protected both the PE1 region and the two AT-rich regions at positions −374 to −385 and −387 to −396 (Bruce et al., 1991). However, when we used quantitative mobility shift assays to compare the binding to PF1 of a 300-bp oat *PHYA3* promoter fragment containing P-PE1 with linker substitution derivatives of PE1 in the same promoter fragment (P-614, P-615, and P-616), we observed that although some binding occurred at high protein concentrations to P-615 and P-616, binding was greatest at low protein concentrations to the wild-type P-PE1 (Figure 5). In addition, in competition experiments, wild-type OT-PE1 was able to compete effectively with the binding of PF1 to P-PE1, but none of the mutant derivatives of PE1 (OT-614, OT-615, and OT-616) was able to do so (Figure 6).

Overall these data indicate that although PF1 can bind to AT-rich DNA sequences other than PE1, the protein has highest affinity for PE1 among the AT-rich motifs in the *PHYA3* promoter.

Figure 10. (continued).

in **(B)**. The protein blot was exposed 10 times longer than in **(B)**. Dilution confirms the presence of the protein product of the ΔC4 construct. Molecular mass markers are given at left in kilodaltons.

(D) Mobility shift assay of in vitro–synthesized proteins. Lane 1, no protein extract (−); lane 2, product of reactions using pR4 in antisense orientation as template (A); lane 3, full-length recombinant PF1; lanes 4, 5, 6, and 7 are ΔC1, ΔC2, ΔC3, and ΔC4 C-terminal deletions of recombinant PF1, respectively. End-labeled 300-bp oat *PHYA3* gene promoter fragment containing PE1 region (position −415 to −116 bp; see sequence in Figure 4) was used as the probe. Equal amounts of ^{35}S-labeled protein (normalized for the number of methionines present in each construct) were applied to each lane in **(B)** and **(C)** and to each reaction in **(D)**.

F, free DNA; B1 and B2 are protein–DNA complexes.

It appears, therefore, that PF1 may possess sequence-selective DNA binding properties leading to preferential interaction with certain AT-rich motifs relative to others. Such selectivity has been observed for other proteins in this class (Churchill and Travers, 1991). This DNA binding behavior of PF1 is consistent with the observation that a *PHYA3* promoter fragment that contains the two upstream AT-rich motifs (−374 to −385 and −387 to −396), but is mutated in the PE1 motif, has only background activity in a transfection assay (Bruce et al., 1991). We suggest that the higher affinity of PF1 for PE1 might lead to functionally productive interactions in vivo, whereas the weaker interaction of PF1 with the other AT-rich motifs in the *PHYA3* gene promoter may be unproductive.

The size of the pR4 cDNA is similar to the size of its mRNA, indicating that the pR4 cDNA clone is close to full length. Additional observations that support this argument are that we were unable to obtain a larger clone either by rescreening the same rice cDNA library or by amplification of cDNAs by PCR at their 5′ ends. Most directly, protein blot analysis using a polyclonal antiserum raised against the recombinant rice PF1 indicated that the recombinant PF1 made in vitro comigrates on SDS-PAGE with a rice nuclear-extract protein recognized by this antiserum (Figure 3).

Examination of the predicted sequence of rice PF1 indicates the presence of four repeats of the A-T hook motif, as well as of sequence similarity to histone H1. The A-T hook, also known as the GRP motif, is a DNA binding motif found in the mammalian HMG I-Y chromosomal proteins that is necessary and sufficient for binding to the narrow minor groove of stretches of AT-rich DNA (Reeves and Nissen, 1990; Churchill and Travers, 1991). Thus, the presence of this motif in PF1 is consistent with its DNA binding behavior. The A-T hook motif is repeated three times in mammalian HMG Y and HMG I proteins (Eckner and Birnstiel, 1989; Johnson et al., 1989) and four times in the protein encoded by the soybean SB16 cDNA that also binds to AT-rich DNA (Laux et al., 1991).

The rice PF1 protein matches in 18 out of 21 positions (residues 28 to 48) a short, published peptide sequence determined for a purified wheat protein designated HMGa (Arwood and Spiker, 1990). This wheat protein has also recently been found to contain A-T hook motifs and other regions of sequence similarity to PF1 but is smaller (26 kD) on SDS gel electrophoresis (B. Wakim and S. Spiker, personal communication). Together with the observation that the purified wheat HMGa protein binds preferentially to AT-rich DNA tracts (Pedersen et al., 1991), the data suggest that HMGa could be functionally as well as structurally related to PF1. We have also isolated an oat cDNA clone, pO2, from an expression library by the same DNA binding method used in the present paper. pO2 encodes an A-T hook–containing protein, designated oat PF1, that has high sequence similarity to rice PF1 (J. Nieto-Sotelo and P.H. Quail, unpublished data).

