Identification and Characterization of a Novel RNA Binding Protein That Associates with the 5‘-Untranslated Region of the Chloroplast psbA mRNA †

ARTICLE *in* BIOCHEMISTRY · AUGUST 2004

Impact Factor: 3.02 · DOI: 10.1021/bi035909j · Source: PubMed

7 AUTHORS, INCLUDING:

Ekem Efuet
University of Texas MD Anderson Cancer Center
12 PUBLICATIONS **222** CITATIONS
SEE PROFILE

Stephen Mayfield
University of California, San Diego
128 PUBLICATIONS **8,063** CITATIONS
SEE PROFILE

Identification and Characterization of a Novel RNA Binding Protein That Associates with the 5′-Untranslated Region of the Chloroplast *psbA* mRNA†

Dwight Barnes,[‡,§] Amybeth Cohen,*[,‡,||] Richard K. Bruick,[§,⊥] Katherine Kantardjieff,[#] Sean Fowler,[||] Ekem Efuet,[§,△] and Stephen P. Mayfield[§]

Department of Biological Science, Department of Chemistry and Biochemistry, and W. M. Keck Foundation Center for Molecular Structure, California State University Fullerton, Fullerton, California 92834-6850, and Department of Cell Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received October 27, 2003; Revised Manuscript Received March 17, 2004

ABSTRACT: Binding of proteins to chloroplast-encoded mRNAs has been shown to be an essential part of chloroplast gene expression. Four nuclear-encoded proteins (38, 47, 55, and 60 kDa) have been identified that bind to the 5′-untranslated region of the *Chlamydomonas reinhardtii psbA* mRNA with high affinity and specificity. We have cloned a cDNA that represents the 38 kDa protein (RB38) and show that it encodes a novel RNA binding protein that is primarily localized within the chloroplast stroma. RB38 contains four 70 amino acid repeats with a high percentage of basic amino acids, as well as an amino-terminal extension predicted to act as a chloroplast import sequence. We demonstrate that the 38 kDa precursor protein is imported into isolated chloroplasts and interacts with high specificity to uridine-rich regions within the 5′-untranslated region of the *psbA* mRNA. While database searches have identified hypothetical proteins from several other eukaryotic species with high sequence similarity to the deduced amino acid sequence of RB38, no proteins with homology to RB38 have been biochemically characterized. Bioinformatic analysis of the RB38 sequence, together with structure analysis using circular dichroism and protein modeling, suggests that the 70 amino acid repeats within RB38 are similar in fold to previously identified RNA binding motifs, despite limited sequence homology.

Regulation of eukaryotic gene expression involves a number of posttranscriptional processes, including those affecting mRNA stability and translation. The interaction between RNA binding proteins and the corresponding RNA elements found within particular mRNAs has been shown to be an essential part of translational regulation. Conserved themes emerge among a variety of organisms when evaluating mRNA binding proteins and the RNA elements they recognize. The expanding database of RNA structures and protein complexes suggests that unique protein−RNA interactions are based on a relatively small set of functional groups (*1*), and these interactions can have profound effects on the expression of specific genes.

The 5′- and 3′-untranslated regions (UTRs)[1] of both prokaryotic and eukaryotic mRNAs play important roles in translation. Often, these RNA elements are bound by specific proteins as a means to regulate the translation of the adjacent coding region. For example, poly(A) binding protein (PABP) has been shown to play an important role in the initiation of eukaryotic protein synthesis through an interaction with eIF4G (yeast and plants) and eIF4B (plants) at the 5′ end of the mRNA and direct interaction with the poly(A) tail on the 3′ end of the mRNA (*2*−*4*). Prokaryotic systems also utilize mRNA binding proteins for translational regulation. However, these proteins typically inhibit translation after binding to the 5′-UTR (*5*). Thus, RNA binding protein interaction with specific RNA elements within UTRs is a universal mechanism to enhance or inhibit translation.

The majority of mRNA binding proteins described to date contain β sheets that act to stabilize the interaction of charged and aromatic side chains of the proteins with the RNA (*1*). Most of these proteins contain a combination of β sheets and α helices that form the well-characterized RNA recognition motif (RRM) that binds to target RNA. PABP contains a total of four RRMs that are positioned within conserved amino acid sequences. Deo et al. (*6*) have shown by X-ray

† This study was supported by funds from the Department of Energy (ER15313) and the National Institutes of Health (GM54659) to S.P.M. and by NIH Grant GM58808-01 to A.C. E.E. was supported by a Skaggs postdoctoral fellowship.

* To whom correspondence should be addressed. E-mail: acohen@fullerton.edu. Fax: (714) 278-3426. Phone: (714) 278-2178.

‡ D.B. and A.C. contributed equally to this work.

§ The Scripps Research Institute.

|| Department of Biological Science, California State University Fullerton.

⊥ Present address: Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390-9152.

Department of Chemistry and Biochemistry and W. M. Keck Foundation Center for Molecular Structure, California State University Fullerton.

△ Present address: The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030.

[1] Abbreviations: UTR, untranslated region; PABP, poly(A) binding protein; RRM, RNA recognition motif; KH, K homology; RBD, RNA binding domain; RBS, ribosome binding site; PDI, protein disulfide isomerase; wt, wild type; IVT, in vitro translation; PAGE, polyacrylamide gel electrophoresis; PDB, Protein Data Bank.

10.1021/bi035909j CCC: $27.50
Published on Web 06/12/2004

crystallography that the β sheets within the two N-terminal RRMs of human PABP form a long trough in which poly(A) RNA binds. Other well-characterized classes of RNA binding proteins include the K homology (KH) and double-stranded RNA binding domain (dsRBD) superfamilies. Although these three major classes of RNA binding proteins share a common $\alpha\beta$ structural theme, they use a number of different secondary structures in the recognition of RNAs and in the control of posttranscriptional regulation of gene expression.

Molecular genetic studies have shown that translational regulation of chloroplast-encoded proteins is dictated by the interaction of nuclear-encoded proteins with RNA elements located in the 5′- and 3′-UTRs of chloroplast mRNAs (*7*, *8*). Several nuclear-encoded factors have been shown to bind to A−U-rich regions in the 5′-UTRs of chloroplast-encoded mRNAs (*9*). The 5′-UTR in the chloroplast-encoded *psbA* mRNA (encodes the D1 protein of photosystem II) in *Chlamydomonas reinhardtii* is recognized by a set of nuclear-encoded RNA binding proteins (*10*−*12*). These proteins (38, 47, 55, and 60 kDa), which are termed RB for *R*NA *b*inding, demonstrate light-regulated binding to the 5′-UTR of the mRNA (*10*). This RNA binding activity can be modulated in vitro by redox potential and ADP-dependent protein phosphorylation (*11*, *13*). Site-directed mutagenesis of the 5′-UTR of the chloroplast-encoded *psbA* mRNA indicates that specific RNA sequences, including a stem−loop region and an adjacent Shine−Dalgarno ribosome binding site (RBS), are required for translation of the *psbA* mRNA in vivo (*14*, *15*). Examination of *C. reinhardtii* nuclear mutants deficient in *psbA* translation has shown that the RB47 protein, which binds an A-rich region in the 5′-UTR of the *psbA* mRNA (*16*), is required for ribosome association and subsequent translation of this mRNA (*12*, *17*).

Two of the *psbA* mRNA binding proteins, RB47 and RB60, have been well characterized. The RB47 protein shows high homology to PABP (*16*), and the RB60 protein is homologous to protein disulfide isomerases (PDI; *18*, *19*). Each of these proteins contains a chloroplast import signal and appears to play an important role in *psbA* mRNA translation. RB60 is capable of regulating the binding activity of RB47 by changing the oxidation−reduction state of RB47 via breakage or formation of disulfide bonds (*19*, *20*). The reduction of RB47 in the light results in an increased affinity of the protein for the *psbA* 5′-UTR. It has been hypothesized that the reducing potential utilized by RB60 to regulate RB47 binding is generated from the light reactions of photosynthesis.

