A Repressor with Similarities to Prokaryotic and Eukaryotic DNA Helicases Controls the Assembly of the CAAT Box Binding Complex at a Photosynthesis Gene Promoter*

Received for publication, December 5, 2000, and in revised form, March 21, 2001
Published, JBC Papers in Press, March 26, 2001, DOI 10.1074/jbc.M010945200

Staver Bezhani, Irena Sherameti, Thomas Pfannschmidt, and Ralf Oelmüller‡

From the Institut für Allgemeine Botanik, Lehrstuhl für Pflanzenphysiologie, Friedrich-Schiller-Universität Jena, Dornburger Strasse 159, D-07743 Jena, Germany

A single nucleotide exchange in a promoter region located immediately upstream of the CAAT box of the spinach photosynthesis gene AtpC (gene product is subunit γ of the chloroplast ATP synthase) prevents the formation of a secondary structure and causes an unregulated, constitutive high level of expression (Kusnetsov, V., Landsberger, M., Oelmüller, R. (1999) J. Biol. Chem. 274, 36009–36014). We have isolated cDNAs for ATPC-2, a new polypeptide with homologies to pro- and eukaryotic helicases, which specifically binds to this promoter region. Binding of ATPC-2 competes strongly with that of a CAAT box binding factor (CBF), consistent with the idea that both complexes cannot be formed simultaneously because of steric reasons. In gel mobility shift assays, high binding activities of ATPC-2 and low binding activities of CBF were observed with nuclear extracts from tissue with low AtpC expression levels, and the opposite was observed with extracts from tissues with high AtpC expression levels. Binding of ATPC-2 to the mutant sequence, which directs a constitutively high level expression in vivo and prevents the formation of a secondary structure in vitro, is significantly weaker than binding to the wild-type sequence. Again, the opposite results were obtained for the CBF. Thus, we conclude that the assembly of the CBF-DNA complex stimulates transcription of AtpC and that CBF binding is prevented if ATPC-2 is bound to the promoter region. The novel mechanism of gene regulation and the role of the helicase-like protein ATPC-2 as a potential transcriptional repressor is discussed in relation to its modular structure.

The plastid ATP synthase of higher plants and algae consists of nine different subunits, and three of them are encoded by the nuclear genes AtpC, AtpD, and AtpG (the gene products are the subunits γ and δ of the CF1 moiety and CF1-II) (2). Comparable with other nuclear-encoded genes for plastid proteins, expression of the spinach ATP synthase genes and chimeric promoter::reporter gene fusions in transgenic tobacco is strongly regulated by light, phytohormones, in particular cytokinins, or the stage of the plastids in cells in which these genes are expressed (cf. Refs. 3–7). However, unusually for photosynthesis genes, the essential cis-elements determining the regulated expression of AtpC are positioned in close vicinity to the respective transcription start sites (4), and crucial nucleotides for the regulated expression appear to be located immediately upstream of the CAAT box (4). This region also forms a complex with a CAAT box binding factor, CBF* (1). However, in contrast to many CBFs from metazoa (8–10) the binding activity in in vitro assays is regulated, and the binding activities correspond to the promoter activities in vivo (1). Surprisingly, a single nucleotide exchange uncoupled the promoter activity from regulatory pathways and resulted in a high constitutive expression. This indicates that low promoter activities (e.g. in darkness or in photobleached tissue) are caused by an inhibitory effect that is no longer active under conditions of high promoter activities, e.g. in light or after cytokinin treatments. Here we describe the isolation of a DNA that codes for a protein with sequence-specific binding characteristics to this promoter region. We demonstrate that the binding activity is high in tissue with low AtpC expression levels and, vice versa, that this behavior is opposite to the binding features described previously for CBF (1), and that the recombinant protein competes with CBF for binding. The data are consistent with a model in which this protein functions as a repressor by preventing the assembly of CBF at the CAAT box and thus repressing AtpC transcription.