AT-rich sequences that have been shown to have a functional role in the stimulation of transcription of a number of different genes have been found in yeast (Struhl, 1985), animals (Fashena et al., 1992; Thanos and Maniatis, 1992), and plants (Bruce and Quail, 1990; Bruce et al., 1991; Czarnecka et al., 1992; Wang et al., 1992). A larger number of plant gene promoters have been reported to contain AT-rich regions to which nuclear proteins bind in vitro (Bustos et al., 1989; Datta and Cashmore, 1989; Jordano et al., 1989; Forde et al., 1990; Lam et al., 1990; Pedersen et al., 1991; Cushman and Bohnert, 1992; Czarnecka et al., 1992; Echeverria et al., 1992). In some of these systems, these AT-rich DNA binding proteins have been biochemically characterized as having the operational features of HMG proteins (Maier et al., 1990; Pedersen et al., 1991; Czarnecka et al., 1992). In soybean leaf and nodule nuclei, these AT-rich DNA binding proteins have been proposed to be HMG I-like proteins (Jacobsen et al., 1990), and more recently cDNA clones encoding HMG Y-like proteins that bind to d(AT) tracks 5′ to soybean seed protein genes have been cloned (Laux et al., 1991). A tobacco cDNA clone encoding 3AF1, a nuclear factor that binds an AT-rich element very similar to PE1 in the ribulose bisphosphate carboxylase small subunit *rbcS-3A* promoter, has been isolated (Lam et al., 1990). 3AF1 requires metal to bind DNA and is proposed to contain zinc-fingerlike DNA binding domains (Lam et al., 1990). 3AF1 lacks consensus A-T hooks and shows no significant sequence similarity to either rice PF1 or human HMG I and HMG Y proteins. Together, the above data indicate that two different families of DNA binding proteins may interact with AT-rich DNA in plant promoters: the HMG-like family and the 3AF1 type.

How many A-T hooks are necessary for high-affinity DNA binding? It has been observed that a synthetic oligopeptide having the sequence of just one A-T hook binds to DNA in a manner similar to intact HMG I proteins (Reeves and Nissen, 1990; Churchill and Travers, 1991). However, the binding of the synthetic peptide to DNA showed dissociation constant (K_d) values several orders of magnitude higher (10 μM) than the K_d for intact HMG I proteins (1 nM). These results suggest a cooperative interaction between the multiple A-T hooks present in a given HMG I molecule in its binding to DNA (Reeves and Nissen, 1990). Our preliminary experiments with C-terminal deletions of rice PF1 indicated that the presence of the two N-terminal A-T hooks is sufficient for this protein to bind DNA with high affinity (Figure 10). However, the data lack the resolution necessary to determine whether removal of one or both of the remaining N-terminal A-T hooks would reduce this affinity.

Rice PF1 contains a stretch of 11 amino acids that are present in mammalian HMG I but not HMG Y proteins (amino acid residues 109 to 119 of rice PF1). The function of this domain is not known but it has been speculated that it could be a dimerization domain for the formation of homodimers and/or heterodimers due to its hydrophobic character (Eckner and Birnstiel, 1989). It is possible that this domain of PF1 could be involved in protein–protein interactions, perhaps with other factors involved in transcriptional activation.

The elucidation of the function of HMG I and HMG Y proteins in mammals has just begun. It is known that these chromosomal proteins are expressed at elevated levels in proliferating, undifferentiated cells and may be implicated in cell division in vivo and in the maintenance of the

undifferentiated state (Johnson et al., 1990; Manfiolетti et al., 1991). Recently, the direct involvement of HMG I and HMG Y proteins in transcription has been established in mammalian cells. It was demonstrated that HMG I proteins bind to poly d(A-T)–rich upstream activator sequences that are regulatory components in the expression of the lymphotoxin gene in murine cells (Fashena et al., 1992). Also HMG I-Y proteins are required along with NF-κB for virus induction of the human interferon-β gene (Thanos and Maniatis, 1992). HMG I-Y proteins also antagonized histone H1-mediated general repression of transcription in an in vitro T7 RNA polymerase transcription assay, suggesting a possible role in chromatin opening (Zhao et al., 1993).

The region in which rice PF1 shows a high degree of similarity to the published wheat HMGa peptide sequence (Arwood and Spiker, 1990) resides in the N-terminal domain where rice PF1 shows the highest degree of similarity (10 identical amino acids of 21) to wheat histone H1 (Brandt and Von Holt, 1986), pea histone H1 (Gantt and Key, 1987), and chicken histone H5 (Wu et al., 1986). The protein encoded by the soybean SB16 cDNA also resembles pea histone H1 in the N-terminal 50 amino acids (Laux et al., 1991). H1 histone proteins are presumed to play a major role in the formation of chromatin higher order structure, specifically in the formation of the chromatosome (Pederson et al., 1986). Unlike the histone proteins that form the octamer in the core particle, H1 histones are known to be a heterogeneous group of proteins that differ also in their ability to condense chromatin (Cole, 1984). Recently, H1 histones have been shown to regulate transcription as repressors by altering the available linker DNA and therefore regulating the interaction of transcription factors with the promoter (reviewed by Grunstein, 1990). The significance of the similarity of rice/oat PF1, wheat HMGa, and soybean SB16 proteins to both H1 histone and HMG I-Y proteins remains to be explored. It is conceivable that their resemblance to histone H1 may allow these proteins to interact with the chromatin in a manner similar to histone H1.

