Planta (1997) 203: 373–380

Planta

Identification and purification of a spinach chloroplast DNA-binding protein that interacts specifically with the plastid *psaA-psaB-rps14* promoter region

M.C. Cheng[1], S.P. Wu[2], L.F.O. Chen[1], S.C.G. Chen[1]

[1]Institute of Botany, Academia Sinica, Taipei, Taiwan, Republic of China
[2]Institute of Biochemistry, National Yang-Ming University, Taipei, Taiwan, Republic of China

Received: 12 February 1997 / Accepted: 13 March 1997

Abstract. We have previously shown the presence in chloroplasts of sequence-specific DNA-binding proteins that interact specifically with two regions located downstream and upstream from the 5′-transcription start site of the plastid *psaA-psaB-rps14* operon. As part of an effort to elucidate the regulatory mechanism of plastid transcription during plant development, we report here the purification and characterization of the chloroplast DNA-binding protein from spinach (*Spinacia oleracea* L. var. *spinosa* Ashers et Graeden) leaves that specifically recognizes sequences between positions +64 to +83 relative to the transcription start site. This DNA-binding protein has been highly purified from chloroplasts by using a combination of high-salt extraction, ammonium sulfate precipitation, heparin-agarose chromatography, and sequence-specific DNA-affinity chromatography. The protein exhibited an apparent molecular weight of 59–60 kDa on the basis of gel filtration. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by Southwestern blot analysis further indicated that this DNA-binding protein is dimeric and composed of two ≈31-kDa subunits. We discuss the properties of this protein in relation to the known chloroplast DNA-binding factors for plastid gene expression.

Key words: Chloroplast – DNA-binding protein (purification) – Operon (*psaA*) – *Spinacia* (chloroplast) – Trans-acting factor – Transcription

Introduction

The biosynthesis and assembly of the photosynthetic apparatus, including photosystems I and II in higher-plant chloroplasts require the coordinated interaction of nuclear and plastid genomes. In green plants, the photosystem I complex consists of more than 13 individual polypeptides, which bind the reaction center P700, electron acceptors A_0 and A_1, iron-sulfur centers, and light-harvesting chlorophyll pigments (Malkin 1982; Parrett et al. 1989). The two largest core subunits within the complex (82–84 kDa), carrying P700 and approx. 90–100 molecules of chlorophyll *a*, are believed to dimerize through a leucine zipper to form the iron-sulfur center F_X (Kössel et al. 1990; Webber and Malkin 1990). The genes for the two subunits, *psaA* and *psaB*, are located adjacent to each other in the large single-copy region of the circular plastid genome in higher plants (Fish et al. 1985; Lehmbeck et al. 1986). Gene *psaB* is followed by *rps14* encoding the chloroplast ribosomal protein CS14 (Kirsch et al. 1986; Shinozaki et al. 1986; Chen et al. 1990). The *psaA-psaB-rps14* gene cluster was found to co-transcribe into a 5- to 6-kb polycistronic mRNA in spinach (Westhoff et al. 1983), tobacco (Meng et al. 1988), and rice (Chen et al. 1992). We have performed a detailed transcriptional analysis of the promoter of the rice *psaA-psaB-rps14* operon with deleted mutants in vitro. Two functional promoters denoted as "−175" and "−129" were revealed. However, only the "−175" promoter was utilized in vivo to initiate transcription at 175 bp upstream of rice *psaA* (Chen et al. 1993). The *Esherichia coli*-like "−10" (TAATAG) and "−35" (TTAGGC) promoter elements, spaced by a distance of 17 bp, were found in the region upstream from the transcriptional initiation site −175 (Chen et al. 1992; Meng et al. 1988).

The chloroplast transcription apparatus has been shown to contain a basic organelle-encoded enzymatic machinery surrounded by factors that may be encoded by nuclear genes (reviewed by Igloi and Kössel 1992). An *E. coli*-like RNA polymerase, with core subunits encoded by plastid *rpo* genes, has been well reported (Little and Hallick 1988; Hu and Bogorad 1990; Hu et al. 1991). Promoter selection by the plastid-encoded enzyme is dependent on nuclear-encoded sigma-like factors, and the transcription activity may be modulated by nuclear-encoded transcription factors (see reviews by Gruissem

Correspondence to: S.C.G. Chen;
E-mail: boscchen@ccvax.sinica.edu.tw;
Fax: 886(2)7827954; Tel: 886(2)7899590 ext 426

and Tonkyn 1993; Link 1994; Troxler et al. 1994). Recent evidence points to the existence of another type of chloroplast RNA polymerase. Using targeted deletion of the plastid *rpoB* gene in tobacco, Allison et al. (1996) demonstrated the presence of non-*E. coli*-like transcription machinery which preferentially transcribes genetic-system genes. As yet, the intracellular genetic origin, subunit composition, and promoter specificities of this non-*E. coli*-like transcription system are uncertain (Hess et al. 1993; Lerbs-Mache 1993; Pfannschmidt and Link 1994; Allison et al. 1996). Furthermore, evidence for several chloroplast sequence-specific DNA-binding proteins which may serve as positive or negative transcriptional regulators has recently emerged (Lam et al. 1988; Zaitlin et al. 1989; Baeza et al. 1991; Sato et al. 1993; Kim and Mullet 1995; To et al. 1996), indicating the complicated regulation of plastid transcription during plant development.