RB38 is a novel RNA binding protein that binds directly to the 5′-UTR of the *psbA* mRNA and interacts with U-rich sequences, as opposed to the A-rich sequences previously shown to be recognized by RB47 (*16*). RB38 is imported into chloroplasts and is primarily located within the stroma. The protein contains four 70 amino acid repeats, which contain a number of basic and aromatic residues that may play a role in RNA binding. Several hypothetical proteins have been identified from *Arabidopsis thaliana* and other species that contain repeats and domain structures similar to those identified within RB38. Structural modeling of RB38 suggests that the protein may be a divergent member of the RBD or KH superfamily.

MATERIALS AND METHODS

Cell Growth Conditions. Wild-type (wt) *C. reinhardtii* (137c), the cell wall-deficient strain (cw15), and a *psbA*-deficient strain (cc744) were grown in Tris−acetate−phosphate medium (TAP; *21*) to a density of approximately 5×10^6 cells/mL under constant light at 25 °C. Cells were harvested by centrifugation at 4 °C for 5 min at 4000*g*. Cells were either used immediately or frozen in liquid N_2 for storage at −70 °C.

Cloning of the RB38 cDNA. The RB38 protein was purified according to Danon and Mayfield (*10*). The protein was digested with trypsin or proteinase Lys-C, and the peptide fragments were separated on a HPLC and microsequenced (John Lysek, Worcester Foundation for Experimental Biology, Worcester, MA). Three peptide sequences were obtained (RDETLANWEQLR, SNPDEWYDNR, QAAEAANWEALR). Degenerate oligos were designed against two of the peptides, end-labeled with [γ-^{32}P]ATP, and used to screen a *C. reinhardtii* phage library according to Sambrook et al. (*22*). Plaques that hybridized with each oligo were selected, and two of the several 2.0 kb cDNAs obtained were subcloned for sequence analysis.

Southern Analysis. Twenty micrograms of genomic DNA from wt cells was digested with restriction endonucleases and separated by size on a 0.8% agarose and 1× TBE gel. The separated DNA was blotted to a nylon membrane according to Sambrook et al. (*22*). The membrane was hybridized with a random-primed, ^{32}P-labeled RB38 cDNA, and autoradiography was performed according to the procedures described by Mayfield et al. (*15*, *23*).

Total RNA Isolation and Analysis. Total RNA was isolated from wt cells according to the methods of Cohen et al. (*24*). Electrophoresis of RNA, blotting, hybridizations with a random-primed, ^{32}P-labeled RB38 cDNA, and autoradiography were performed according to the procedures described by Mayfield et al. (*15*, *23*).

Expression and Purification of RB38 Domains 1 and 2, RB38, and RB47. A DNA fragment containing the RB38 RNA binding domains 1 and 2 (D1/D2) and the full-length RB38 and RB47 cDNAs, all without the import sequences, were amplified by PCR and subcloned into pCR2.1-TOPO (Invitrogen, Carlsbad, CA). The clones were sequenced to ensure that no errors were introduced, subcloned into a pET23a expression vector (Invitrogen), and subsequently transformed into *Escherichia coli* BL21 LysC cells. The cells were allowed to grow to a density of 0.6 (OD_{600}), prior to the addition of isopropyl β-D-thiogalactosidase (1 mM final concentration). Cells were grown for an additional 4 h and pelleted by centrifugation. The expressed proteins were isolated from the *E. coli* pellet by Ni−agarose chromatography as described (*20*). The proteins were dialyzed overnight into 100 mM NaCl, 20 mM Tris (pH 8.0), and 2 mM DTT and further purified using MonoQ and Superdex 200 columns (Amersham Biosciences, Piscataway, NJ).

RNA Gel Shift Assays. RNA binding analysis was carried out on 10−1000 nM purified recombinant protein in buffer containing 25 mM Tris (pH 7.5), 100 mM KCl, 0.2 mM EDTA, 4 mM $MgCl_2$, and 3% glycerol. The recombinant RB38 or RB47 (rRB38, rRB47) proteins were incubated with 100 pM ^{32}P-labeled *psbA* 5′-UTR and 1 μg of wheat germ tRNA (Sigma-Aldrich, St. Louis, MO) at 4 °C for 20 min.

For the competition assays, 500 ng of unlabeled total RNA from cc744 or wt cells, or 100 ng of unlabeled poly(U) and poly(A) RNA (Sigma-Aldrich), was added to the reaction mix containing 100 nM rRB38 or 150 nM rRB47. Unlabeled competitor RNAs (10× and 100×) were used in reactions containing 150 nM rRB38. Following incubation, the RNA−protein complexes were separated from free RNA by electrophoresis on 7% native polyacrylamide gels containing 1× TBE. The RNA−protein complexes were visualized on a Cyclone phosphorimager (Packard Bioscience, Meriden, CT).

Sequence Analysis. The sequence of RB38 was analyzed for similarity to other genomic sequences using a number of standard similarity search algorithms (*25*, *26*). Putative domains were identified using Domain Fishing (*27*) and Predict Protein (*28*). Several widely used algorithms were used to predict secondary structure profiles (*27*, *29*−*32*) and to identify fold similarities (*33*−*38*). Putative domains were also subjected to the same sequence and fold similarity searches in an attempt to improve upon the sequence and structural homology predictions.

Circular Dichroism. rRB38 D1/D2 and rRB38 (both minus the transit sequence) were expressed and purified as previously described. The proteins were dialyzed into 10 mM Tris-HCl (pH 7.5) in microdialysis cells, and protein concentrations for molar ellipticity calculations were estimated using absorbance at 280 nm and amino acid composition. CD spectra (190−240 nm) were acquired at room temperature with a Jasco J-720 spectropolarimeter interfaced with Jasco data acquisition and standard analysis software. Protein solutions of ~0.5 mg/mL were placed in 0.1 cm cylindrical quartz cells. Data were converted to molar ellipticity, and secondary structure composition was calculated using JFIT (B. Rupp; http:/www-structure.llnl.gov/cd/cdtutorial.htm).

Chloroplast and Protein Isolation and Immunoblotting. Crude protein isolation, electrophoresis, and immunoblotting were performed as described by Mayfield et al. (*15*). Briefly, wt cells were resuspended in extraction buffer [750 mM Tris (pH 8.0), 15% sucrose, and 100 mM β-mercaptoethanol] and disrupted by sonication with a microtip probe. Membrane proteins were pelleted by centrifugation and resuspended in a small volume of extraction buffer. Soluble proteins were retained in the supernatant.

Chloroplasts fractions were isolated from cw15 cells. Chloroplasts were isolated according to Harris (*21*), Nickelsen et al. (*39*), and Lawrence and Kindle (*40*), with slight modifications. Cells were harvested by slow-speed centrifugation, washed in 20 mM Hepes−NaOH (pH 7.5), and resuspended in breaking buffer [0.33 M sorbitol, 50 mM Hepes−NaOH (pH 7.2), 2 mM EDTA, 1 mM $MgCl_2$, and 1% BSA]. The cells were incubated on ice for 10 min and then broken in a cell disruption bomb (Parr Instrument Co.) using 50 psi for 3 min. Intact chloroplasts were obtained by layering the broken cells over a 40%:70% Percoll gradient in breaking buffer (1:1 v/v) and centrifuging at 2500*g* for 7 min at 4 °C. The chloroplast pellets were recovered from the 40:70 interface and washed with breaking buffer. The intact chloroplasts were lysed by sonication, and the membrane fraction was separated from the soluble phase by centrifugation at 10000*g* for 15 min at 4 °C.

Protein separation and Western blot analysis were carried out as described by Cohen et al. (*24*).

In Vitro RNA Transcription Reactions. High specific activity RNA probes were synthesized using T7 RNA polymerase (New England Biolabs, Beverly, MA) according to the procedures of Melton et al. (*41*). Reactions were carried out using the *psbA Rsa*I−*Nde*I DNA template containing the full-length 90 nucleotide 5′-UTR (*10*), *psbA* short 5′-UTR (30 nucleotides from the RBS to the initiation codon), *psbA* 3′-UTR, and full-length *atpA*, *rbcL*, *psbD*, and *petD* 5′-UTRs cloned downstream of the T7 promoter into pUC18. Unlabeled transcription reactions were carried out according to the New England Biolabs protocol.

In Vitro Translation and Chloroplast Import. Chloroplasts were isolated from cw15 cells as described previously and resuspended at a concentration of 1 mg of chlorophyll/mL in 1× import buffer [0.33 M sorbitol and 50 mM Hepes−KOH (pH 7.5)].