Does PF1 have a role in transcriptional activation? The primary evidence favoring such a role is that PF1 binds in an apparently sequence-selective fashion to the functionally defined PE1 element and does not bind efficiently to mutant derivatives of this element that are functionally inactive in gene transfer experiments (Bruce et al., 1991). However, this question is yet to be addressed directly by domain swapping experiments with the different regions of the factor to assay for possible activator domains. Inspection of the predicted PF1 sequence indicates the lack of acidic activation domains common in many eukaryotic transcription factors (Ptashne, 1988). A proline-rich region in the transcription factor CTF/NF-1 has been found to function as an activator domain (Mermod et al., 1989). It is not clear what the functional significance of the proline-rich regions in PF1 may be, whether as activation domains, as DNA binding domains, or both. Most likely, however, PF1 may be involved in activation of transcription through its synergistic interactions with other factors, a common mechanism in eukaryotes (Herbomel, 1990). This notion is strongly supported by experiments with the rice transient expression assay that showed that both the PE1 and PE3 regions are necessary for high-level expression in low-Pfr cells (Bruce and Quail, 1990; Bruce et al., 1991). Our earlier model predicts such synergistic interaction between PF1 and PF3, the factor or factors that bind PE3 (Bruce et al., 1991).

In the rice *PHYA* promoter, the PE1 element is not present. Instead, an element, designated GT2 box, is found upstream of the TATA box and the PE3 element (Dehesh et al., 1992). The GT2 box has been shown to be essential along with the TATA box and PE3 element for expression in the context of the rice *PHYA* promoter (Dehesh et al., 1990). A cDNA encoding the GT2 box binding protein, designated GT-2, has been isolated and sequenced (Dehesh et al., 1990, 1992). Analysis of the predicted structure of the protein encoded by this cDNA indicates that it contains a new class of DNA binding motif designated as the trihelix motif (Dehesh et al., 1992). Thus, although PF1 and GT2 are totally different proteins that bind to different elements in their respective promoters, they would seem to be involved in transcriptional regulation by means of similar interactions with PF3, TATA box-associated factors, or other components of the transcriptional machinery.

The evidence presented here indicates that rice PF1 is a member of a relatively new class of DNA binding proteins that have the potential to function in transcriptional activation through recognition of AT-rich target sequences. Direct verification of this potential role must await further analysis, such as that provided by in vitro transcription assays.

METHODS

Plant Material and Growth Conditions

Soaked rice (*Oryza sativa* IR36) seeds were grown on moist termiculite at 26°C in the dark for the periods indicated. Harvesting of tissue took place under green safe light.

RNA Isolation

Total RNA was isolated from 7-day-old dark-grown seedlings. Using scissors, the top 2 to 3 cm of shoot tissue was harvested, immediately weighed, frozen under liquid nitrogen, and kept at −80°C until RNA isolation. Total RNA was isolated by the method of Rochester et al. (1986). Poly(A)$^+$ RNA was selected by oligo(dT) cellulose chromatography, as described previously by Theologis et al. (1985).

Construction and Screening of cDNA Library

A cDNA library was constructed using a commercially available cDNA synthesis kit (Promega), EcoRI adaptors (Promega), and λgt11 arms (Stratagene). Starting with 4 μg of poly(A)$^+$ RNA, 1.6 μg of cDNA was obtained. From 50 ng of cDNA ligated to λgt11 arms, 12.8×10^6 recombinant phage were obtained.

Screening of the cDNA library was done by the method of Vinson et al. (1988) using a double-stranded O-PE1 probe that was prepared