We have previously demonstrated the existence of *psaA* promoter-specific binding proteins in the chloroplasts by competitive gel retardation assay (Chen et al. 1993). Two binding regions were mapped upstream (termed region U) and downstream (termed region D) of the transcription start site by exonuclease III protection analysis of the DNA fragment containing −95 to +223 of the rice *psaA* operon. In order to obtain greater insight into the possible regulatory role of the chloroplast sequence-specific DNA-binding proteins involved in the leaf-specific expression of the plastid *psaA* operon, we have made efforts to identify and isolate the *psaA* promoter-specific binding protein from higher plants, including rice, tobacco and spinach. Since it is easier to isolate chloroplasts from spinach than from other plants, and with better yield, the present work describes the purification and biochemical characterization of a chloroplast protein from spinach leaves that specifically binds to the 5′-untranslated region D (from positions +64 to +83) on the spinach *psaA* operon.

Materials and methods

Isolation of spinach chloroplast DNA and RNA. Intact chloroplasts were isolated essentially as described by Orozco et al. (1985) and To et al. (1996). Deveined spinach (*Spinacia oleracea* L. var. *spinosa* Ashers et Graeden) leaves (200 g) were homogenized in ten times the volume of chilled grinding buffer [50 mM Hepes (pH 8.0), 1 mM $MgCl_2$, 1 mM $MnCl_2$, 2 mM EDTA, 330 mM sorbitol, 5 mM sodium ascorbate]. After filtration through two layers of Miracloth (Calbiochem, La Jolla, Calif., USA), the filtrate was briefly centrifuged at 4500 · *g*. The pellet was suspended in a small volume of grinding buffer and overlaid on top of a 40%/85% Percoll step gradient in grinding buffer. After centrifugation at 4000 rpm for 15 min with an SRP28 rotor (Hitachi, Tokyo, Japan), the green interface layer of chloroplasts was collected. The intact chloroplasts were then subjected to DNA or RNA isolation as described in the previous reports (Chen et al. 1990, 1992).

Sequencing of DNA and primer extension. The upstream region of the spinach *psaA* operon was isolated by polymerase chain reaction (PCR) cloning and sequenced in both directions by the dideoxy method (Maniatis et al. 1989). The 5′ end of the transcript for the *psaA* operon was precisely determined by primer extension. The primer corresponding to the sequence complementary to the region extending from +66 to +90 (see Fig. 1) was 5′-end-labeled with [γ-^{32}P]ATP (Amersham International, Bucks., UK), annealed with spinach chloroplast RNA, and extended using AMV reverse transcriptase (Promega Corporation, Madison, Wis., USA). The product was analyzed by electrophoresis through a 6% polyacrylamide-8 M urea sequencing gel and subsequent autoradiography on Kodak X-Omat AR film. Products of the sequencing reaction with the same primer were run in parallel to generate DNA sequence ladders.

Gel retardation assay. For preparation of the binding probe, the synthetic oligonucleotide and its complementary strand were annealed in TNE buffer [10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 1 mM EDTA] by sequential heat treatment (88 °C, 2 min; 65 °C, 10 min; 37 °C, 10 min) and finally placed at room temperature for 15 min. The double-stranded oligomer was then end-labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase, followed by purification through a Centricon-3 column (Amicon, Beverly, Mass., USA).

The binding reaction mixture (20 µl) consisted of the end-labeled oligonucleotide probe (1 ng) and spinach chloroplast protein fraction in binding buffer [10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 50 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 0.1% Nonidet P-40, 5% glycerol) containing 2 µg of nonspecific competitor DNA poly(dI-dC) · poly(dI-dC) (Pharmacia LKB Biotechnology, Uppsala, Sweden). For the competition binding assay, varying amounts of cold-annealed oligonucleotide were included in the reaction. Following incubation at 25 °C for 30 min, the DNA-protein complexes were analyzed on a native 5% polyacrylamide gel (acrylamide:bisacrylamide 29:1) in 0.5 × TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8.0) and subsequent autoradiography. One unit of binding activity is defined as the amount of protein required to retard 0.1 ng of labeled probe.

Isolation of sequence-specific DNA-binding factor from spinach chloroplasts. All the isolation and purification procedures were conducted at 0–4 °C. The chloroplasts isolated from 200 g of deveined spinach leaves were lysed by adding five times the volume of lysis buffer [20 mM Hepes (pH 8.0), 5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine, 5 mM ε-amino-*n*-caproic acid] and centrifuged at 4000 · *g* to collect the pellet. The pellet was extracted with lysis buffer containing 1 M NaCl and 15% glycerol for 30 min at 4 °C, followed by ultracentrifugation at 100 000 · *g* for 30 min. The supernatant containing about 65 mg protein was then subjected to ammonium sulfate fractionation. The 25–65% fraction possessing both region-U- and region-D-binding activities was resuspended and dialyzed overnight against buffer H [30 mM sodium phosphate (pH 6.8), 0.5 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT, and 10% glycerol]. The resulting protein preparation (about 30 mg) was then stored in aliquots at −70 °C for further purification. The protein concentration was estimated by the Bradford method (1976).

Purification of region-D-binding protein by affinity column chromatography. The above-described ammonium sulfate fraction (30 mg protein) was then loaded onto a column of heparin-agarose gel (Bio-Rad Laboratories, Hercules, Calif., USA; 30-ml bed volume) for purification of region-D-binding protein. The column was thoroughly washed with buffer H to remove the unbound proteins. The column was sequentially eluted with buffer H containing 0.3 M, 0.5 M, 0.7 M, and 1.0 M KCl. Each fraction was dialyzed against buffer Z [25 mM Hepes (pH 7.6), 12.5 mM $MgCl_2$, 1 mM DTT, 0.1% Nonidet P-40, and 20% glycerol), followed by concentration through a Centricon-3 column.