RNA was transcribed from 0.5 μg of an *E. coli* plasmid (pBS-SK+) containing the coding region of the RB38 cDNA or the OEE1 cDNA (*42*) cloned downstream of the T7 promoter. These RNAs and luciferase RNA (Promega, Madison, WI) were incubated in a rabbit reticulocyte in vitro translation (IVT) system (Promega) at 37 °C for 50 min according to the manufacturer's instructions.

Import assays were performed by incubating 10 μL of the RB38, OEE1, or luciferase IVT mix in a reaction containing chloroplasts at 1 mg of chlorophyll/mL, 100 mM MgATP, and 25 mM methionine in 1× import buffer for 10 min at room temperature in the presence of light (*43*). After import, chloroplasts were isolated over 40% Percoll cushions in 1× import buffer or treated with 2 mg/mL thermolysin for 15 min on ice, followed by the addition of EDTA to a final concentration of 50 mM. Thermolysin-treated chloroplasts were isolated over 40% Percoll containing 5 mM EDTA by washing in 1× import buffer, sedimenting at 1500*g* for 5 min, and resuspending in 1× import buffer. These chloroplasts were lysed by sonication with two 15 s pulses from a microprobe at 35% strength. Sonicated proteins were placed on ice for 30 min in the presence or absence of 2 mg/mL thermolysin. Chloroplast proteins were separated by SDS−polyacrylamide gel electrophoresis (PAGE), and imported proteins were detected by autoradiography after treatment of the gel with Enhance (Perkin-Elmer NEN, Boston, MA) according to the manufacturer's instructions.

RESULTS

Isolation and Microsequence Analysis of the RB38 Protein. The RB38 protein was isolated from *C. reinhardtii* cells using *psbA* RNA affinity chromatography as described in Danon and Mayfield (*10*). Isolated RB38 was digested with proteases, and individual peptide fragments were sequenced by Edmond degradation. The amino acid sequence from two of three sequenced peptides was utilized to design oligonucleotides for use in cloning of the RB38 cDNA. Screening of a cDNA library with the oligos detected several putative RB38 clones that hybridized to both oligonucleotides. Many of the clones contained inserts of 2 kb in length that were determined to be identical by restriction analysis. Using one of these cDNAs as a probe, a single band was observed in Southern analysis using high- and low-stringency washes (Figure 1A), indicating that RB38 is encoded by a single

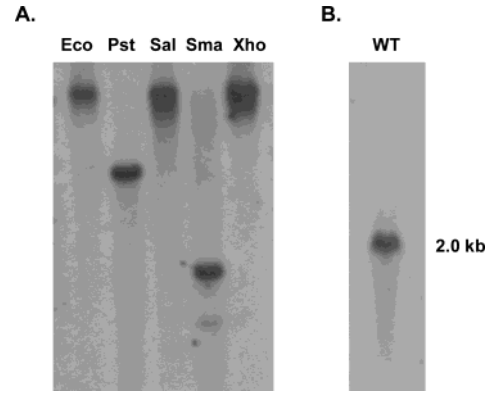

FIGURE 1: (A) Southern blot analysis of *C. reinhardtii* genomic DNA. Twenty micrograms of wt genomic DNA was digested with *Eco*RI, *Pst*I, *Sal*I, *Sma*I, and *Xho*I, separated on an agarose gel, blotted to nylon membrane, and hybridized with a RB38 cDNA probe. (B) Northern blot analysis of wt total RNA. Twenty micrograms of RNA was separated on a denaturing agarose gel, blotted to nylon membrane, and hybridized with a RB38 cDNA probe.

gene and is not a member of a large gene family. Northern analysis with total RNA from wt cells identified a single 2.0 kb mRNA (Figure 1B).

The RB38 Deduced Amino Acid Sequence Reveals a Novel Protein. Analysis of the deduced amino acid sequence of the RB38 cDNA (Figure 2A) showed that each of the three sequenced peptides obtained from the isolated RB38 protein was present in the predicted protein. The first 18 amino acids in RB38 were identified as a chloroplast import sequence by the ChloroP 1.1 Prediction Server (http://www.cbs.dtu.dk/services/ChloroP/). The deduced amino acid sequence did not show a high degree of similarity with other previously identified proteins. Homologues have been identified from *A. thaliana* EST libraries, as shown in Figure 2B, but these theoretical proteins have no known biochemical function (*25*). RB38 also exhibits weak (<30%) sequence homology to segments of several eukaryotic initiation factor proteins, as well as viral pol polyproteins (data not shown), consistent with the hypothesis that RB38 contains regions involved in nucleic acid binding.

RB38 has several characteristics reminiscent of previously described RNA binding proteins. The protein has a large number of basic amino acids (18.2%), with arginine constituting 15.4% of the total protein, almost twice any other residue type, aside from alanine (12.4%). RB38 also contains 7% aromatic amino acids, with tryptophan occurring most frequently within the protein. There are four conserved domains of approximately 70 amino acids within RB38 (Figure 2A), separated by linking sequences of approximately 12−16 residues. Many RNA binding proteins contain multiple domains showing sequence similarities, with RNA binding activity often determined cooperatively by two, or occasionally more, domains (*44*).

When standard sequence search methods (*25*, *26*) failed to detect a significant relationship between RB38 and a protein of known function, we analyzed RB38 using hybrid fold recognition methods, which combine sequence-derived properties with evolutionary information (*45*). From this analysis, each of the repeat domains in RB38 was predicted to belong to the mixed α and β class of proteins, a common structural theme of RNA binding protein domains (*29*, *32*, *46*).

To improve upon the sequence and structural homology predictions of RB38, methods based upon profile neural networks and hybrid fold recognition motifs were used (*28*). Weak resemblance to a structural variant of the dsRBD binding motif, which is a member of the KH superfamily (*47*), was identified along with potential (∼15%) sequence homology between domain 1 and the poly(A) binding protein (PDB 1CVJ_E; $E = 0.5$; *30*).

To assign plausible structures to the RB38 sequence, we used LOOPP, which performs sequence to structure threading as well as structure to structure alignments (*37*). LOOPP identified a high-scoring endotoxin (1CBY) template with "good" structural alignment. After further refinement of side chain positions with SCWRL3 (*48*), the resulting model for RB38 domain 1 had quite good structure and fold quality scores for a low-homology model, as evaluated by PROCHECK (*49*) and PROSAII (*50*). A comparison of a classic RRM (PDB 1QM9), PABP (PDB 1CVJ), the Nova-2 Kh3 RBD (PDB 1DTJ), and RB38 domain 1 is shown in Figure 3. The model of RB38 domain 1 (Figure 3D) appears to be similar to known RNA binding motifs, with a core antiparallel β sheet flanked by α helices. The charge surface of the RB38 domain 1 model (data not shown) is predominantly positive in specific regions that are known to make molecular contacts with RNA, as is seen in PABP and the KH domains (Figure 3B,C). The location of this region in the RB38 domain 1 model, which is rich in arginines, as well as the overall α and β architecture of the fold, suggests that the mode of RNA binding for RB38 domain 1 may be similar to that found in the KH domains. Although RB38 domain 1 and other RNA binding proteins show low (∼13%) sequence identity, they share a structural homology consistent with the predicted secondary structure and CD spectrum of domain 1 (Figure 4). Domains 2, 3, and 4 could also be modeled (threaded) with good reliability using LOOPP (data not shown). Each domain contained an α and β architecture similar to domain 1 with backbone folds reminiscent of KH domains.

Circular Dichroism and Calculated Secondary Structure Content. The secondary structure composition of RB38 domain 1 as calculated from the CD spectrum (Figure 4A) is 38% helix, 25.4% sheet, and 36.6% coil, with a residual of 6.9%. The composition of the four binding domains of RB38 (Figure 4B) is 47.8% helix, 8.5% sheet, and 43.7% coil, with a residual of 14.9%. Using the PROF algorithm (*28*), the secondary structure content of RB38 domain 1 is calculated to be 27.2% helix, 17.4% sheet, and 55.4% coil, and the four domains are 38.7% helix, 8.9% sheet, and 52.4% coil. Each of these calculated structural characteristics is clearly within error of the content predicted from the CD spectrum. For comparison, the secondary structure content of the KH domain in Figure 3C using the DSSP algorithm (*51*) is 38.1% helix, 28.9% sheet, and 33% coil, which is in close agreement with the predicted secondary structure generated from the modeled RB38 protein domain 1 (Figure 3D).