by annealing two complementary synthetic oligonucleotides 5′-GGCTGGAAATAGCAAATGTTAAAAATAAA-3′ and 5′-AGCCTTTATTTTTAACATTTGCTATTTCC-3′ that were previously gel purified. After annealing, oligonucleotides were phosphorylated with kinase, ligated, and radiolabeled with phosphorus-32 and a nick translation kit (Amersham) using standard techniques to a specific activity of at least 10^8 cpm/mg of DNA (Sambrook et al., 1989). A total of 750,000 recombinant phages was screened, and two positive clones were obtained after three rounds of screening. Characterization of the DNA binding properties of protein made by recombinant phage was performed using a filter binding assay by comparing the binding of the oligonucleotide of PE1 (O-PE1) probe to mutant probes. Filters were incubated in the presence of equal amounts of cpm of probes that were radiolabeled to comparable specific activities. Mutant probes were synthetic double-stranded oligonucleotides O-614 (5′-GGCTGCTCGAGTCAAATGTTAAAAATAAA-3′ and 5′-AGCCTTTATTTTTAACATTTGACTCGAGC-3′), O-615 (5′-GGCTGGAAATAGGCTCGAGTAAAAATAAA-3′ and 5′-AGCCTTTATTTTTACTCGAGCCTATTTCC-3′), and O-616 (5′-GGCTGGAAATAGCAAATGTGCTCGAGTAA-3′ and 5′-AGCCTTACTCGAGCACATTTGCTATTTCC-3′) that were prepared in a manner similar to O-PE1. DNA bound to filters containing blotted phage protein was estimated by liquid scintillation counting and by autoradiography. cDNA inserts were subcloned as EcoRI fragments in two orientations in the EcoRI site of pBluescript SK+ vector (Stratagene) for sequence analysis, and in the pPO-9 vector for overexpression of the encoded protein in *Escherichia coli* (for more details, see the section entitled Expression of PF1 Protein in Bacteria and Protein Extract Preparation).

RNA Gel Blot Analysis

Gel blot analysis was conducted as described previously by Colbert et al. (1985) using 1.25 μg of poly(A)$^+$ RNA per lane.

Rice Nuclear Extract Preparation

Harvesting and homogenization of tissue were conducted under green safe light at 4°C. The top 2 to 3 cm of the shoot of 13-day-old rice seedlings was harvested (50 g) and homogenized in 200 mL of extraction buffer (1 M hexylene glycol, 10 mM Pipes-KOH, pH 7.0, 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 0.8 mM phenylmethylsulfonyl fluoride) (Watson and Thompson, 1986) in a blender (Waring, New Hartford, CT) fitted with a tower of new razor blades (Kannangara et al., 1977). The homogenate was filtered through 1000-, 149-, and 62-μm pore-sized nylon mesh (Small Parts, Inc., Miami, FL) wetted in extraction buffer. The filtrate was adjusted to 0.5% Triton X-100, sedimented at 3000*g* for 10 min, and resuspended in 16 mL of gradient buffer (0.5 M hexylene glycol, 10 mM Pipes-KOH, pH 7.0, 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 0.5% Triton X-100, 0.8 mM phenylmethylsulfonyl fluoride) (Watson and Thompson, 1986). Resuspended material was sedimented again at 3000*g* for 5 min, resuspended in 2 mL of gradient buffer; 670 μL of 80% glycerol was added, and samples were frozen at −80°C. Isolated nuclei were thawed on ice, spun at 3000*g* for 5 min, and resuspended in 2 mL of nuclei freezing buffer; nuclear extracts were prepared as described previously by Bruce et al. (1991). Protein concentration was estimated by the Bradford assay (Bradford, 1976).

Expression of PF1 Protein in Bacteria and Protein Extract Preparation

The cDNA insert from recombinant λgt11 bacteriophage λR4 was subcloned as an EcoRI fragment into plasmid vector pPO-9 (Rottmann et al., 1991), which is a modified version of the pET-3a vector originally described by Rosenberg et al. (1987). Vector pPO-9 keeps the EcoRI cDNA insert in the same frame as it was in the original λgt11 phage and, as its parent plasmid pET-3a, expression of the cDNA is under the control of a T7 promoter. Sense and antisense orientation versions of pR4 cDNA insert subcloned in pPO-9 were transformed into *E. coli* BL21(DE3) cells. Transformed bacteria were induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at an OD of 600 nm between 0.5 to 1.0 for 3 hr, harvested, and resuspended in 1 × binding buffer (20 mM Hepes, pH 7.9, 3 mM $MgCL_2$, 40 mM KCl, 1 mM DTT, 20% glycerol, 1 μg/mL leupeptin, 1 μg/mL antipain). Protein extracts were obtained by freezing and thawing three times between liquid N_2 and room temperature followed by sonication on ice. The suspension was centrifuged at 80,000 rpm for 20 min at 4°C (TL 100 ultracentrifuge and rotor TLA100.3; Beckman), and the supernatant fraction was saved and stored in small aliquots at −80°C. Protein concentration was measured by the Bradford assay (Bradford, 1976).

DNA Binding Assays in Vitro

Protein Blot Probed with DNA Assay

Protein extracts were resolved in 11% polyacrylamide–SDS gels and transferred to nitrocellulose. Blots were denatured/renatured with guanidine-HCl, hybridized to nick-translated O-PE1 catenated double-stranded oligonucleotides for 3 to 4 hr at 4°C, washed, and autoradiographed as described by Vinson et al. (1988).