The 0.3 M KCl fraction from the heparin-agarose column was found to contain most of the region-D-binding activity and was thus utilized for further purification on a sequence-specific affinity column. The affinity resin was prepared by coupling the catenated oligo D20 (see Fig. 1) to Sepharose 4B following the protocol of Wu et al. (1988) with some modification. The annealed oligo D20 with a GATC overhang at the 5′ ends (460 µg) was phosphorylated

and ligated to approx. 10- to 15-mers, followed by chemical coupling to CNBr-activated Sepharose 4B (Pharmacia). The procedure for DNA affinity column chromatography was essentially the same as described by Kadonaga (1991). The packed Bio-Rad Econo column (1-ml bed volume) was pre-equilibrated with buffer Z containing 0.1 M KCl. The protein fraction purified from the heparin-agarose column (5 mg) in 200 µl of buffer Z containing 0.1 M KCl was preincubated with poly(dI-dC) · poly(dI-dC) (final concentration 60 µg · ml^{-1}) at 0 °C for 10 min. Subsequently the mixture was passed through the column slowly at a flow rate of 3.5 ml · h^{-1}, and the column was washed with 8 ml of the starting buffer to remove the unbound proteins. Region-D-binding activity was then eluted with 1.4 ml of buffer Z containing 1.0 M KCl. The eluate was desalted with Centricon-3, adjusted to 0.1 M KCl, and mixed again with competitor DNA at 35 µg · ml^{-1} for reapplication to the second round of the sequence-specific affinity column (0.5-ml bed volume).

Analytical fast protein liquid chromatography. An FPLC system with a Mono Q HR5/5 column (Pharmacia) was employed to dissect the region-U- and region-D-binding activities of the chloroplast proteins. The 0.3 M KCl fraction from the heparin-agarose column (total protein 500 µg) was loaded onto the column and eluted with a linear 0–500 mM KCl gradient in 20 mM Tris-HCl (pH 7.5) at a flow rate of 1.0 ml · min^{-1}. Each fraction was desalted and concentrated with Centricon-3, and then assayed for the oligoprobe-binding activities.

The apparent molecular weight of region-D-binding protein was estimated by FPLC using a Superose-12 HR 10/30 column (Pharmacia). The column was pre-calibrated with standard proteins including blue dextran, aldolase, bovine serum albumin, ovalbumin, chymotrypsinogen A and ribonuclease A. The 0.3 M KCl fraction from the heparin-agarose column (total protein 2.5 mg) was then loaded onto the column and eluted with 50 mM sodium phosphate buffer (pH 7.0) containing 150 mM NaCl at a flow rate of 0.2 ml · min^{-1}. Each fraction was concentrated and assayed for region-D-binding activity. The molecular weight of the activity-peak fraction was estimated according to K_{av} versus log MW standard curve.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis and Southwestern blot analysis of region-D-binding protein. The final region D-binding protein preparation from the second round of sequence-specific affinity column was pretreated with denaturation buffer [60 mM Tris-HCl (pH 6.8), 2% SDS, 5% 2-mercaptoethanol, 10% glycerol and 0.01% bromophenol blue], followed by electrophoresis on two identical SDS-polyacrylamide gels (14%). The molecular size markers (Sigma Chemical Co., St. Louis, Mo., USA) were run in parallel with the sample. After running, the gel was silver-stained and a duplicate gel was subjected to Southwestern analysis as described by Chen et al. (1993). The gel was equilibrated with blotting buffer [25 mM Tris-HCl (pH 8.3), 190 mM glycine, 20% methanol, 0.01% SDS] for 30 min, and then electroblotted onto a nitrocellulose membrane (BA85; Schleicher & Schuell, Dassel, Germany). The membrane blot was then incubated with renaturation buffer [10 mM Hepes (pH 7.5), 100 mM KCl, 1 mM DTT, 0.1 mM EDTA, 10 mM $MgCl_2$, 0.1 mM $ZnSO_4$ and 5% nonfat milk] at 4 °C for 18 h. After washing with TNE-50 buffer [10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM EDTA, 1 mM DTT], the plot was incubated in TNE-50 buffer containing ^{32}P-labeled oligoprobe D20 (1.5×10^6 cpm · ml^{-1}) and nonspecific competitor poly(dI-dC) · poly(dI-dC) (10 µg · ml^{-1}) for 6 h at 25 °C. The blot was then washed three times with TNE-50 at room temperature for 30 min followed by autoradiography.

Results

Presence of region-D- and region-U-binding proteins in spinach chloroplasts. To investigate the sequence-specific DNA-binding proteins of the *psaA* operon in spinach chloroplasts, the upstream sequence of spinach *psaA* was determined first. Kirsch et al. (1986) have reported the spinach *psaA-psaB* sequences starting from +4 as depicted in Fig. 1. We continued the sequencing work by PCR cloning, and the upstream sequences obtained are shown in Fig. 1. The corresponding rice chloroplast DNA sequences were aligned for comparison. There is overall a 77% homology between spinach and rice sequences in this 245-bp region. The 5′ end of spinach *psaA-psaB-rps14* transcript was finely mapped by primer extension analysis of spinach chloroplast RNA, using a 5′ end-labeled primer corresponding to the sequence complementary to the region from +66 to +90. A major extended product mapped the 5′ end at the C residue lying 181 bp upstream of the *psaA* ATG codon (Fig. 2), which is numbered as +1. This 5′ end coincides well with the 5′ end of the rice *psaA-psaB-rps14* transcript that was determined by Chen et al. (1992).

