Published online 29 September 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 18* ***5337–5351***
doi:10.1093/nar/gkl611

A nucleosome assembly protein-like polypeptide binds to chloroplast group II intron RNA in *Chlamydomonas reinhardtii*

Stephanie Glanz, Astrid Bunse, Andrea Wimbert, Carsten Balczun and Ulrich Kück*

Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum, D-44780 Bochum, Germany

Received May 2, 2006; Revised and Accepted August 4, 2006

ABSTRACT

In the unicellular green alga *Chlamydomonas reinhardtii*, the chloroplast-encoded *tscA* RNA is part of a tripartite group IIB intron, which is involved in *trans*-splicing of precursor mRNAs. We have used the yeast three-hybrid system to identify chloroplast group II intron RNA-binding proteins, capable of interacting with the *tscA* RNA. Of 14 candidate cDNAs, 13 encode identical polypeptides with significant homology to members of the nuclear nucleosome assembly protein (NAP) family. The RNA-binding property of the identified polypeptide was demonstrated by electrophoretic mobility shift assays using different domains of the tripartite group II intron as well as further chloroplast transcripts. Because of its binding to chloroplast RNA it was designated as NAP-like (cNAPL). *In silico* analysis revealed that the derived polypeptide carries a 46 amino acid chloroplast leader peptide, in contrast to nuclear NAPs. The chloroplast localization of cNAPL was demonstrated by laser scanning confocal fluorescence microscopy using different chimeric cGFP fusion proteins. Phylogenetic analysis shows that no homologues of cNAPL and its related nuclear counterparts are present in prokaryotic genomes. These data indicate that the chloroplast protein described here is a novel member of the NAP family and most probably has not been acquired from a prokaryotic endosymbiont.

INTRODUCTION

Group II introns have been detected in eukaryotic organelles as well as in eubacterial genomes as intervening sequences of protein, tRNA or rRNA coding genes. Common to all group II introns is a conserved secondary structure that consists of six double-helical domains (DI-DVI) radiating from a central wheel (1). Group II introns splice via two sequential transesterifications that *in vitro* occur in some cases autocatalytically. *In vivo*, however, splicing is dependent on protein co-factors as was shown by mutant analyses (2,3).

In *Chlamydomonas reinhardtii* chloroplasts, two *trans*-splicing introns were found in the *psaA* gene encoding the major P700 chlorophyll a/b-binding protein (4). In the case of the first intron of the *psaA* gene, three independently transcribed RNAs associate via tertiary interactions to form a functional group II intron, resulting in *trans*-splicing of the flanking exon 1 and exon 2 (5). Part of this tripartite group II intron is the *tscA* RNA that is processed from a chloroplast-encoded precursor RNA. Secondary structure predictions revealed that *tscA* contains domain DII and DIII as well as partial domains DI and DIV of the conserved secondary group II intron structure (6).

Mutant work in *C.reinhardtii* has shown that >14 nuclear genes affect the chloroplast *trans*-splicing reaction, and complementation of some of the characterized mutants led to the identification of the molecular determinants. Although some of the identified polypeptides are related to proteins involved in nucleic acid metabolisms, this function seems to be non-essential for intron splicing (7,8). Similar to nuclear mRNA splicing, these protein components might be part of a chloroplast splicing complex, because they were found in large stromal or membrane-bound RNA-containing complexes (9–11).

To define further components of a putative 'chloroplast spliceosome' in *C.reinhardtii*, we used the *tscA* RNA as bait to identify novel intron RNA-binding proteins. Using the yeast three-hybrid system, we isolated a polypeptide that seems to be a novel member of the multifunctional nucleosome assembly protein (NAP) family. These well-conserved eukaryotic histone chaperones facilitate, for example, the nucleosome assembly and remodelling of chromatin, and are implicated in transcriptional regulation and cell cycle regulation (12,13). Here, we assign a novel function to a NAP-like protein. Laser scanning confocal fluorescence microscopy (LSCFM) and *in vitro* binding

*To whom correspondence should be addressed. Tel: +49 234 3226212; Fax: +49 234 3214184; Email: ulrich.kueck@ruhr-uni-bochum.de
Present address:
Astrid Bunse, Qiagen GmbH, Qiagen Straße 1, D-40724 Hilden, Germany