Mobility Shift Analysis

Protein extracts were incubated in 1 × binding buffer with 1 to 4 μg of salmon sperm DNA and an end-labeled probe at room temperature for 15 min. Samples were analyzed by electrophoresis on 6% or 10% polyacrylamide gels in 1 × Tris-borate-EDTA buffer. End-labeled probes were a 300-bp fragment of the oat phytochrome A3 gene (*PHYA3*) promoter (−415 to −116 relative to transcription start site) or linker substitution mutants of this same 300-bp promoter fragment. The 300-bp native promoter probe corresponds to the BamHI-SalI fragment of clone 449 (Bruce and Quail, 1990), whereas the linker substitution mutants correspond to the BamHI-SalI 300-bp fragments of clones 614, 615, and 616 (Bruce et al., 1991).

DNase I Footprinting Assays

Plasmid DNA used as a source of probe was clone 449*. Clone 449* was obtained by introducing a SalI site to the 449 clone by polymerase chain reaction at positions −301 to −306 bp. The 119-bp BamHI-SalI fragment was subcloned into pBluescript SK+. The resulting subclone was named 449*. Plasmid DNA from subclone 449* was linearized with BamHI. Ends were radiolabeled by filling in with the Klenow fragment of DNA polymerase I in the presence of α-^{32}P-dGTP and α-^{32}P-dATP and chased with 2.5 mM of unlabeled deoxynucleotide triphosphates. A 259-bp DNA insert was released with PvuII and purified by agarose gel electrophoresis. This fragment contains 109 bp

of the oat *PHYA3* gene promoter (positions −415 to −307), 6 bp of the introduced SalI fragment, and 144 bp of the pBluescript SK+ sequences between the SalI and PvuII sites at position 532 of pBluescript SK+. The fragment was eluted from the gel, phenol/chloroform extracted, ethanol precipitated, and resuspended in 10 mM Tris, pH 8.0, 20 mM NaCl, 1 mM EDTA. The resulting probe was labeled on the bottom strand at the BamHI end. The footprinting reaction was performed as described previously by Bruce et al. (1991) using 1 μg of salmon sperm DNA as nonspecific competitor DNA instead of poly(dI-dC). G and G+A reactions were performed according to the method of Maxam and Gilbert (1980).

Antibody Production

PF1 overproduced in *E. coli* was purified by electroelution after electrophoretic separation in preparative SDS–polyacrylamide gels and injected into three Swiss Webster mice to raise anti-PF1 polyclonal antibodies, as described by Hunter et al. (1990). Production of the polyclonal antibody was performed at the University of California-Berkeley Hybridoma Facility.

Protein Gel Blotting

Proteins were separated electrophoretically on SDS–polyacrylamide gels (Laemmli, 1970). After blotting onto nitrocellulose membranes, the filter was blocked overnight at 4°C in milk buffer (4% powdered milk, 20 mM Tris, pH 7.5, 0.15 mM NaCl, 0.15% Nonidet P-40). Incubation with primary antibody (1:7500 dilution in milk buffer) depleted of *E. coli*–reactive antibodies was for 3 hr at room temperature followed by three 10-min washes with milk buffer. Incubation with secondary antibody (goat anti–mouse alkaline phosphatase–conjugated antibody [Promega]) at a 1:5000 dilution in milk buffer was for 1 hr followed by three washes in milk buffer and development of alkaline phosphatase activity in 100 mM Tris, pH 8.9, 100 mM NaCl, 5 mM $MgCl_2$, 75 μg/mL 5-bromo-4-chloro-3-indolyl phosphate, 150 μg/mL nitro blue tetrazolium.

Coupled in Vitro Transcription-Translation

Subcloning into the pPO-9 vector of C-terminal deletion derivatives of pR4 cDNA in pBluescript SK+ was done by cutting each deletion derivative with MluI and PvuII generating a DNA fragment with a 5′ end at bp 51 of the cDNA and a 3′ end at the PvuII site proximal to the SacI site of pBluescript SK+. These fragments contain the 3′ end of each deletion derivative cDNA plus 218 bp of pBluescript SK+ vector sequences between the SacI site and PvuII site. The resulting MluI-PvuII fragment was ligated to a pR4 cDNA fragment cloned in pPO-9 previously cut with MluI and BamHI (blunt ended). This procedure regenerates a pPO-9 vector containing a C-terminal deletion derivative of the original pR4 cDNA insert with stop codons provided by the pBluescript SK+ sequences between the SacI and PvuII sites. Plasmid DNA from pR4 cDNA or C-terminal deletion derivatives ΔC1, ΔC2, ΔC3, and ΔC4 subcloned in pPO-9 vector were transcribed and translated in vitro in the presence of ^{35}S-methionine using a coupled in vitro transcription–translation kit (TNT lysate; Promega). Products were analyzed by autoradiography after electrophoretic separation on a Tricine SDS–polyacrylamide gel system (Schägger and Von Jagow, 1987) followed by blotting to nitrocellulose membrane (Millipore Corp.). Synthesized protein was estimated by measuring incorporation of Sulfur-35 in the TCA-precipitable fraction by liquid scintillation counting. Both pR4 cDNA and C-terminal deletion derivatives expressed from pPO-9 are predicted to contain 20 extra amino acids at their N-terminal end as a result of translation being initiated from ATG in the vicinity of the pPO-9 EcoRI cloning site. Therefore, all in vitro–translated constructs initiate with the following sequence at their N terminus: MASEFRCCRSLPLFTPRLTP. In addition, C-terminal deletion derivatives are predicted to have a few extra amino acids at their C terminus due to the position of the nearest stop codon in the original pBluescript SK+ vector from where these cDNA inserts came. Deletion ΔC1 has eight extra amino acids at the C terminus: GSSSFCSL. Deletions ΔC2 and ΔC4 have 12 extra amino acids at the C terminus: GGAPAFVPFSEG. Deletion ΔC3 has two extra amino acids at the C terminus: EG.