We have previously reported the existence of chloroplast DNA-binding proteins that interacted specifically with two regions of the rice *psaA* operon (region U and region D shown in Fig. 1) (Chen et al. 1993). Region D,

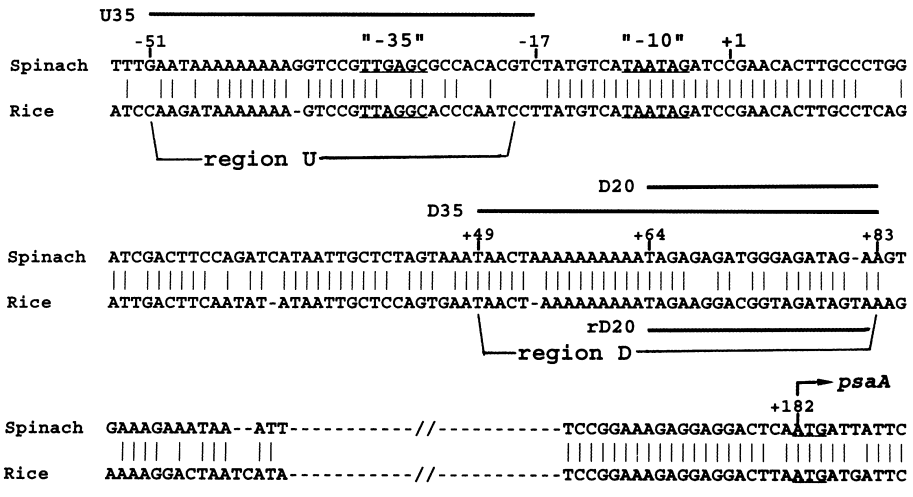

Fig. 1. Nucleotide sequences of the *psaA-psaB-rps14* promoter region of spinach and rice. The transcription start site is numbered as *+1*. The ATG initiation codon for *psaA* is underlined. The prokaryotic-like "−10" and "−35" elements are marked and *underlined*. The sequences in the *brackets* are the two binding regions U and D. The regions overlined or underlined with *thick lines* were used to make oligoprobes D35, D20, rD20, and U35

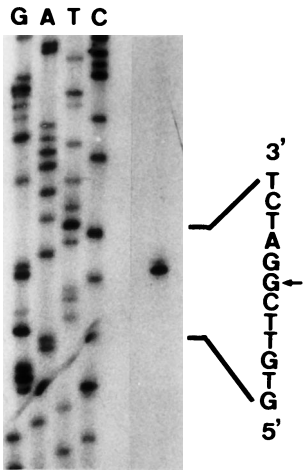

Fig. 2. 5′-End mapping of the transcript for the spinach *psaA-psaB-rps14* operon by primer extension. The primer corresponding to the sequence complementary to the +66 to +90 region depicted in Fig. 1 was employed in the primer extension reaction on spinach chloroplast RNA. The extended products were run in parallel with the sequence ladders of the DNA strand complementary to the primer. The *arrow* indicates the major extension product

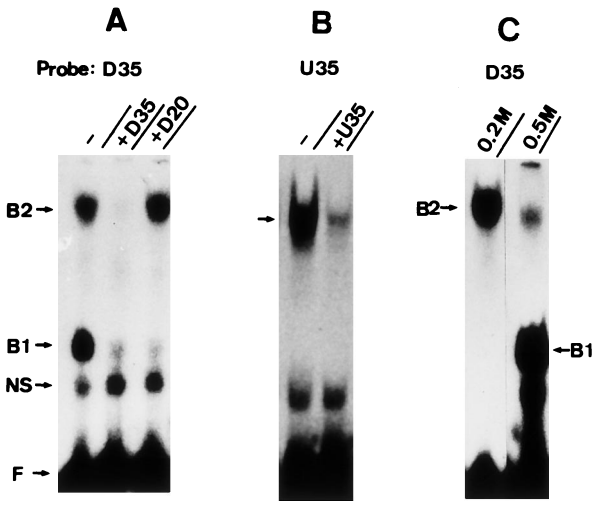

Fig. 3A–C. Identification of region-D- and region-U-binding proteins in spinach chloroplasts by gel retardation assay. **A** The end-labeled oligoprobe D35 (1 ng) was incubated with 30 μg of spinach chloroplast ammonium sulfate fraction (25–65%) in the absence or presence of 400-fold molar excess of unlabeled D35 or D20. The resulting DNA-protein complexes were resolved on a native 5% polyacrylamide gel. The free probe (*F*) and the retarded bands (*NS*, *B1*, *B2*) are indicated at the left margin. **B** The end-labeled oligoprobe U35 (1 ng) was incubated with 30 μg of spinach ammonium sulfate fraction (25–65%) in the absence or presence of 400-fold molar excess of unlabeled U35. **C** The spinach ammonium sulfate fraction (25–65%) was fractionated on a Mono Q-FPLC column with KCl-gradient elution. Each fraction was subjected to the binding assay using oligoprobe D35. The fractions eluted at 0.2 M KCl and 0.5 M KCl formed the B2 complex and B1 complex, respectively

located downstream from the transcription start site, has about 84% sequence homology between rice and spinach. Region U, located upstream from the transcription start site containing the "−35" element, has about 65% homology between these two plants. It is necessary to examine first whether any DNA-binding proteins in spinach chloroplasts interact specifically with the two corresponding regions of the spinach *psaA* operon. For this purpose, two synthetic oligoprobes of 35-bp sequences corresponding to +49 to +83 (oligo D35) and −51 to −17 (oligo U35) of the spinach *psaA* operon were prepared. The spinach chloroplast high-salt extracts were subjected to ammonium sulfate fractionation (0–25%, 25–65%, 65–80%) and tested for oligoprobe-binding activities by gel retardation assays. Both region-D- and region-U-binding activities were found to be present in the 25–65% fraction (Fig. 3). Incubation of the 25–65% fraction with oligoprobe D35 resulted in the formation of several complexes, as revealed on the native gel (retarded bands B1, B2, NS on Fig. 3A). The DNA-protein complexes of B1 and B2 were fully competed by adding 400-fold molar excess of unlabeled oligo D35, while band NS was a nonspecific complex. This competition assay indicated that retarded bands B1 and B2 are due to specific DNA-protein interactions. Incubation of oligo U35 with the 25–65% fraction resulted in the formation of a specific DNA-protein complex of low-mobility, which is fully competed by adding 400-fold molar excess of unlabeled U35 (Fig. 3B). Preliminary analysis of region-U-binding protein by gel filtration chromatography estimated an apparent molecular weight for the native protein of approx. 64 kDa (data not shown). This observation suggests that region-U-binding protein is not related to chloroplast core RNA polymerase or sigma-like factor that binds to DNA only in the presence of RNA polymerase. The identity of the spinach region-U-binding protein remains to be investigated.