**EXPERIMENTAL PROCEDURES**

**Plant Growth**—Tobacco seeds were surface-sterilized, planted on 1/2 Murashige and Skoog medium supplemented with 2% sucrose in the presence or absence of Norflurazon ($10^{-6}$ M) and kept in a cold room in darkness for 2 days to synchronize germination. Seedlings were either kept in darkness for 10 days at 22 °C or they were transferred to white light for 40 h prior to harvesting in the absence or presence of Norflurazon. Alternatively, etiolated seedlings were transferred to cytokinin-containing plates ($N^\text{6}$-benzylaminopurine, $10^{-5}$ M) in darkness 40 h prior to harvest.

**Double-stranded Oligonucleotides, Nuclear Extracts from Tobacco, Southwestern Hybridization**—Oligonucleotides for wild-type and mutant sequences were annealed (11) and cloned into the Smal site of pBSC-1 (Stratagene, San Diego) prior to sequencing. Plasmids were chosen in which the 5'-end of the promoter fragment was oriented toward the BamHI site of pBSC-1. Thus all promoter fragments had the following 5'- and 3' extensions: 5'-GATCCCC[Insert]GGGCTGCA-3' and 3'-GGGG[Insert]CCCG-5'. After digestion of the plasmids with PstI and BamHI, the BamHI site was filled-in with all four radiolabeled nucleotides and the Klenow enzyme, and the fragments were gel-purified on agarose gels.

For the binding studies in the presence of CBF-C (1), nuclear protein extracts were isolated from six-week-old tobacco or Arabidopsis plants grown in the greenhouse. For physiological studies, seedlings were grown as described above, and the nuclear protein fractions were prepared as described (12).

---

* This work was supported by the German Research Foundation and the "Fond der Chemischen Industrie." The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 0049-3641-949231; Fax: 0049-3641-949232; E-mail: Ralf.Oelmuller@rz.uni-jena.de.

1 The abbreviations used are: CBF, CAAT box binding factor; GST, glutathione S-transferase.
A Repressor Prevents CAAT Box Complex Assembly

Approximately 190,000 phases of an Arabidopsis thaliana cDNA expression library were screened with the radiolabeled AtPC promoter region (−68 to −39; see below). Two positive phases were plaque-puriﬁed, and pBSC was excised according to manufacturer’s instructions (Stratagene) prior to sequencing.

DNA sequence hybridization was performed at 4 °C. The filters were ﬁrst blocked with blocking buffer by gentle shaking for 3 h (SW buffer, 1% bovine serum albumin, 1.5 mg/ml denatured salmon sperm DNA; SW buffer = 25 mM HEPES-NaOH, pH 7, 6, 50 mM NaCl, 5 mM MgCl₂, 1 mM EDTA, 5% γ-irradiated glycerol, 1 mM dithiothreitol). After three washing steps (10 min each) with HS buffer (SW buffer with 1 mg/ml denatured salmon sperm DNA), the ﬁlters were incubated overnight with SW buffer in the presence of the radiolabile DNA fragments. After washing with SW buffer, the signals were visualized on a PhosphorImage (Storm 820, Molecular Dynamics, Krefeld, Germany).

Overexpression and Puriﬁcation of GST Fusion Proteins, Filter Binding Assays—The isolated cDNA as well as the cDNA encoding CBF-C (1) were N-terminally fused to GST according to the manufacturer’s instructions (pGEX vector system from Amersham Pharmacia Biotech). Plasmids encoding GST fusion proteins were transformed into Escherichia coli strain BL21 (DE3) or S30. Cultures were grown overnight at 37 °C in the presence of 40 μg/ml chloramphenicol and 100 μg/ml ampicillin, diluted 100-fold, and grown at 30 °C to an optical density of 0.6 at 600 nm. Protein expression was induced by isopropyl-β-D-thiogalactopyranoside, with a ﬁnal concentration of 0.5 mM at 27 °C. Cells were harvested 4 h after induction, resuspended in 1× phosphate-buffered saline, and sonicated. After centrifugation (12,000 g for 15 min at 4 °C), aliquots of the supernatant were applied directly onto freshly prepared glutathione-Sepharose 4B columns and further processed according to the manufacturer’s instructions (Amersham Pharmacia Biotech). The puriﬁed GST and GST fusion proteins were separated by SDS-polyacrylamide gel electrophoresis (12%) and transferred onto nitrocellulose. Southern analysis was performed as described previously (1) with the radiolabeled DNA fragments. For gel mobility shift assays, equal amounts of the puriﬁed proteins were used.