DNA Sequencing and Sequence Analyses

The nucleotide sequence of the cDNA insert subcloned into the pBluescript SK+ vector was determined on both strands by the dideoxy chain terminating method (Sanger et al., 1977) as modified for sequencing double-stranded DNA preparations (Seto, 1990) using Sequenase (U.S. Biochemical Corp.). Sequence of the coding strand was obtained by generating a series of overlapping 3′ deletions using an ExoIII-mung bean nuclease kit (Stratagene). Noncoding strand sequence was obtained by using specific synthetic oligonucleotides as the continuing primers. DNA sequence was analyzed using programs of the Genetics Computer Group Sequence Analysis Software Package (Devereux et al., 1984).

ACKNOWLEDGMENTS

We thank Dr. Katayoon Dehesh for advice on protocols for screening λ expression libraries, Dr. Xing-Wang Deng for advice on nuclei isolation methods, and Dr. Paul Oeller for providing us with the pPO-9 vector. We are indebted to Dr. Alex Karu and his colleagues of the University of California at Berkeley, College of Natural Resources Hybridoma Facility for production of polyclonal antibodies. Finally, we would like to thank the members of our laboratory for their helpful criticism and discussion during the course of this work. This work was supported by National Science Foundation Grant No. MCB-9220161 and U.S. Department of Agriculture/Agricultural Research Service Current Research Information Service Grant No. 5335-21000-006-00D.

Received October 4, 1993; accepted December 9, 1993.

REFERENCES

Arwood, L.J., and Spiker, S. (1990). Binding of wheat and chicken high mobility group chromosomal proteins to DNA and to wheat and chicken mononucleosomes. J. Biol. Chem. **265,** 9771–9777.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72,** 248–254.

Brandt, W.F., and von Holt, C. (1986). Variants of wheat histone H1 with N- and C-terminal extensions. FEBS Lett. **194,** 282–286.

Bruce, W.B., and Quail, P.H. (1990). *cis*-Acting elements involved in photoregulation of an oat phytochrome promoter in rice. Plant Cell **2,** 1081–1089.

Bruce, W.B., Christensen, A.H., Klein, T., Fromm, M., and Quail, P.H. (1989). Photoregulation of a phytochrome gene promoter from oat transferred into rice by particle bombardment. Proc. Natl. Acad. Sci. USA **86,** 9692–9696.

Bruce, W.B., Deng, X.-W., and Quail, P.H. (1991). A negatively acting DNA sequence element mediates phytochrome-directed repression of *phyA* gene transcription. EMBO J. **10,** 3015–3024.

Bustos, M.M., Guiltinan, M.J., Jordano, J., Begum, D., Kalkan, F.A., and Hall, T.C. (1989). Regulation of β-glucuronidase expression in transgenic tobacco plants by an A/T-rich, *cis*-acting sequence found upstream of a French bean β-phaseolin gene. Plant Cell **1,** 839–853.

Churchill, M.E.A., and Travers, A. A. (1991). Protein motifs that recognize structural features of DNA. Trends Biochem. Sci. **16,** 92–97.

Colbert, J.T., Hershey, H.P., and Quail, P.H. (1985). Phytochrome regulation of phytochrome mRNA abundance. Plant Mol. Biol. **5,** 91–102.

Cole, R.D. (1984). A minireview of microheterogeneity in H1 histone and its possible significance. Anal. Biochem. **136,** 24–30.

Cushman, J.C., and Bohnert, H.J. (1992). Salt stress alters A/T-rich DNA-binding factor interactions within the phosphoenolpyruvate carboxylase promoter from *Mesembryanthemum crystallinum.* Plant Mol. Biol. **20,** 411–424.

Czarnecka, E., Ingersoll, J.C., and Gurley, W.B. (1992). AT-rich promoter elements of soybean heat shock gene *Gmhsp17.5E* bind two distinct sets of nuclear proteins in vitro. Plant Mol. Biol. **19,** 985–1000.