To further dissect the sequence specificities of region-D-binding proteins, oligo D20 corresponding to +64 to +83 without the front 15-bp A-rich sequences was used as competitor in the binding reaction. Oligo D20 was found to specifically abolish the formation of DNA-protein complex B1 but showed no effect on the formation of complex B2 (Fig. 3A). It appears that the front A-rich region is necessary for the formation of complex B2. When the spinach chloroplast proteins were fractionated on a Mono Q-FPLC column, the proteins responsible for the formation of complex B1 and complex B2 were well separated by KCl-gradient elution. The protein of complex B2 eluted at about 0.2 M KCl, while the protein of complex B1 eluted at a higher concentration of 0.5 M KCl (Fig. 3C). We then focused on the identity and property of the protein of complex B1 that recognizes sequences +64 to +83 of region D hereafter. Oligo D20 was thus employed as probe in gel retardation assays to monitor purification of the sequence-specific binding protein from spinach chloroplasts.

Purification of region-D-binding protein. The ammonium sulfate fraction (25–65%) was employed as the starting material for purifying region-D-binding protein through a series of column-chromatographic steps. The 25–65% fraction containing approx. 30 mg protein was first applied onto a heparin-agarose column (30 ml) and

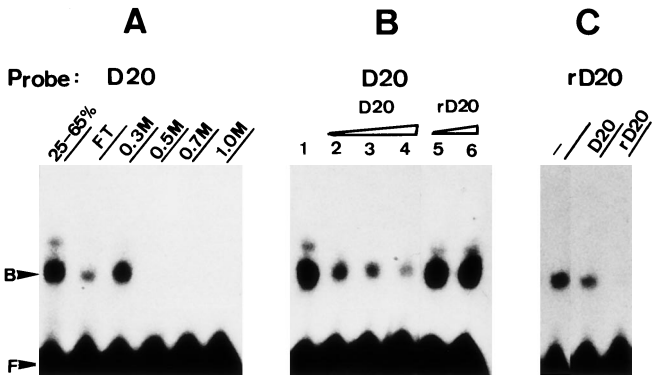

Fig. 4A–C. Heparin-agarose column chromatography of spinach chloroplast protein extracts for region-D-binding protein. **A** The 25–65% ammonium sulfate fraction (30 mg in 1 ml) was fractionated on a heparin-agarose column (30 ml bed volume) as described under *Materials and methods*. Two microliters of each fraction eluted with a KCl step gradient in buffer H containing 0.3 M, 0.5 M, 0.7 M and 1.0 M KCl was used for binding assay with oligoprobe D20. *FT*, flow-through. **B**, **C** Competition assays were performed for sequence-specificity of region-D-binding protein of 0.3 M KCl eluate. Oligoprobe D20 (1 ng, **B**) was incubated with 5 μg of 0.3 M KCl eluate in the absence (*lane 1*) or presence of cold competitor D20 (100-, 200-, 500-fold molar excess, *lanes 2–4*) or competitor rD20 (500-, 1000-fold molar excess, *lanes 5,6*). Oligoprobe rD20 (1 ng, **C**) was incubated with 5 μg of 0.3 M KCl eluate in the absence or presence of 500-fold molar excess of cold competitor D20 or rD20

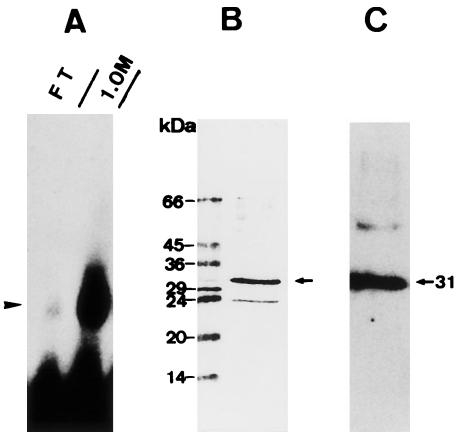

Fig. 5A–C. Purification of spinach chloroplast region-D-binding protein by sequence-specific DNA affinity chromatography. **A** The 0.3 M KCl fraction from the heparin-agarose column (5 mg) was preincubated with nonspecific competitor DNA poly(dI-dC) · poly(dI-dC) and then passed through two cycles of D20-specific affinity resin as described under *Materials and methods*. Both flow-through (*FT*) and the 1.0 M KCl eluate (10 ng in 1.5 μl) were assayed for oligoprobe D20-binding activity. **B** The affinity-purified region-D-binding protein was denatured and electrophoresed on an SDS-polyacrylamide gel (14%) (0.1 μg/lane), followed by silver staining. The protein size markers are indicated at the left margin. **C** A duplicate SDS-polyacrylamide gel (0.3 μg/lane) was subjected to Southwestern blot analysis. After electrophoresis, the protein bands were electroblotted onto a nitrocellulose membrane and renatured as described under *Materials and methods*. The blot was then incubated with oligoprobe D20. The *arrowhead* indicates the 31-kDa protein band interacting with probe D20

eluted with a step gradient composed of 0.3 M, 0.5 M, 0.7 M and 1.0 M KCl in buffer H (Fig. 4A). The 0.3 M KCl eluate possessing most of the oligo-D20-binding activity (retarded band B) was collected and concentrated by precipitation with ammonium sulfate to 80% saturation. The resulting protein fraction (about 6.1 mg protein) was dialyzed overnight against buffer Z and stored at −70 °C.