The Endogenous RB38 Protein Is Localized in the Chloroplast Stroma. Immunoblot analysis was used to further confirm that the RB38 cDNA represented the endogenous protein. As shown in Figure 5A, a 38 kDa protein is

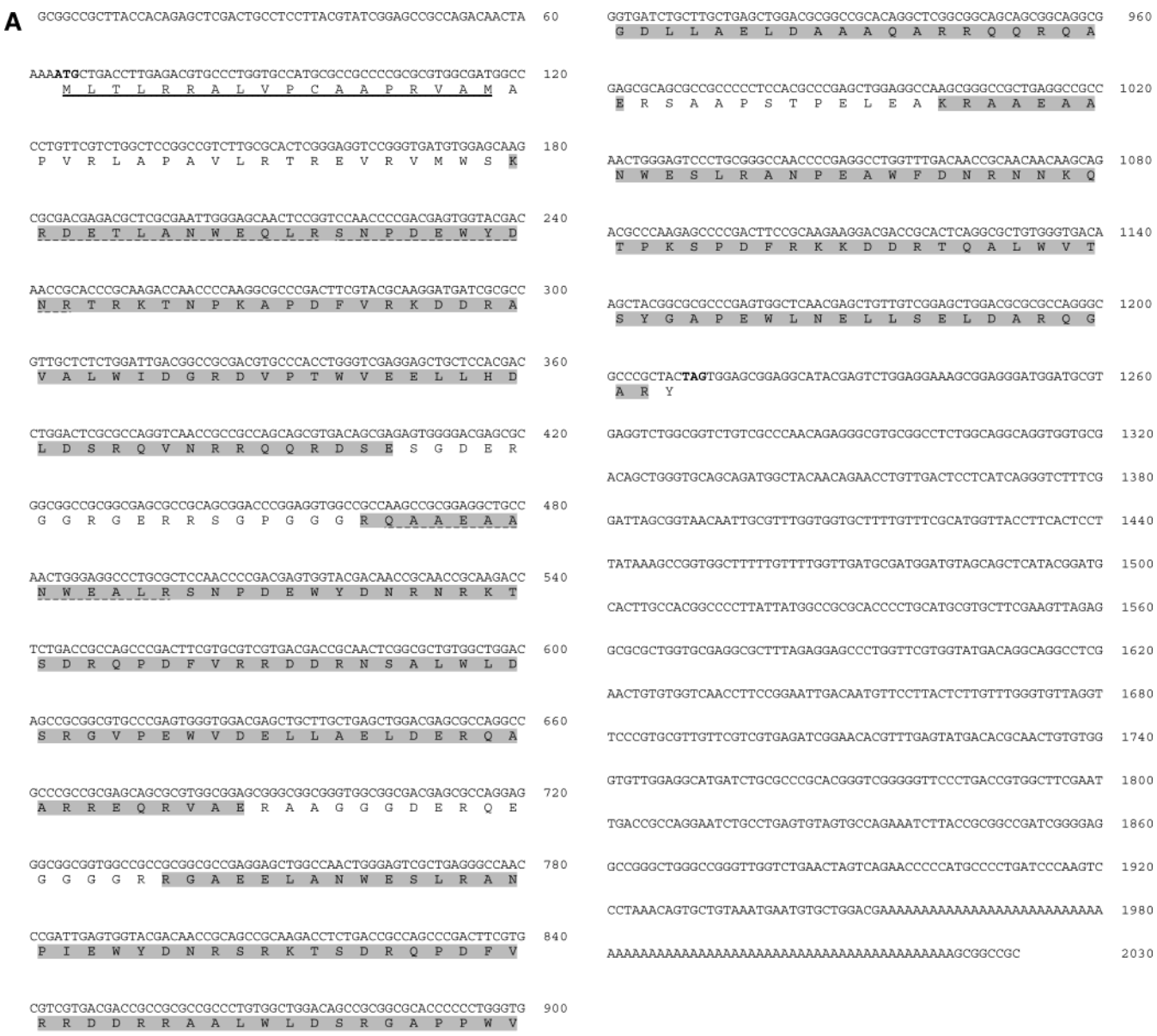

FIGURE 2: (A) The cDNA and deduced amino acid sequence of RB38. The start codon (ATG) and stop codon (TAG) of the deduced amino acid sequence are indicated in bold. The four 70 amino acid repeats with high sequence similarity are shaded. The predicted chloroplast import sequence is underlined, while peptides obtained from microsequence of the endogenous RB38 from chloroplasts are indicated by a dashed underline. The accession number of RB38 is AY124882. (B) Alignment of the four 70 amino acid repeats in RB38 (1−4) with each other and five hypothetical proteins from *A. thaliana* (identified by their NCBI accession numbers). Identical amino acids are indicated in black and marked with an asterisk. Amino acid similarity is indicated in gray. Both AC002342 and AL021637 show two repeats similar to the four RB38 repeats, while AC000107, AC007519, and AC012463 contain a single repeat.

recognized in isolated chloroplasts from *C. reinhardtii* by antisera raised against rRB38 protein. Expression of the RB38 cDNA resulted in the production of a recombinant protein with a molecular mass of 45 kDa, resulting from the additional amino acids of the transit peptide (4 kDa) and the His tag and linker (3 kDa).

Separation of chloroplasts into soluble and membrane fractions revealed that RB38 is found primarily in the soluble stroma and is not associated with the thylakoid membranes of the chloroplast (Figure 5B). The large subunit of Rubisco also accumulated in the soluble fractions of whole cells and chloroplasts, while the light-harvesting complex protein

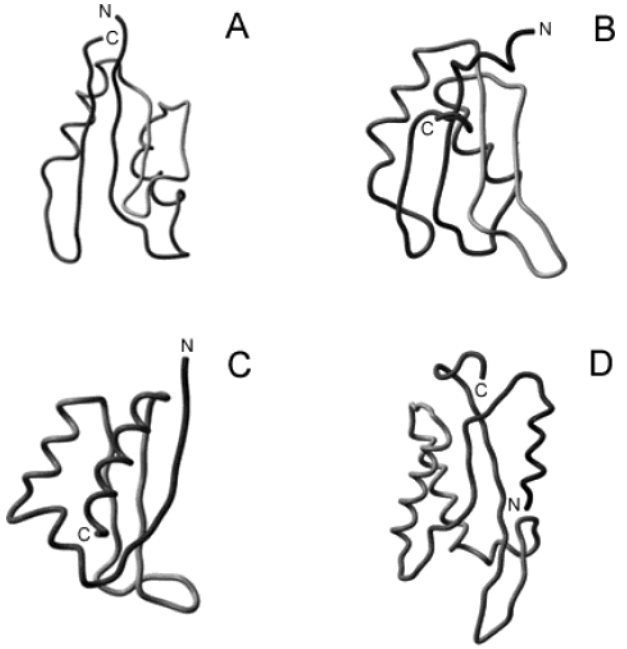

FIGURE 3: Three-dimensional structures of RNA recognition motifs. The direction of the folds is indicated (N to C), with the N-terminus colored in darkest gray. (A) A classic RRM, the polypyrimidine tract binding protein (PDB 1QM9). (B) The canonical RBD, PABP (PDB 1CVJ). (C) the Nova-2 Kh3 domain (PDB 1DTJ). (D) RB38 domain 1 model. All folds generally have a mixed α and β topology, with a core β sheet flanked by α helices. Domain 1 includes the N-terminal transit peptide, which is predicted to be and models as helical. Graphics were rendered with ICM-Browser-Pro, Molsoft L.L.C. (*61*).