Datta, N., and Cashmore, A.R. (1989). Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. Plant Cell **1,** 1069–1077.

Dehesh, K., Bruce, W.B., and Quail, P.H. (1990). A *trans*-acting factor that binds to a GT-motif in a phytochrome gene promoter. Science **250,** 1397–1399.

Dehesh, K., Hung, H., Tepperman, J.M., and Quail, P.H. (1992). GT-2: A transcription factor with twin autonomous DNA-binding domains of closely related but different target sequence specificity. EMBO J. **11,** 4131–4144.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. **12,** 387–395.

Echeverria, M., Delcasso-Tremousaygue, D., and Delseny, M. (1992). A nuclear protein binding to dA/dT-rich sequences upstream from the radish DNA promoter. Plant J. **2,** 211–219.

Eckner, R., and Birnstiel, M.L. (1989). Cloning of cDNAs coding for human HMG I and HMG Y proteins: Both are capable of binding to the octamer sequence motif. Nucl. Acids Res. **17,** 5947–5959.

Fashena, S.J., Reeves, R., and Ruddle, N.H. (1992). A poly(dA-dT) upstream activating sequence binds high-mobility group I protein and contributes to lymphotoxin (tumor necrosis factor-β) gene regulation. Mol. Cell Biol. **12,** 894–903.

Forde, B.G., Freeman, J., Oliver, J.E., and Pineda, M. (1990). Nuclear factors interact with conserved A/T-rich elements upstream of a nodule-enhanced glutamine synthetase gene from French bean. Plant Cell **2,** 925–939.

Gantt, J.S., and Key, J.L. (1987). Molecular cloning of a pea H1 histone cDNA. Eur. J. Biochem. **166,** 119–125.

Gilmartin, P.M., Sarokin, L., Memelink, J., and Chua, N.-H. (1990). Molecular light switches for plant genes. Plant Cell **2,** 369–378.

Grunstein, M. (1990). Histone function in transcription. Annu. Rev. Cell Biol. **6,** 643–678.

Herbomel, P. (1990). Synergistic activation of eukaryotic transcription: The multiacceptor target hypothesis. New Biol. **2,** 1063–1070.

Hershey, H.P., Barker, R.F., Idler, K.B., Murray, M.G., and Quail, P.H. (1987). Nucleotide sequence and characterization of a gene encoding the phytochrome polypeptide from *Avena.* Gene **61,** 339–348.

Hunter, B.G., Richardson, J., Dietzgen, R.G., Karu, A.E., Sylvester, E.S., Jackson, A.O., and Morris, J.T. (1990). Purification and characterization of strawberry crinkle virus. Phytopathology **80,** 282–287.

Jacobsen, K., Laursen, N.B., Jensen, E.Ø., Marcker, A., Poulsen, C., and Marcker, K.A. (1990). HMG I-like proteins from leaf and nodule nuclei interact with different AT motifs in soybean nodulin promoters. Plant Cell **2,** 85–94.

Johnson, K.R., Lehn, D.A., and Reeves, R. (1989). Alternative processing of mRNAs encoding mammalian chromosomal high-mobility-group proteins HMG-I and HMG-Y. Mol. Cell. Biol. **9,** 2114–2123.

Johnson, K.R., Disney, J.E., Wyatt, C.R., and Reeves, R. (1990). Expression of mRNAs encoding mammalian chromosomal proteins HMG-I and HMG-Y during cellular proliferation. Exp. Cell Res. **187,** 69–76.

Jordano, J., Almoguera, C., and Thomas, T.L. (1989). A sunflower helianthinin gene upstream sequence ensemble contains an enhancer and sites of nuclear protein interaction. Plant Cell **1,** 855–866.

Kannangara, C.G., Gough, S.P., Hansen, B., Rasmussen, J.N., and Sampson, K.J. (1977). A homogenizer with replaceable razor blades for bulk isolation of active barley plastids. Carlsberg Res. Commun. **42,** 431–434.

Kendrick, R.E., and Kronenberg, G.H.M. (1986). Photomorphogenesis in Plants. (Dordrecht: Martinus Nijhoff).

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227,** 680–685.

Lam, E., Kano-Murakami, Y., Gilmartin, P., Niner, B., and Chua, N.-H. (1990). A metal-dependent DNA-binding protein interacts with a constitutive element of a light-responsive promoter. Plant Cell **2,** 857–866.

Laurent, B.C., Treitel, M.A., and Carlson, M. (1991). Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in transcriptional activation. Proc. Natl. Acad. Sci. USA **88,** 2687–2691.

Laux, T., Seurinck, J., and Goldberg, R.B. (1991). A soybean embryo cDNA encodes a DNA binding protein with histone and HMG-protein-like domains. Nucl. Acids Res. **19,** 4768.