The highly specific D20-binding property of the 0.3 M KCl eluate was further examined by the competition binding assays. Addition of an approx. 200- to 500-fold molar excess of unlabeled D20 abolished most of the D20-binding activity, while addition of a high concentration of rice homologous sequences rD20 at approx. 500–1000 molar excess did not significantly affect the formation of the D20-protein complex (Fig. 4B). Because rice rD20 sequences shows 75% homology to spinach D20 sequences, it is necessary to investigate whether the spinach 0.3 M KCl fraction has rD20-binding ability. Using rD20 as the probe, Fig. 4C shows that the 0.3 M KCl eluate interacted with rD20 much less strongly than with D20, and the interaction is also specifically inhibited by addition of the rD20 competitor.

The 0.3 M KCl fraction obtained from the heparin-agarose column was further purified by sequence-specific DNA affinity chromatography. The affinity resin was prepared by coupling the catenated D20 to Sepharose 4B with CNBr and then equilibrated in a column (1 ml) with buffer Z containing 0.1 M KCl. The protein fraction from the heparin-agarose column (5–10 mg protein) was adjusted to 0.1 M KCl, preincubated with nonspecific competitor poly(dI-dC) · poly (dI-dC), then passed through the column. The 1.0 M KCl eluate containing DNA-binding activity was desalted and subjected to a second round of DNA affinity chromatography (Fig. 5A). In a typical experiment (Table 1), region-D-binding protein was purified approx. 1900-fold with about 3.2% yield (1.1 μg) from the chloroplast high-salt extracts prepared from 200 g of spinach leaves by two sequential DNA-affinity-chromatographic steps.

A combination of SDS-PAGE and silver staining was employed to check the purity of the DNA-affinity-purified fraction. A major protein band of approx. 31 kDa was observed together with a minor band migrating at the 24-kDa position (Fig. 5B). To correlate the protein band with the DNA-binding activity, Southwestern blot analysis was performed. The separated proteins on the gel were electroblotted onto a nitrocellulose membrane and renatured. The DNA-binding activity of the protein was tested by incubating the membrane filter with oligoprobe D20. Figure 5C was the resulting autoradiogram, and shows the 31-kDa protein band with D20-binding activity. A faint band was also observed at approx. 60-kDa, which is thought to be the dimeric form of region-D-binding protein resulting from incomplete denaturation before loading onto the SDS-polyacrylamide gel. To examine this problem further, the apparent molecular weight of region-D-binding protein was estimated by gel filtration chromatography using a Superose 12 column. The

Table 1. Purification of the *psaA* region-D-binding protein from spinach chloroplasts. The data presented are from a typical purification experiment

Fraction	Total protein (mg)	Specific activity (units · (mg protein)$^{-1}$)	Purification (-fold)	Yield (%)
Chloroplast extracts	65.2	5.1	1	100
Ammonium sulfate (25–65%)	30.7	10.3	2	95
Heparin-agarose	6.1	27.8	5	51
First DNA affinity chromatography	0.0032[a]	6320	1239	6.1
Second DNA affinity chromatography	0.0011[a]	9700	1901	3.2

[a]Estimated values for protein content were based on silver-staining intensity on SDS gels

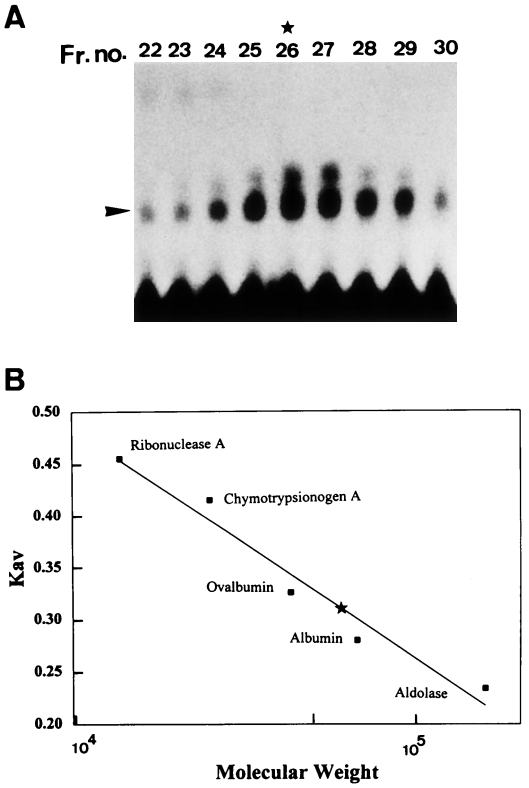

Fig. 6A,B. Molecular-weight estimation of spinach chloroplast region-D-binding protein by gel filtration chromatography. **A** The partially purified region-D-binding protein preparation from the heparin-agarose column (about 500 μg in 200 μl) was fractionated on a Superose-12 column as described under *Materials and methods*. A 2-μl aliquot of each fraction (0.5 ml) was assayed for oligoprobe D20-binding activity. **B** The apparent molecular weight of region-D-binding protein (activity-peak fraction 26) is estimated with the standard curve of K_{av} versus log MW. The position of region-D-binding protein relative to the standard curve is indicated with an *asterisk*

partially purified region-D-binding protein obtained from the heparin-agarose column was utilized for this purpose. The D20-binding activity eluted mostly in fractions 25–28 (Fig. 6A) with maximum activity in fraction 26, corresponding to an apparent molecular weight for the native protein of 59–60 kDa (Fig. 6B). Taken together, the results indicate that spinach chloroplast region-D-binding protein is a dimeric protein composed of two 31-kDa subunits.