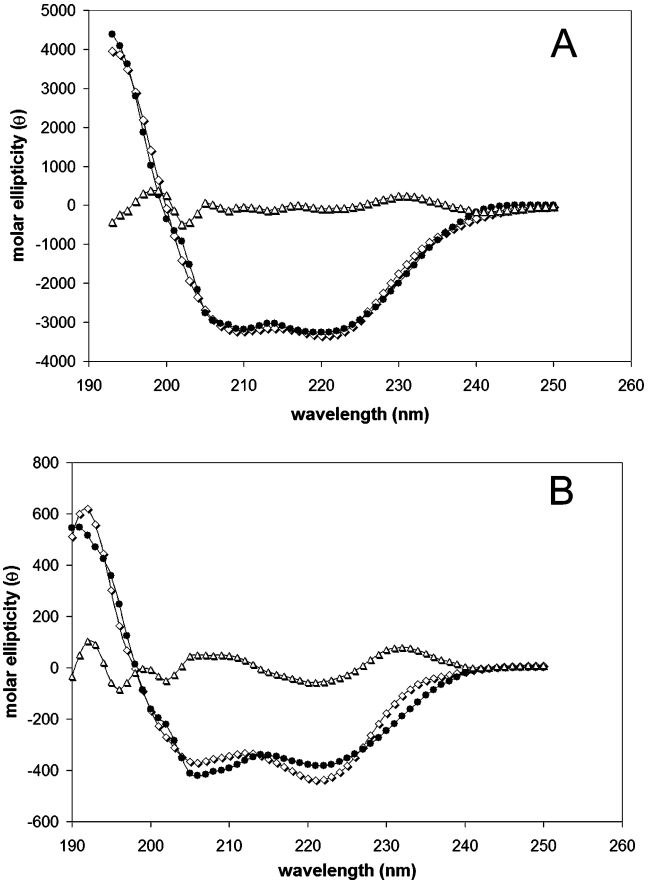

FIGURE 4: Circular dichroism spectroscopy. (A) RB38 domain 1 minus the transit peptide. The secondary structure composition of RB38 domain 1 as calculated from the CD spectrum with JFIT (B. Rupp, http://www-structure.llnl.gov/cd/cdtutorial.htm) is 38% helix, 25.4% sheet, and 36.6% coil, with a residual of 6.9%. (B) RB38 full sequence minus the transit peptide. The composition of the four binding domains of RB38 is calculated to be 47.8% helix, 8.5% sheet, and 43.7% coil, with a residual of 14.9%.

accumulated only in the membrane fractions of whole cells and isolated chloroplasts, both as expected. The cytoplasmic translation initiation factor eIF4A was only present in the soluble fraction of whole cells and not in isolated chloroplasts, demonstrating that the chloroplasts were free of cytoplasmic contamination.

RB38 Binds to the psbA 5′-UTR. To directly assay the ability of RB38 to bind to the 5′-UTR of the *psbA* mRNA, in vitro gel mobility shift assays (*10*) were performed with purified rRB38. A Coomassie-stained SDS−PAGE gel (Figure 6A) of the purified RB38 domains 1 and 2 (rRB38 D1/D2), all four domains (rRB38), and full-length rRB47 (including the RNA binding domains and the carboxy unique domain; 70 kDa total) indicated that all the proteins are free of bacterial contaminants. As seen in Figure 6B, the full-length rRB38 and rRB38 D1/D2 proteins bind radiolabeled *psbA* 5′-UTR, forming RNA−protein complexes that migrate slower than free RNA during electrophoresis in nondenaturing gels. A single RNA binding domain alone does not bind radiolabeled *psbA* 5′-UTR (data not shown). The rRB38 D1/D2 protein initially binds the RNA at 100 nM, presumably as a dimer, and forms a faster migrating RNA−protein complex at a protein concentration of 300 nM (Figure 6B). This faster migrating complex may represent a D1/D2 monomer−RNA complex that forms at greater protein concentrations. The initial shift of rRB38 D1/D2 appears to migrate similarly to the full-length rRB38, implying that a protein consisting of four domains may be responsible for this higher affinity binding. The full-length rRB38 protein also initially binds RNA at low protein concentration. Increasing the rRB38 protein concentration reveals two RNA−protein complexes, suggesting a high-affinity binding site at 100 nM and a second, low-affinity binding site at 300 nM (Figure 6B). The binding constants for these two sites are calculated to be 63 and 488 nM for the high-affinity and low-affinity sites, respectively.

To investigate the specificity of the RB38 protein for the *psbA* 5′-UTR, competition gel shift assays were performed. As shown in Figure 6C, the addition of 500 ng of total RNA from a wt strain of *C. reinhardtii* or from a *psbA*-deficient strain (cc744) reduced the binding of rRB38 to the radiolabeled *psbA* 5′-UTR. The addition of 100 ng of poly(U) RNA nearly abolished RB38 binding, while the addition of 100 ng of poly(A) RNA had little effect (Figure 6C). A comparison between binding of rRB38 and rRB47 shows reduced binding of each protein in the presence of 500 ng of unlabeled wt total RNA (Figure 6C). In contrast to rRB38 binding, rRB47 binding is competed by the addition of poly(A) RNA (*16*) but not by the addition of poly(U) RNA. rRB47 binding is also poorly competed by the addition of total RNA from a *psbA*-deficient strain (cc744). These RNA binding assays clearly show that rRB38 binding is impacted by a U-rich sequence, which is different from the previously reported preference of rRB47 for A-rich sequences (*16*). The decreased binding of rRB38 to the *psbA* 5′-UTR in the

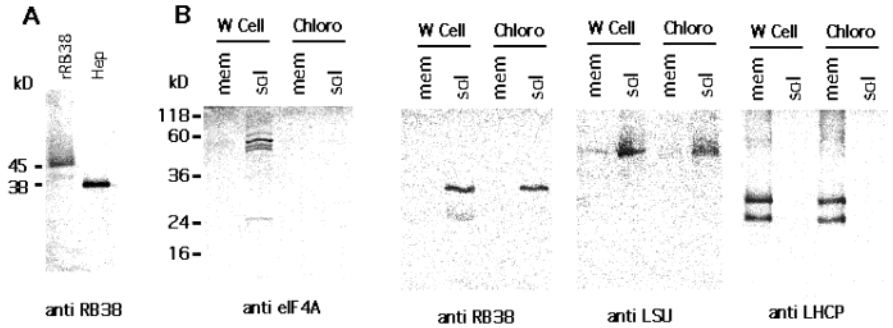

FIGURE 5: Protein accumulation of RB38. (A) Immunoblot analysis of rRB38 and endogenous RB38 isolated from *C. reinhardtii* chloroplasts and purified over heparin−agarose columns (Hep). rRB38 migrates as a 45 kDa protein, while the endogenous RB38 migrates as a 38 kDa protein. (B) Immunoblot analysis of duplicate gel membrane (mem) and soluble (sol) proteins isolated from whole cells (W cell) or chloroplasts (chloro) and decorated with antisera raised against eIF4A (eIF4), RB38, large subunit of Rubisco (LSU), or light-harvesting complex (LHCP).

presence of cc744 RNA suggests that RB38 may be a more general RNA binding protein than RB47.

To further examine the specificity of RB38 for the *psbA* 5′-UTR, different regions of the *psbA* RNA were used as competitors in binding reactions (Figure 6D). Complex formation was greatly reduced by the addition of a 10-fold excess of the unlabeled, full-length *psbA* 5′-UTR and completely eliminated by a 100-fold excess. Addition of a 100-fold excess of the short *psbA* 5′-UTR, containing the 30 nucleotides from the RBS to the initiation codon, reduced binding only slightly, as did competition with the *psbA* 3′-UTR. These data suggest that RB38 binds a region of the *psbA* 5′-UTR located 5′ of the RBS, which contains two stretches of U-rich sequence.

The decreased binding of rRB38 to the *psbA* 5′-UTR in the presence of total RNA suggested that RB38 may bind other chloroplast mRNAs in addition to *psbA*. To examine the specificity of RB38 binding, unlabeled RNA from four other chloroplast mRNAs was used in competition assays. As shown in Figure 6D, the *rbcL* 5′-UTR had no effect on binding while the addition of the *atpA* and *psbD* 5′-UTRs each resulted in a slight decrease in *psbA* 5′-UTR−protein complex formation. The addition of a 100-fold excess of unlabeled *petD* 5′-UTR competed with binding to approximately the same level as a 10-fold excess of the *psbA* 5′-UTR (Figure 6D) but did not compete as well as a 100-fold excess. Thus, rRB38 shows specificity for the *psbA* 5′-UTR but also has some affinity for other chloroplast 5′-UTRs. While the *petD* 5′-UTR has some U-rich regions, a comparison of the *psbA* and *petD* 5′-UTRs showed no significant homology.