Lissemore, J.L., and Quail, P.H. (1988). Rapid transcriptional regulation by phytochrome of the genes for phytochrome and chlorophyll a/b-binding protein in *Avena sativa.* Mol. Cell. Biol. **8,** 4840–4850.

Maier, U-G., Grasser, K.D., Haass, M.M., and Feix, G. (1990). Multiple proteins bind to the P2 promoter region of the zein gene pMS1 of maize. Mol. Gen. Genet. **221,** 164–170.

Manfioletti, G., Giancotti, V., Bandiera, A., Buratti, E., Sautière, P., Cary, P., Crane-Robinson, C., Coles, B., and Goodwin, G.H.

(1991). cDNA cloning of the HMGI-C phosphoprotein, a nuclear protein associated with neoplastic and undifferentiated phenotypes. Nucl. Acids Res. **19,** 6793–6797.

Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65,** 499–560.

Mermod, N., O'Neill, E.A., Kelly, T.J., and Tjian, R. (1989). The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. Cell **58,** 741–753.

Pedersen, T.J., Arwood, L.J., Spiker, S., Guiltinan, M.J., and Thompson, W.F. (1991). High mobility group chromosomal proteins bind to AT-rich tracts flanking plant genes. Plant Mol. Biol. **16,** 95–104.

Pederson, D.S., Thoma, F., and Simpson, R.T. (1986). Core particle, fiber, and transcriptionally active chromatin structure. Annu. Rev. Cell Biol. **2,** 117–147.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature **335,** 683–689.

Quail, P.H. (1991). Phytochrome: A light-activated molecular switch that regulates plant gene expression. Annu. Rev. Genet. **25,** 389–409.

Reeves, R., and Nissen, M.S. (1990). The A·T-DNA-binding domain of mammalian high mobility group I chromosomal proteins: A novel peptide motif for recognizing DNA structure. J. Biol. Chem. **265,** 8573–8582.

Rochester, D.E., Winter, J.A., and Shah, D.M. (1986). The structure and expression of maize genes encoding the major heat shock protein, hsp70. EMBO J. **5,** 451–458.

Rosenberg, A.H., Lade, B.N., Chui, D.-S., Lin, S.-W., Dunn, J.J., and Studier, F.W. (1987). Vectors for selective expression of cloned cDNAs by T7 RNA polymerase. Gene **56,** 125–135.

Rottmann, W.H., Peter, G.F., Oeller, P.W., Keller, J.A., Shen, N.F., Nagy, B.P., Taylor, L.P., Campbell, A.D., and Theologis, A. (1991). 1-Aminocyclopropane-1-carboxylate synthase in tomato is encoded by a multigene family whose transcription is induced during fruit and floral senescence. J. Mol. Biol. **222,** 937–961.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-termination inhibitors. Proc. Natl. Acad. Sci. USA **74,** 5463–5467.

Schägger, H., and Von Jagow, G. (1987). Tricine-sodium dodecylsulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166,** 368–379.

Seto, D. (1990). An improved method for sequencing double stranded plasmid DNA from minipreps using DMSO and modified template preparation. Nucl. Acids Res. **18,** 5905–5906.

Shaw, D.R., Richter, H., Giorda, R., Ohmachi, T., and Ennis, H.L. (1989). Nucleotide sequences of *Dictyostelium discoideum* developmentally regulated cDNAs rich in (AAC) imply proteins that contain clusters of asparagine, glutamine, or threonine. Mol. Gen. Genet. **218,** 453–459.

Struhl, K. (1985). Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. USA **82,** 8419–8423.

Thanos, D., and Maniatis, T. (1992). The high mobility group protein HMG I(Y) is required for NF-κB-dependent virus induction of the human IFN-β gene. Cell **71,** 777–789.

Theologis, A., Huynh, T.V., and Davies, R.W. (1985). Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. J. Mol. Biol. **183,** 53–68.

Thompson, W.F., and White, M.J. (1991). Physiological and molecular studies of light-regulated nuclear genes in higher plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **42,** 423–466.

Vinson, C.R., LaMarco, K.L., Johnson, P.F., Landschulz, W.H., and McKnight, S.L. (1988). *In situ* detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. **2,** 801–806.

Wang, Y., Zhang, W., Cao, J., McElroy, D., and Wu, R. (1992). Characterization of *cis*-acting elements regulating transcription from the promoter of a constitutively active rice actin gene. Mol. Cell Biol. **12,** 3399–3406.

Watson, J.C., and Thompson, W.F. (1986). Purification and restriction endonuclease analysis of plant nuclear DNA. Methods Enzymol. **118,** 57–75.

Wu, R.S., Panusz, H.T., Hatch, C.L., and Bonner, W.M. (1986). Histones and their modifications. CRC Crit. Rev. Biochem. **20,** 201–263.

Zhao, K., Käs, E., Gonzalez, E., and Laemmli, U.K. (1993). SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. EMBO J. **12,** 3237–3247.