Discussion

In the present study, we have employed affinity column chromatography to purify a spinach chloroplast DNA-binding protein which specifically interacts with region D (from position +64 to +83) on the *psaA-psaB-rps14* operon. Biochemical characterization indicates that this sequence-specific DNA-binding protein is a dimeric protein composed of two approx. 31-kDa subunits. Whether it is a homodimer or heterodimer remains to be investigated. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis demonstrated that the highly purified DNA-binding protein preparation contains a small amount of a 24-kDa protein which does not possess a D20-binding property as judged from Southwestern blot analysis (Fig. 5). Copurification of the 24-kDa protein with region-D-binding protein was observed even after two rounds of sequence-specific affinity column chromatography. The protein content was estimated based on silver-staining intensity on SDS gels during the later stages of affinity purification. Generally speaking, our purification procedures resulted in about 1–2 μg of region-D-binding protein starting from 200 g of deveined spinach leaves, and the DNA-binding activity was highly unstable during the later stages of purification (Table 1). Purification on a larger scale is required in the near future to get enough protein for N-terminal microsequencing, antibody induction, and more-detailed functional studies.

The highly sequence-specific interaction between region D and the DNA-binding protein has been demonstrated by competitive gel retardation assays. Spinach D20 sequence binds to spinach region-D-binding protein much more strongly than the rice homologous sequence rD20, which shows 75% homology to D20 (Fig. 4B,C). Purified spinach region-D-binding protein is very likely the "31-kDa polypeptide" identified in our previous study that binds to the rice *psaA* DNA probe extending from −95 to +223 (Chen et al. 1993).

A

```
                +64                  +83
spinach      AAA TAGAGAGATGGGAGATAG-AA GTG
tobacco      AAA TAGATAGATGGGAGATAGAAA AAT
rice         AAA TAGAAGGACGGTAGATAGTAA AGA
maize        AAA TAGAAGGACGGTAGATAGTAA AGA
```

B

```
spinach      AATGAATAAGAGGCTCGTGGGATTGACG
 CDF2
barley         GGAAAGAAGCATAAAGTAAGTAGACC
 AGF
```

Fig. 7. **A** Sequence comparison of region D (+64 to +83) among higher plants. Region D (+64 to +83) of the spinach *psaA-psaB-rps14* operon is aligned with similar sequences located upstream of *psaA* in other plants. **B** Sequences of the spinach CDF2- and barley AGF-binding sites

To our knowledge, spinach region-D-binding protein is a newly identified and highly purified sequence-specific DNA binding protein from higher-plant chloroplasts. This protein is distinct from plastid sigma-like factors that bind to DNA only in the presence of chloroplast RNA polymerase to modulate promoter recognition (Tiller and Link 1993; Troxler et al. 1994). Two types of DNA-binding protein that directly bind to chloroplast DNA in a sequence-specific manner have been documented. The first type is PEND protein (130 kDa), detected in the chloroplast envelope membrane of young pea leaves, which binds to AT-rich sequences within the *psbM*, *rpoC2* and *petA* genes (Sato et al. 1993). This envelope-membrane PEND protein may function to anchor chloroplast DNA during replication or segregation in young plastids. Another type of chloroplast sequence-specific binding protein includes several factors identified both in mono- and dicotyledonous species. The pea 115-kDa protein (denoted as CDF1) is specific for sequences in the intergenic region of *rbcL* and *atpB* (Lam et al. 1988). The maize BF protein is specific for the *psbA* promoter region (Zaitlin et al. 1989). Two spinach proteins of 35 kDa and 33 kDa (CDF2) are specific for a 14-bp sequence located in front of the 16 S rRNA initiation site (Baeza et al. 1991) and seem to act as repressors of rDNA transcription (Iratni et al. 1994). The AGF protein of barley and rice is specific for the *psbD-C* light-responsive promoter (Kim and Mullet 1995; To et al. 1996). Very interestingly, we found that the recognition sites for CDF2, AGF, and region D of *psaA* (+64 to +83) are all AG-rich, although the sequence homology is low (Fig. 7). CDF2, AGF and region-D-binding protein are all composed of polypeptides of similar sizes (31–36 kDa) (Baeza et al. 1991; To et al. 1996). Although the precise biological functions of these factors and region-D-binding protein with respect to the *psaA* operon remain to be established, these protein factors may be members of a larger family of DNA-binding proteins and act as regulators of plastid gene expression in a tissue- or cell-type-specific manner during plant development.

This work was financially supported by Academia Sinica and the National Science Council of the Republic of China.

References

Allison LA, Simon LD, Maliga P (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. EMBO J 15: 2802–2809

Baeza L, Bertrand A, Mache R, Lerbs-Mache S (1991) Characterization of a protein binding sequence in the promoter region of the 16S rRNA gene of the spinach chloroplast genome. Nucleic Acids Res 19: 3577–3581

Bradford MM (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

Chen SCG, Cheng MC, Chen J, Hwang LY (1990) Organization of the rice chloroplast *psaA-psaB-rps14* gene and the presence of sequence heterogeneity in this gene cluster. Plant Sci 68: 213–221

Chen SCG, Cheng MC, Chung KR, Yu NJ, Chen MC (1992) Expression of the rice chloroplast *psaA-psaB-rps14* gene cluster. Plant Sci 81: 93–102

Chen MC, Cheng MC, Chen SCG (1993) Characterization of the promoter of rice plastid *psaA-psaB-rps14* operon and the DNA-specific binding proteins. Plant Cell Physiol 34: 577–584

Fish LE, Kück U, Bogorad L (1985) Two partially homologous adjacent light-inducible maize chloroplast genes encoding polypeptides of the P700 chlorophyll *a*-protein complex of photosystem I. J Biol Chem 260: 1413–1421

Gruissem W, Tonkyn JC (1993) Control mechanisms of plastid gene expression. Crit Rev Plant Sci 12: 19–55