Gel mobility shift assays were performed essentially as described (1, 12). CBF-C was incubated with nuclear protein extracts (1 μg) in binding buffer (25 mM HEPES/ROH, pH 7, 5, 50 mM KCl, 5 mM MgCl₂, 5% glycerol) for 30 min at 37 °C. Finally, the DNA was added and the mixture incubated at 37 °C for 2 h prior to gel electrophoresis. For competition studies, the puriﬁed AtPC-2 GST fusion protein, the CBF-C GST fusion protein, or GST alone was added to the preformed radiolabeled protein-DNA complex and treated as described in the ﬁgure legends. In all of the studies the CBF-C fusion was preincubated with nuclear extract as described above, whereas the AtPC-2 fusion was added directly.

RESULTS

The AtPC Promoter Segment—69 to −39 Speciﬁcally Interacts with AtPC-2—A double-stranded oligonucleotide from the −68 to −39 AtPC region (5′-TTTACCTCAGAAATCTAATTGGCAAAATCT-3′) harboring the expression-relevant AAAAT (bold) and the CAAT (italics) motifs was used initially to screen an expression library from Arabidopsis. Positive signals were obtained only when the screen was performed in the presence of 1 mg/ml nuclear extract. The four isolated cDNAs encoded CBF-C, the subunit C of a new CAAT box binding complex isolated from plants (1). Therefore, a second screen was initiated with an oligonucleotide in which the CAAT sequence was replaced by GATTA, an oligonucleotide that fails to bind to CBF-C (1). Two positive phases with signiﬁcantly lower binding strength were obtained. Both of them also bind to the wild-type sequence harboring the CAAT motif but not to a mutant sequence with a single nucleotide exchange (A → G) in the AAAAT sequence (data not shown; see below). DNA sequence analysis revealed that both phases carried identical DNA insertions that code for the C-terminal part of a novel protein named AtPC-2. After the complete nucleotide sequence of this gene became available in the data banks, we polymerase chain reaction-ampliﬁed the complete cDNA from our library and conﬁrmed that the protein is encoded by a genomic sequence without introns (Fig. 1). A homologous cDNA was also isolated from our tobacco cDNA library; however, it is not full-length (data not shown).

Data bank searches revealed that the deduced amino acid sequence exhibits homologies to DNA helicases from pro- and eukaryotic organisms (Fig. 2). Surprisingly, the strongest homologies were observed to the immunoglobulin S Mu-binding protein 2 from mouse, a DNA helicase from human, and a putative DNA-binding protein from yeast (GenBank™ accession numbers P40694, L24544, and Z98951, respectively), whereas an Arabidopsis protein (GenBank™ accession number AB026643) with an unknown function exhibits less sequence similarity. Interestingly, the isolated protein also exhibits striking similarities to various bacterial proteins (amino acid sequence identity > 30%); however, the function of these proteins has not yet been determined.

Motif searches did not reveal any obvious functions of AtPC-2 except that the protein harbors 29 (partially overlapping) putative phosphorylation sites for signal transduction components (Fig. 1). Closer inspection uncovered that the similarities to the mouse, human, and yeast proteins include the N-terminal ﬁve well characterized motifs known from DNA helicases (boxes I, II, III, and IV, Fig. 2). The protein as well as eight homologues with unknown function that are present in the data banks end behind box IV (Fig. 2). Furthermore, the C-terminal parts of the latter protein group contain a segment of −60 amino acids exhibiting between 65 and 85% amino acid sequence identity to each other (data not shown). This suggests that the Arabidopsis polypeptide belongs to a novel class of proteins present in various species and that this class of proteins exhibits striking similarities to the N termini of some DNA helicases (Fig. 2; cf. “Discussion”). In addition, a phylogenetic tree indicates that the eukaryotic proteins are closely related to each other, whereas the bacterial proteins appear to be more distantly related (data not shown).