RB38 Is Imported into C. reinhardtii Chloroplasts. The nuclear-encoded RB38 protein contains a predicted chloroplast import sequence (Figure 2A). To determine whether recombinant RB38 could be imported into isolated chloroplasts, cw15 chloroplasts were purified and used in protein import assays with an IVT precursor of RB38. As shown in Figure 7, the major product of the RB38 IVT reaction is a polypeptide of approximately 42 kDa. Incubation of this protein for 10 min with isolated cw15 chloroplasts revealed that the protein bound to the chloroplast envelope and was subsequently imported into the chloroplast. The imported protein was resistant to degradation with thermolysin but was degraded after import if the chloroplasts were lysed by sonication and given a second thermolysin treatment. Sonication alone had no effect on the stability of the imported RB38. The competency of isolated chloroplasts for import of proteins was established by the efficient import of a *C. reinhardtii* OEE1 IVT product into chloroplasts (Figure 7). The imported RB38 protein was not processed to the mature 38 kDa protein during our in vitro import assay, suggesting that additional time, or other factors, might be required for RB38 processing following import. The failure of an imported protein to be processed during in vitro import has been observed with RB60 (*18*) and with a chloroplast-localized nuclear-encoded heat-shock protein 22 (CHS62; *52*).

DISCUSSION

We have cloned and characterized a novel mRNA binding protein from *C. reinhardtii* chloroplasts. This 38 kDa protein (RB38) was first identified as a member of an RNA binding complex that recognizes the 5′-UTR of the chloroplast-encoded *psbA* mRNA (*10*). Other members of this complex include RB47, an RNA binding protein of the PABP family (*16*), RB60, a PDI (*19*), and RB55, a protein with homology to nuclear assembly proteins (Mayfield et al., unpublished results). This protein complex exhibits light-regulated binding to the 5′-UTR of the chloroplast *psbA* mRNA, and this binding can be modulated by both redox potential and ADP-dependent protein phosphorylation (*10*, *11*, *13*, *19*). Genetic analysis has shown that RB47 accumulation is required for *psbA* translation initiation (*12*, *17*), while biochemical studies have shown that RB60 is capable of modulating RB47 binding activity in a redox-dependent manner (*19*).

RB38 is capable of binding directly to the *psbA* mRNA (Figure 6B). However, RB38 binds to a different region of the *psbA* 5′-UTR than RB47, as the binding of each protein is affected by a different set of competitors. Binding of rRB38 is competed with poly(U) RNA but not with poly(A) RNA, while the binding of rRB47 is competed by the addition of poly(A) RNA but not by the addition of poly(U) RNA (Figure 6C; *16*). RB47 binding protects a 36-base fragment that is A-rich and located adjacent to the RBS. As rRB38 binding is competed by poly(U), and more efficiently with the full-length *psbA* 5′-UTR than the short 5′-UTR (30 nucleotides between the RBS to the AUG) (Figure 6D), the protein would appear to be binding to a U-rich region that is immediately 5′ of the RBS (*15*). It should be noted that the short *psbA* 5′-UTR may not fold into the same secondary structure as the full-length 5′-UTR. Thus, it is possible that

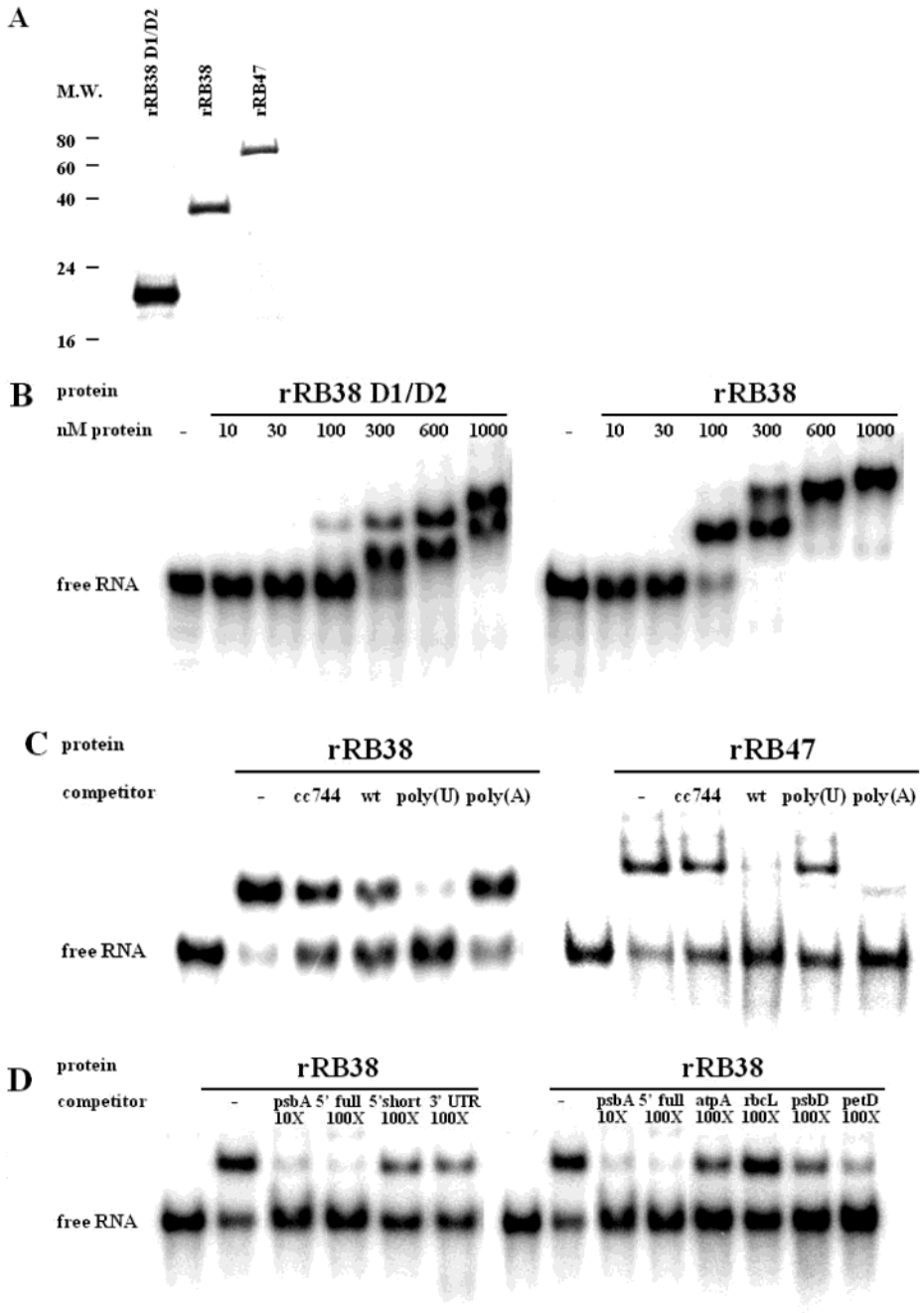

FIGURE 6: Gel mobility shift assays of rRB38 D1/D2, rRB38, and rRB47. (A) Coomassie-stained SDS−PAGE gel of the purified RB38 domains 1 and 2 (rRB38 D1/D2), all four domains (rRB38), and full-length rRB47 (including the RNA binding domains and the carboxy unique domain; 70 kDa total). The corresponding molecular weights are indicated. (B) Affinity of rRB38 D1/D2 and rRB38 for the *psbA* 5′-UTR. (C) Binding of rRB38 (100 nM) and rRB47 (150 nM) to the *psbA* 5′-UTR in the presence of no competitor or 500 ng of unlabeled cc744 and wt total RNA or 100 ng of poly(U) and poly(A) RNA. (D) Affinity of rRB38 (150 nM) for the *psbA* 5′-UTR in the presence of 10× and 100× unlabeled competitor RNA: the full-length *psbA* 5′-UTR, short *psbA* 5′-UTR, *psbA* 3′-UTR, and 5′-UTRs of *atpA*, *rbcL*, *psbD*, and *petD*.

the differences observed in binding of the full-length and short *psbA* 5′-UTRs are the result of changes in RNA folding and do not directly relate to sequence element differences between these RNAs.