Table 1. Plasmids used in this work

Plasmid	Characteristics	Reference
Yeast three-hybrid		
pADRevM10	pACTII derivative, expression of fusion protein GAL4-AD and HIV1 RevM10	(21)
pDBRevM10	Yeast shuttle vector (TRP1, LEU2), cloning of hybrid RNA transcription unit cassette of pPGKRRE and expression cassette of fusion protein GAL4-DB and HIV1 RevM10	(21)
pGAD10 derivatives	cDNA fragments of *C.reinhardtii* CC406 cw15 in pGAD10 (BD Biosciences Clontech)	(20)
pPGKRRE	Intermediate to clone hybrid RNAs as a transcription unit cassette into pDBRevM10	(21)
pRev3′Lhcb	Fusion protein GAL4-DB and RevM10; hybrid RNA RRE and *Lhcb1* 3′-UTR (pBQ2.4) (37)	(8)
pRevpsaADV+VI	Fusion protein GAL4-DB and RevM10; hybrid RNA RRE and domains DV+VI of *psaA* intron 1	This work
pRevR2	RRE–RRE transcription unit cassette	(21)
pRevrI1DIV-VI	Fusion protein GAL4-DB and RevM10; hybrid RNA RRE and domains DIV-VI of intron rI1 of *S.obliquus* (36)	(8)
pRevtscA1	Fusion protein GAL4-DB and RevM10; hybrid RNA RRE and processed *tscA* RNA (6)	(8)
cNAPL cDNA		
pGAD6-19	1.6 kb *cNapl* cDNA fragment in pGAD10	This work
EMSA		
pQE70-cNAPL	1 kb SphI/BglII coding sequence of cNAPL in pQE70 with His_6 tag (Qiagen)	This work
pdI	347 bp fragment of domain DI of intron rI1 of *S.obliquus* in pT3T7BM/SmaI	This work
pdII+III	139 bp fragment of domain DII and DIII of intron rI1 of *S.obliquus* in pT3T7BM/SmaI	This work
pdIV	57 bp fragment of domain DIV of intron rI1 of *S.obliquus*, complementary oligonucleotides in pT3T7BM	(22)
pdV+VI	73 bp fragment of domain DV and DVI of intron rI1 of *S.obliquus* in pT3T7BM/SmaI	(22)
p734	211 pb fragment of *rps4* 5′-UTR in plasmid pBIIKS+	(22)
pT3T73′UTRLhcb1T7	109 nt EcoRI/BamHI fragment of *Lhcb1* 3′-UTR in pT3T7BM	This work
pT3T7psaADV+VIT7	400 nt MluI fragment of pRevpsaADV+VI in pT3T7BM/SmaI (Boehringer, Mannheim, Germany)	(23)
pT3T7Rps18T7	136 nt EcoRI/BamHI fragment of exon 2 of *Rps18* (46) in pT3T7BM	This work
pT3T7Tba1AT7	127 nt EcoRI/BamHI fragment of exon 3 of *Tba1A* in pT3T7BM	This work
pT3T7tscADIIT7	156 nt EcoRI/BamHI fragment of domain DII of *tscA* RNA in pT3T7BM	This work
pT3T7tscADIIIT7	193 nt EcoRI/BamHI fragment of domain DIII of *tscA* RNA in pT3T7BM	This work
pT3T7tscADII+IIIT7	338 nt EcoRI/BamHI fragment of domain DII and DIII of *tscA* RNA in pT3T7BM	This work
LSCFM		
pCr1	*HSP70A/RBCS2* promoter (pCB740) (49) and *Lhcb1* 3′-UTR terminator, *ASL* for selection	This work
pCr1g	0.7 kb NheI/BglII fragment of *cgfp* (47) in pCr1	This work
pCr1g_cNaplN84	0.3 kb NheI fragment, N-terminal 84 amino acids of cNAPL in pCr1g encoding for fusion protein N84-cGFP	This work
pCr1g_cNaplN84Δ8-39	0.2 kb NheI fragment, N-terminal 84 amino acids of cNAPL without targeting signal in pCr1g encoding for fusion protein N84Δ8-39-cGFP	This work
pCr1g_Rps18	0.5 kb NheI fragment of *Rps18* coding sequence in pCr1g encoding for fusion protein Rps18-cGFP	This work
pGAD-Rps18	*Rps18* cDNA fragment in pGAD424 (BD Biosciences Clontech)	This work
pMF59	Encoding for fusion protein Ble-cGFP	(17)

AD, DNA activation domain; *ASL*, arginiosuccinate lyase; DB, DNA-binding domain; EMSA, electromobility shift assay; LSCFM, laser scanning confocal fluorescence microscopy

assays demonstrate that chloroplasts of *C.reinhardtii* contain a NAP-like protein that specifically binds to organellar group II intron RNA and U-rich chloroplast transcripts. Our data further support the view that during evolution, chloroplasts have acquired novel nuclear components that most probably were not delivered by endosymbiont gene transfer (14).