Binding of AtPC-2 to the AtPC Promoter in Vitro—To test the DNA sequence requirement for the formation of the AtPC-2 DNA complex, a series of mutant sequences for which the transcriptional activity is known from chimeric promoter::transcriptional activity is known from chimeric promoter::uidA gene fusions in transgenic tobacco (1) was tested for its ability to bind to AtPC-2 and CBF-C. Both proteins were overexpressed in E. coli as GST fusion proteins, isolated ﬁrst by glutathione 4B-Sepharose chromatography and then by SDS-gel electrophoresis (Fig. 3A). The isolated proteins were separated by gel before transfer to nitrocellulose membranes. Filter binding assays (Fig. 3B) in the absence (left panels) or presence (right panels) of nuclear extracts demonstrate that the wild-type sequence (AAAAATCTAAT, top panels) binds to AtPC-2 in both instances. As previously described (1), binding of CBF-C requires the addition of nuclear extracts (Fig. 3B, right top panel). The mutant sequence with an A → G exchange (AAGATCTAAT) fails to bind to AtPC-2, whereas CBF-C is still capable of binding in the presence of nuclear extracts.
motifs are arbitrary (the boxes show residues that are not consistent with the consensus amino acid sequences. The relative positions of the motifs and spacing between residue (Ser, Thr, Asp, Glu, Asn, Gln, Lys, and Arg); x, residue not restricted to hydrophobic or hydrophilic. The lowercase letters in the ATPC-2 boxes show residues that are not consistent with the consensus amino acid sequences. The relative positions of the motifs and spacing between motifs are arbitrary (cf. Ref. 30). The conserved region between boxes IV and V of the helicase superfamily I is indicated.

Fig. 2. Schematic presentation of the modular organization of helicases belonging to superfamily I (top row) and ATPC-2 (bottom row). The boxes represent the conserved helicase motifs, and letters inside the boxes are the consensus amino acid sequences of each motif. The labels above the boxes are the names assigned to the motifs. +, hydrophobic residue (Ile, Leu, Val, Phe, Tyr, and Trp); o, charged or polar residue (Ser, Thr, Asp, Glu, Lys, and Arg); x, residue not restricted to hydrophobic or hydrophilic. The lowercase letters in the ATPC-2 boxes show residues that are not consistent with the consensus amino acid sequences.

Fig. 3. DNA binding activities of promoter and mutant sequences to ATPC-2 and CBF-C. A, Coomassie Blue-stained gel showing the purified CBF-C-GST fusion protein (1), the ATPC-2-GST fusion protein (2), and GST alone (3). B, DNA filter binding assays with the three proteins shown in A. Equal amounts of protein (10 mg) were loaded per lane, separated electrophoretically, and transferred to nylon membranes prior to incubation with radiolabeled oligonucleotides and autoradiography. Filters shown on the left side of panel B were incubated with the radiolabeled oligonucleotide alone and those shown on the right with the oligonucleotide plus nuclear extract (1 mg/ml solution). Oligonucleotides used for the incubation are: top panels, AAAATTC; middle panels, AAGATTCAAT; bottom panels, AAAATTGTTA.

In contrast, the CAAT mutant sequence (AAAAATTGTTA) fails to bind to CBF-C, whereas it does bind to ATPC-2 (bottom panels). This was confirmed by gel mobility shift assays. Fig. 4 demonstrates that the binding activity is not affected by the CAAT box mutation (compare lanes 2 and 3 with lanes 4 and 5) and that the binding activity of ATPC-2 to the oligonucleotide with the A → G exchange is completely lost (lanes 6 and 7). A shorter oligonucleotide, which contains the AAAATT sequence but lacks the CAAT box sequence, also fails to bind to ATPC-2 (lanes 8 and 9); this indicates that ATPC-2 binding to the AtPC promotor is sequence-specific and requires additional nucleotides located 3’ to the AAAATT motif. Furthermore, we performed competition experiments between purified ATPC-2 and CBF with the wild-type oligonucleotide sequence (Fig. 5). A 10-fold molar excess of the ATPC-2-GST fusion protein cannot displace CBF-C-GST from the oligonucleotide, whereas the same molar excess of CBF-C-GST is sufficient for self-competition (Fig. 5, top panel). The same is true for the opposite experiment; if DNA preloaded with ATPC-2-GST is used for competition studies; again the binding activity cannot be competed with a 10-fold molar excess of CBF-C/GST, whereas ATPC-2-GST is effective (Fig. 5, bottom panel). This indicates that both binding activities are tight and that either ATPC-2 or CBF can bind to this promotor region. The assembly of both complexes simultaneously is not possible for stererical reasons (cf. “Discussion”).