The decreased binding of rRB38 in the presence of 500 ng of cc744 RNA, which does not affect RB47 binding (Figure 6C), suggests that RB38 may be a more general RNA binding protein than RB47. Competition gel shift assays show that RB38 is specific for the *psbA* 5′-UTR over the *rbcL*, *atpA*, and *psbD* 5′-UTRs (Figure 6D). However, the competition with *petD* shows that RB38 may have affinity for other chloroplast 5′-UTRs.

RB38 contains four 70 amino acid repeats, each containing a number of basic and aromatic amino acids (Figure 2). These repeats are suggestive of RNA binding proteins, as is the high percentage of arginine and aromatic amino acids, which often participate directly in binding to the RNA. RB38, however, cannot be placed into any one category of previously described RNA binding proteins. Database searches have identified a number of hypothetical proteins from *A. thaliana* that have a significant degree of similarity with RB38, but none of these hypothetical proteins have been assigned a biochemical function. Fold recognition methods predict domains with mixed α and β architecture, typical of

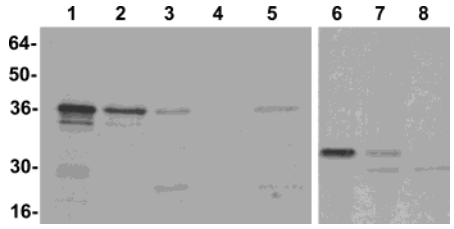

FIGURE 7: SDS−PAGE of proteins following import of in vitro translated RB38 and OEE1 proteins into isolated *C. reinhardtii* chloroplasts. In vitro translated RB38 protein (left panel) and OEE1 protein (right panel) were incubated with chloroplasts. Lanes: 1, RB38 IVT product; 2, chloroplasts incubated with the RB38 IVT product; 3, chloroplast import of RB38 IVT product followed by thermolysin treatment; 4, chloroplast import of RB38 IVT product followed by thermolysin treatment, sonication, and second thermolysin treatment; 5, chloroplast import of RB38 IVT product followed by thermolysin treatment and sonication; 6, OEE1 IVT product (33 kD); 7, chloroplasts incubated with OEE1 IVT product; 8, wt chloroplast import of OEE1 IVT product followed by thermolysin treatment. The molecular weight of the markers is indicated at the left.

RNA binding proteins, and subsequent structural modeling of RB38 suggests the protein may be a divergent member of the RBD superfamily, most similar to KH domains. KH domain proteins have been shown to recognize pyrimidine-rich sequences within their target RNAs (*53*) and also to interact with the 5′-UTR of several mRNAs to regulate translation and stability of the mRNA (*54*−*56*). Furthermore, RB38 contains several R/G/P-rich areas in the putative loop regions, motifs that have previously been associated with RNA binding regions of proteins (*57*, *58*). Thus, RB38 appears to be a multidomain protein with a mixed α and β architecture, containing RBD variants, along with perhaps other RNA recognition motifs, that recognize the U-rich RNA sequences in the 5′-UTR of the *psbA* mRNA.

To date, several models have been put forth for translational regulation of the chloroplast *psbA* mRNA based on the proposed activity of RB47 and RB60 (*18*−*20*). In each of these models, redox potential generated from the light reactions of photosynthesis is donated to RB60 (PDI), which in turn catalyzes the breakage of disulfide bonds in RB47 (PABP). RB47, in a reduced form, is then able to recognize the A-rich 5′-UTR of *psbA*, allowing *psbA* translation to occur. Fong et al. (*20*) showed that single cysteine residues located in RRM2 and RRM3 of RB47 were necessary for disulfide bond formation and redox regulation. RB38 lacks cysteine residues that could act in disulfide bond formation, and in vitro analysis failed to identify any redox regulation of RB38 RNA binding activity (data not shown).

Several critical regulatory elements in the *psbA* 5′-UTR have been identified as being important for *psbA* translation, including a RBS, an AU-box, and a stem−loop structure 5′ and adjacent to the RBS (*15*, *59*). In order for the *psbA* mRNA to be translated efficiently into the D1 protein, the stem−loop must be removed by a processing event during mRNA maturation (*14*). The binding of RB38 to the U-rich region, immediately upstream of the stem−loop structure, in the *psbA* 5′-UTR may facilitate this processing event in conjunction with the early stages of ribosome assembly. Should RB38 bind the U-rich region between the RBS and the start of translation, it may aid the 30S ribosomal subunit to bridge a 30-nt sequence that is too great for it to span alone. This interaction may be required to localize the ribosome to the initiation codon on the *psbA* mRNA in a manner that is not yet understood. Several sequences have been characterized in prokaryotes that enhance translation, including a poly(U) binding site for the ribosomal S1 protein (*60*).

The prospect of determining the function of RB38 through additional experimentation, including the production of a RB38 knockout strain in *C. reinhardtii*, will allow us to explore the biological role of this novel RNA binding protein. In addition, structural characterization of this novel fold by X-ray crystallography will provide new insights into the principles, patterns, and diversity of strategies associated with RNA recognition.

ACKNOWLEDGMENT

We thank Mike Hanson for assistance in making the table of contents graphic. The Center for Molecular Structure is supported by the W. M. Keck Foundation and the California State University Program for Education and Research in Biotechnology.

REFERENCES

1. Siomi, H., and Dreyfuss, G. (1997) RNA-binding proteins as regulators of gene expression, *Curr. Opin. Genet. Dev. 7*, 345−353.
2. Gingras, A. C., Raught, B., and Sonenberg, N. (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation, *Annu. Rev. Biochem. 68*, 913−963.
3. Gallie, D. R. (1998) A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation, *Gene 216*, 1−11.
4. Wells, D. R., Tanguay, R. L., Le, H., and Gallie, D. R. (1998) HSP101 functions as a specific translational regulatory protein whose activity is regulated by nutrient status, *Genes Dev. 12*, 3236−3251.
5. McCarthy, J. E., and Gualerzi, C. (1990) Translational control of prokaryotic gene expression, *Trends Genet. 6*, 78−85.
6. Deo, R. C., Bonanno, J. B., Sonenberg, N., and Burley, S. K. (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein, *Cell 98*, 835−845.
7. Somanchi, A., and Mayfield, S. P. (1999) Nuclear-chloroplast signalling, *Curr. Opin. Plant Biol. 2*, 404−409.
8. Zerges, W. (2000) Translation in chloroplast, *Biochimie 82*, 583−601.
9. Barkan, A., and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression, *Biochimie 82*, 559−572.
10. Danon, A., and Mayfield, S. P. (1991) Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins, *EMBO J. 10*, 3993−4001.
11. Danon, A., and Mayfield, S. P. (1994) Light-regulated translation of chloroplast messenger RNAs through redox potential, *Science 266*, 1717−1719.
12. Yohn, C. B., Cohen, A., Danon, A., and Mayfield, S. P. (1996) Altered mRNA binding activity and decreased translational initiation in a nuclear mutant lacking translation of the chloroplast *psbA* mRNA, *Mol. Cell. Biol. 16*, 3560−3566.
13. Danon, A., and Mayfield, S. P. (1994) ADP-dependent phosphorylation regulates RNA-binding in vitro: implications in light-modulated translation, *EMBO J. 13*, 2227−2235.
14. Bruick, R. K., and Mayfield, S. P. (1998) Processing of the *psbA* 5′ untranslated region in *Chlamydomonas reinhardtii* depends on factors mediating ribosome association, *J. Cell Biol. 143*, 1145−1153.
15. Mayfield, S. P., Cohen, A., Danon, A., and Yohn, C. B. (1994) Translation of the *psbA* mRNA of *Chlamydomonas reinhardtii* requires a structured RNA element contained within the 5′ untranslated region, *J. Cell Biol. 127*, 1537−1545.
16. Yohn, C. B., Cohen, A., Danon, A., and Mayfield, S. P. (1998) A poly(A) binding protein functions in the chloroplast as a message-specific translation factor, *Proc. Natl. Acad. Sci. U.S.A. 95*, 2238−2243.