Hess WR, Prombona A, Fieder B, Subramanian AR, Börner T (1993) Chloroplast *rps15* and the *rpoB/C1/C2* gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. EMBO J 12: 563–571

Hu J, Bogorad L (1990) Maize chloroplast RNA polymerase: the 180-, 120-, and 38-kilodalton polypeptides are encoded in chloroplast genes. Proc Natl Acad Sci USA 87: 1531–1535

Hu J, Troxler RF, Bogorad L (1991) Maize chloroplast RNA polymerase: the 78-kilodalton polypeptide is encoded by the plastid *rpoC1* gene. Nucleic Acids Res 19: 3431–3434

Igloi GL, Kössel H (1992) The transcriptional apparatus of chloroplasts. Crit Rev Plant Sci 10: 525–558

Iratni R, Baeza L, Andreeva A, Mache R, Lerbs-Mache S (1994) Regulation of rDNA transcription in chloroplasts: promoter exclusion by constitutive repression. Genes Dev 8: 2928–2938

Kadonaga JT (1991) Purification of sequence-specific binding proteins by DNA affinity chromatography. Methods Enzymol 208: 10–23

Kim M, Mullet JE (1995) Identification of a sequence-specific DNA binding factor required for transcription of the barley chloroplast blue light-responsive *psbD-psbC* promoter. Plant Cell 7: 1445–1457

Kirsch W, Seyer P, Herrmann RG (1986) Nucleotide sequence of the clustered genes for two P700 chlorophyll *a* apoproteins of the photosystem I reaction center and the ribosomal protein S14 of the spinach plastid chromosome. Curr Genet 10: 843–855

Kössel H, Döry I, Igloi G, Maier R (1990) A leucine-zipper motif in photosystem I. Plant Mol Biol 15: 497–499

Lam E, Hanley-Bowdoin L, Chua NH (1988) Characterization of a chloroplast sequence-specific DNA binding factor. J Biol Chem 263: 8288–8293

Lehmbeck J, Rasmussen OF, Bookjans GB, Jepsen BR, Stummann BM, Henningsen KW (1986) Sequence of two genes in pea chloroplast DNA coding for 84 and 82 kD polypeptides of the photosystem I complex. Plant Mol Biol 7: 3–10

Lerbs-Mache S (1993) The 110-kDa polypeptide of spinach plastid DNA-dependent RNA polymerase: single-subunit enzyme or catalytic core of multimeric enzyme complexes? Proc Natl Acad Sci USA 90: 5509–5513

Link G (1994) Plastid differentiation: organelle promoters and transcription factors. In: Nover L (ed) Plant promoters and transcription factors. Springer, Heidelberg, pp 63–83

Little MC, Hallick RB (1988) Chloroplast *rpoA*, *rpoB*, and *rpoC* genes specify at least three components of a chloroplast DNA-dependent RNA polymerase active in tRNA and mRNA transcription. J Biol Chem 263: 14302–14307

Malkin R (1982) Photosystem I. Annu Rev Plant Physiol 33: 455–479

Maniatis T, Fritsch EF, Sambrook J (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

Meng BY, Tanaka M, Wakasugi T, Ohme M, Shinozaki K, Sugiura M (1988) Cotranscription of the genes encoding two P700 chlorophyll *a* apoproteins with the gene for ribosomal protein CS14: determination of the transcriptional initiation site by in vitro capping. Curr Genet 14: 395–400

Orozco EM, Mullet JE, Chua NH (1985) An in vitro system for accurate transcription initiation of chloroplast protein genes. Nucleic Acids Res 13: 1283–1302

Parrett KG, Mehari T, Warren PG, Golbeck JH (1989) Purification and properties of the intact P700 and F_x containing photosystem I core protein. Biochim Biophys Acta 973: 324–332

Pfannschmidt T, Link G (1994) Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (*Sinapis alba* L.) seedlings. Plant Mol Biol 25: 69–81

Sato N, Albrieux C, Joyard J, Douce R, Kuroiwa T (1993) Detection and characterization of plastid envelope DNA-binding protein which may anchor plastid nucleoids. EMBO J 12: 555–561

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5: 2043–2049

Tiller K, Link G (1993) Sigma-like transcription factors from mustard (*Sinapis alba* L.) etioplast are similar in size to, but functionally distinct from, their chloroplast counterparts. Plant Mol Biol 21: 503–513

To KY, Cheng MC, Chen LFO, Chen SCG (1996) Introduction and expression of foreign DNA in isolated spinach chloroplasts by electroporation. Plant J 10: 737–743

To KY, Cheng MC, Suen DF, Mon DP, Chen LFO, Chen SCG (1996) Characterization of the light-responsive promoter of rice chloroplast *psbD-C* operon and the sequence-specific DNA binding factor. Plant Cell Physiol 37: 660–666

Troxler RF, Zhang F, Hu J, Bogorad L (1994) Evidence that σ factors are components of chloroplast RNA polymerase. Plant Physiol 104: 753–759

Webber AN, Malkin R (1990) Photosystem I reaction-center proteins contain leucine zipper motifs. A proposed role in dimer formation. FEBS Lett 264: 1–4

Westhoff P, Alt J, Nelson N, Bottomley W, Bünemann H, Herrmann RG (1983) Genes and transcripts for the P700 chlorophyll *a* apoprotein and subunit 2 of the photosystem I reaction center complex from spinach thylakoid membranes. Plant Mol Biol 2: 95–107

Wu C, Tsai C, Wilson S (1988) In: Setlow JK, Hollaender A (eds) Genetic engineering. Plenum, New York, vol 10, pp 67–74

Zaitlin D, Hu J, Bogorad L (1989) Binding and transcription of relaxed DNA templates by fractions of maize chloroplast extracts. Proc Natl Acad Sci USA 86: 876–880