A correlation of the promoter activity in vivo (1, 3) and its capability to bind to ATPC-2 is presented in Fig. 6. Nuclear extracts from light-grown seedlings showed little binding activity to the wild-type oligonucleotide, whereas extracts from dark-grown seedlings showed substantially higher retardation signals. The same correlation between a low binding activity and a high promoter activity was observed with nuclear ex-
tracts from etiolated seedlings treated with cytokinin (high promoter activity (1)) and from light-grown seedlings treated with the herbicide Norflurazon (lane 3), or dark-grown seedlings with cytokinin (lane 4). Lane 5, Control: extracts from 6-week-old tobacco plants grown in the greenhouse, which were dark-adapted for 2 weeks. F, free fragment.

**DISCUSSION**

In many eukaryotic class II promoters, CCAAT motifs are often found between 50 and 100 nucleotides upstream of the transcription start site (13–16). These motifs are recognized by CCAAT-binding proteins. The CAAAT box sequences in metazoa, which are believed to influence the frequency of transcriptional initiation, can be the target site for regulation, and the assembly of CBP at the CAAAT motif occurs in response to cell-internal or -external signals (8, 17–28). Here we demonstrate that the binding activity of CBP to the CAAAT motif in the ATPC promoter is controlled by ATPC-2. Whenever the promoter is activated, can be the target site for regulation, and the assembly of CBP, resulting in a diminished ATPC transcription. Both factors cannot bind simultaneously because of their overlapping binding regions (1). This concept is supported by crystal structural data from the yeast CBP, which protects a DNA region of at least 120 nucleotides (8). Activators of ATPC transcription, such as light or cytokinin, prevent repressor binding and thus allow for the assembly of CBP, this is consistent with the observation that the A → G mutation, which prevents binding of the repressor but allows the assembly of CBP (Fig. 3), results in a constitutive promoter activity in vitro (1). This novel mechanism of gene regulation might add a new facet to the observations that etiolation is caused by processes that actively prevent the development observed in light (38).

**Acknowledgments**—We thank H. Renz and F. Wicke for technical assistance and Susan Herzog for critically reading the manuscript.

**REFERENCES**

1. Kusnetsov, V., Landsberger, M., Oelmuller, R. (1999) J. Biol. Chem. **274**, 36009–36014

2. Herrmann, R. G., Oelmuller, R., Bichler, J., Schneiderbauer, A., Steppuhn, J., Wedel, N., Tyagi, A. K., and Westhoff, P. (1994) in *Plant Molecular Biology* (Herrmann, R. G., and Larkins, B. A., eds) Vol. 2, pp. 411–427, Plenum Press, New York

3. Bolle, C., Sopory, S., Luberstedt, T., Klosgen, R. B., Herrmann, R. G., and
A Repressor Prevents CAAT Box Complex Assembly