17. Yohn, C. B., Cohen, A., Rosch, C., Kuchka, M. R., and Mayfield, S. P. (1998) Translation of the chloroplast *psbA* mRNA requires the nuclear-encoded poly(A)-binding protein, RB47, *J. Cell Biol. 142*, 435−442.
18. Trebitsh, T., Levitan, A., Sofer, A., and Danon, A. (2000) Translation of chloroplast *psbA* mRNA is modulated in the light by counteracting oxidizing and reducing activities, *Mol. Cell. Biol. 20*, 1116−1123.
19. Kim, J., and Mayfield, S. P. (1997) Protein disulfide isomerase as a regulator of chloroplast translational activation, *Science 278*, 1954−1957.
20. Fong, C. L., Lentz, A., and Mayfield, S. P. (2000) Disulfide bond formation between RNA binding domains is used to regulate mRNA binding activity of the chloroplast poly(A)-binding protein, *J. Biol. Chem. 275*, 8275−8278.
21. Harris, E. H. (1989) *The Chlamydomonas Sourcebook*, Academic Press, San Diego, CA.
22. Sambrook, J., Fritsch, E. F., and Maniatis (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Mayfield, S. P., Bennoun, P., and Rochaix, J. D. (1987) Expression of the nuclear encoded OEE1 protein is required for oxygen evolution and stability of photosystem II particles in *Chlamydomonas reinhardtii*, *EMBO J. 6*, 313−318.
24. Cohen, A., Yohn, C. B., Bruick, R. K., and Mayfield, S. P. (1998) Translational of chloroplast gene expression in *Chlamydomonas reinhardtii*, *Methods Enzymol. 297*, 192−208.
25. Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. (1997) Gapped BLAST and PSI−BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
26. Altschul, S., Gish, W., Miller, W., Meyers, E., and Lipman, D. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403−410.
27. Contreras-Moreira, B., and Bates, P. (2002) Domain Fishing and 3D-JIGSAW, web tools for protein comparative modeling, *Proceedings of the Symposium on Bioinformatics and Computational Biology* (de Las Rivas, J., Ed.) p 95, Universidad de Salamanca, Salamanca, Spain.
28. Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile based neural networks, *Methods Enzymol. 266*, 525−539.
29. Karplus, K., Karchin, R., Draper, J., Casper, J., Mandel-Gutfreund, Y., Diekhans, M., and Hughey, R. (2003) Combining local structure, fold recognition, and new methods for protein structure prediction, *Proteins: Struct., Funct., Genet.* (in press).
30. MacCallum, R., Kelley, L., and Sternberg, M. (2000) SAWTED: Structure assignment with text description, *Bioinformatics 16*, 125−129.
31. Jones, D. (1999) Protein secondary structure prediction based on position j-specific scoring matrices, *J. Mol. Biol. 292*, 195−202.
32. Rychlewski, L., Jaroszewski, L., Li, W., and Goszik, A. (2000) Comparison of sequence profiles. Stategies for structural predictions using sequence information, *Protein Sci. 9*, 232−241.
33. Jones, D. (1999) GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences, *J. Mol. Biol. 287*, 797−815.
34. McGuffin, L., and Jones, D. (2003) Improvement of the GenTHREADER method for genomic fold recognition, *Bioinformatics 19*, 874−881.
35. McGuffin, L., Bryson, K., and Jones, D. (2000) The PSIPRED protein structure prediction server, *Bioinformatics 16*, 404−405.
36. Lo Conte, L., Brenner, S., Hubbard, T., Chothia, C., and Murzin, A. (2002) SCOP database in 2002: refinements accommodate structural genomics, *Nucleic Acid Res. 30*, 264−267.
37. Meller, J., and Elber, R. (2001) Linear programming and a double statistical filter for protein threading protocols, *Proteins 45*, 241−261.
38. Murzin, A., Lo Conte, L., Andreeva, A., Howorth, D., Ailey, B., Brenner, S., Hubbard, T., and Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures, *J. Mol. Biol. 247*, 536−540.
39. Nickelsen, J., van Dillewijn, J., Rahire, M., and Rochaix, J. D. (1994) Determinants for stability of the chloroplast *psbD* RNA are located within its short leader region in *Chlamydomonas reinhardtii*, *EMBO J. 13*, 3182−3191.
40. Lawrence, S. D., and Kindle, K. L. (1997) Alterations in the *Chlamydomonas* plastocyanin transit peptide have distinct effects on in vitro import and in vivo protein accumulation, *J. Biol. Chem. 272*, 20357−20363.
41. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984) Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter, *Nucleic Acids Res. 12*, 7035−7056.
42. Mayfield, S. P., Rahire, M., Frank, G., Zuber, H., and Rochaix, J. D. (1987) Expression of the nuclear gene encoding oxygen-evolving enhancer protein 2 is required for high levels of photosynthetic oxygen evolution in *Chlamydomonas reinhardtii*, *Proc. Natl. Acad. Sci. U.S.A. 84*, 749−753.
43. Cline, K., Werner-Washburne, M., Lubben, T. H., and Keegstra, K. (1985) Precursors to two nuclear-encoded chloroplast proteins bind to the outer envelope membrane before being imported into chloroplasts, *J. Biol. Chem. 260*, 3691−3696.
44. Perez-Canadillas, J. M., and Varani, G. (2001) Recent advances in RNA-protein recognition, *Curr. Opin. Struct. Biol. 11*, 53−58.
45. Fischer, D. (2000) Hybrid Fold Recognition: Combining Sequence Derived Properties with Evolutionary Information, in *Pacific Symposium of Biocomputing*, World Scientific, Hawaii.
46. Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure, *J. Mol. Biol. 313*, 903−919.
47. Connolly, K. M., Wojciak, J. M., and Clubb, R. T. (1998) Site-specific DNA binding using a variation of the double stranded RNA binding motif, *Nat. Struct. Biol. 5*, 546−550.
48. Canutescu, A., Shelenkov, A., and Dunbrack, R. J. (2003) A graph-theory algorithm for protein side-chain prediction, *Protein Sci.* (in press).
49. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
50. Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins, *Proteins 17*, 355−362.
51. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−2563.
52. Grimm, B., Ish-Shalom, D., Even, D., Glaczinski, H., Ottersbach, P., Ohad, I., and Kloppstech, K. (1989) The nuclear-coded chloroplast 22-kDa heat-shock protein of *Chlamydomonas*. Evidence for translocation into the organelle without a processing step, *Eur. J. Biochem. 182*, 539−546.
53. Jensen, K. B., Musunuru, K., Lewis, H. A., Burley, S. K., and Darnell, R. B. (2000) The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5740−5745.
54. Graff, J., Cha, J., Blyn, L. B., and Ehrenfeld, E. (1998) Interaction of poly(rC) binding protein 2 with the 5′ noncoding region of hepatitis A virus RNA and its effects on translation, *J. Virol. 72*, 9668−9675.
55. Gamarnik, A. V., and Andino, R. (2000) Interactions of viral protein 3CD and poly(rC) binding protein with the 5′ untranslated region of the poliovirus genome, *J. Virol. 74*, 2219−2226.
56. Stefanovic, B., Hellerbrand, C., and Brenner, D. A. (1999) Regulatory role of the conserved stem-loop structure at the 5′ end of collagen alpha1(I) mRNA, *Mol. Cell. Biol. 19*, 4334−4342.
57. Draper, D. E. (1999) Themes in RNA-protein recognition, *J. Mol. Biol. 293*, 255−270.
58. Cheng, A. C., Calabro, V., and Frankel, A. D. (2001) Design of RNA-binding proteins and ligands, *Curr. Opin. Struct. Biol. 11*, 478−484.
59. Hirose, T., and Sugiura, M. (1996) Cis-acting elements and trans-acting factors for accurate translation of chloroplast *psbA* mRNAs: development of an in vitro translation system from tobacco chloroplasts, *EMBO J. 15*, 1687−1695.
60. McCarthy, J. E., and Brimacombe, R. (1994) Prokaryotic translation: the interactive pathway leading to initiation, *Trends Genet. 10*, 402−407.
61. Abagyan, R., Totrov, M., and Kuznetsov, D. (1994) ICM: a new method for protein modeling and design. Applications to docking and structure prediction from the distorted native conformation, *J. Comput. Chem. 15*, 488−506.

BI035909J