23789

Oelmüller, R. (1994) Plant Physiol. 105, 1355–1364
4. Bolle, C., Kusnetsov, V. V., Herrmann, R. G., and Oelmüller, R. (1996) Plant J. 9, 21–30
5. Kusnetsov, V. V., Oelmüller, R., Sarwat, M. I., Porfirova, S. A., Cherepneva, G. N., Herrmann, R. G., and Kulaeva, O. N. (1994) Planta 194, 318–327
6. Kusnetsov, V. V., Bolle, C., Lubberstedt, T., Sopory, S., Herrmann, R. G., and Oelmüller, R. (1996) Mol. Gen. Genet. 252, 631–639
7. Oelmüller, R., Schneiderhauer, A., Herrmann, R. G., and Kruppstech, K. (1995) Mol. Gen. Genet. 245, 478–484
8. Maity, S. N., and de Crombrugghe, B. (1998) Trends Biochem. Sci. 23, 174–178
9. McNabb, D. S., Xing, Y., and Guarente, L. (1994) Genes Dev. 9, 47–58
10. Sinha, S., Maity, S. N., Lu, J., and de Crombrugghe, B. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1624–1628
11. Lubberstedt, T., Oelmüller, R., Wanner, G., and Herrmann, R. G. (1994) Mol. Gen. Genet. 242, 602–613
12. Oelmüller, R., Bolle, C., Tyagi, A. K., Niekraszewitz, N., Breit, S., and Herrmann, R. G. (1993) Mol. Gen. Genet. 237, 261–272
13. Dorn, A., Bollekens, J., Staub, A., Benoist, C., and Mathis, D. (1987) Cell 50, 863–872
14. Chodosh, L. A., Baldwin, A. S., Carthew, R. W., and Sharp, P. A. (1988) Cell 53, 11–24
15. Hatamoto, A., Golumbek, P. T., Van Schaftingen, E., and de Crombrugghe, B. (1988) J. Biol. Chem. 263, 5940–5947
16. Bucher, P. (1990) J. Mol. Biol. 212, 563–578
17. Maity, S. N., Golumbek, P. T., Karsenty, G., and de Crombrugghe, B. (1988) Science 241, 582–585
18. Maire, P., Wuarin, J., and Schibler, U. (1989) Science 244, 343–346
19. Hasegawa, T., Zhou, X., Garrett, L. A., Ruteshouser, E. C., Maity, S. N., and de Crombrugghe, B. (1996) Nucleic Acids Res. 24, 3253–3260
20. Brown, A. M., Wright, K. L., and Ting, J. P. (1993) J. Biol. Chem. 268, 26328–26333
21. Goldsmith, M. E., Madden, M. J., Morrow, C. S., and Cowan, K. H. (1993) J. Biol. Chem. 268, 5856–5860
22. Hung, H. L., and High, K. A. (1996) J. Biol. Chem. 271, 2323–2331
23. Li-Weber, M., Davydov, I. V., Krafft, H., and Kramer, P. H. (1994) J. Immunol. 153, 4122–4133
24. Framson, P., and Bornstein, P. (1993) J. Biol. Chem. 268, 4989–4996
25. Wright, K. L., Vilen, B. J., Itzhak-Lindstrom, Y., Moore, T. L., Cricchiello, M., Cogswell, P., Clarke, J. B., and Ting, J. P. (1994) EMBO J. 13, 4042–4053
26. Jackson, S. M., Ericsson, J., Osborne, T. F., and Edwards, P. A. (1995) J. Biol. Chem. 270, 21445–21448
27. Roy, B., Li, W. W., and Lee, A. S. (1996) J. Biol. Chem. 271, 28995–29002
28. Tezuka, K., Denhardt, D. T., Rodan, G. A., and Harada, S. (1996) J. Biol. Chem. 271, 2271–2277
29. Hall, M. C., and Matson, S. W. (1999) Mol. Microbiol. 34, 867–877
30. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989) Nucleic Acids Res. 17, 4713–4729
31. Mizuta, T.-R., Fukita, Y., Miyoshi, T., Shimizu, A., and Honjo, T. (1993) Nucleic Acids Res. 21, 1761–1766
32. Lee, D. G., and Bell, S. P. (2000) Curr. Opin. Cell Biol. 12, 280–285
33. Bird, L. E., Subramanya, H. S., and Wigley, D. B. (1998) Curr. Opin. Struct. Biol. 8, 14–18
34. Mariano, R. J. (1997) Structure 5, 1129–1134
35. Baker, T. A, and Bell, S. P. (1998) Cell 92, 295–305
36. Soultanas, P., and Wigley, D. B. (2000) Curr. Opin. Struct. Biol. 10, 124–128
37. Fukita, Y., Mizuta, T.-R., Shiraizu, M., Ozawa, K., Shimizu, A., and Honjo, T. (1993) J. Biol. Chem. 268, 17463–17470
38. Fankhauser, C., and Chery, J. (1997) Annu. Rev. Cell Dev. Biol. 13, 203–229