MATERIALS AND METHODS

Strains, culture conditions and transformation

C.reinhardtii cell wall depleted, arginine auxotroph strain arg^{b-e}cw15 as well as wild-type strain CC406 cw15 mt- were obtained from the *Chlamydomonas* Center (Duke University, Durham, NC, USA) and were grown as described on Tris–acetate–phosphate (TAP) medium (15). When required, the medium was supplemented with 50 μg arginine per ml. Nuclear transformation was carried out using the glass bead method (16) and was performed according to (8) with the following modifications. A total of 1–3 × 10^6 cells/ml were incubated with 1–5 μg supercoiled or SpeI digested pMF59 (17) plasmid DNA. After transformation, cells were spread on solid TAP medium and incubated in continuous light at 25°C. For selection of transformants carrying pMF59, medium contained 5 μg/ml of zeocin (Invitrogen, Karlsruhe, Germany).

Recombinant plasmids

All recombinant plasmids used for *in vitro* transcription, PCR analysis, protein synthesis, yeast three-hybrid screening, hybridization or generation of transgenic algal strains are listed in Table 1. Sequences of all oligonucleotides are given in Table 2.

Molecular biological techniques

Procedures for standard molecular techniques were performed as described previously (18). *Escherichia coli* strain

Table 2. Oligonucleotide primer pairs used in PCR and RT–PCR experiments

Oligonucleotide	Specific sequence (5′→3′)	Specificity
2428_NheI,BglII,SmaI	GCT AGC AGA TCT CCC GGG TTT GTA TGA GCG TAT AAG CTC TGC	5′-region of *Lhcb1* 3′-UTR
2429_EcoRV	GAT ATC ATC GAT GGC CTA CCC AAA CCC C	3′-region of *Lhcb1* 3′-UTR
for_6-19_BspHI	TAT TCA TGA TGG CAC TAG CTC TGC TCA CTC G	*cNapl*, RT–PCR
rev_6-19_2xMet_HindIII	ATA AAG CTT CTA CAT CAT GTA CTC GCC CTC CTC GTA CTC G	*cNapl*, RT–PCR
for_6-19_ SphI	GCA TGC GCA TGG CAC TAG CTC TGC TCA C	*cNapl*
rev_6-19_BglII	AGA TCT GTA CTC GCC CTC CTC GTA C	*cNapl*
for_ATG_cgfp	CTA GCT AGC ATG GCC AAG GGC GAG GAG	*cgfp*
rev_cgfp	GAA GAT CTT TAC TTG TAC AGC TCG TCC ATG C	*cgfp*
for_cNapl_N84	CTA GCT AGC ATG GCA CTA GCT CTG CTC AC	*cNapl*N84
rev_cNapl_N84	CTA GCT AGC GGC GGT GTT GAG CTC CTC	*cNapl*N84, *cNapl*N84Δ8-39
for_cNapl_N84Δ8-39	CTA GCT AGC ATG GCA CTA GCT CTG CTC ACT TTC CTG GCA GTG GTT GGC GCC	*cNapl*N84Δ8-39
for_3′Lhcb_EcoRI	CCA TCA AGG ACC AGG TCA TC	Fragment of *Lhcb1* 3′-UTR
rev_3′Lhcb_BamHI	AGG AAT TCA AGA GCC CAT TCC AAA GTC	Fragment of *Lhcb1* 3′-UTR
Lhcbm6 T7-1	GTA ATA CGA CAT CAC TAT AGG GCC TGT TTG GCG CTT	Fragment of *Lhcb1* 5′-UTR
Lhcbm6-2	CAT TCC TGC ACG CTC GCT TG	Fragment of *Lhcb1* 5′-UTR
T7-psbA5′	GTA ATA CGA CTC ACT ATA GGG TAC CAT GCT TTT AAT	Fragment of *psbA* 5′-UTR
psbA3′-2054	GAT CCA TGG TCA TAT GTT AAT TTT TTT AAA G	Fragment of *psbA* 5′-UTR
su3131	TGT GCG TTT CTC TTG ATA TGT ACC G	Fragment of *psbD* 5′-UTR
2126	TAA TAC GAC TCA CTA TAG GGA CAC AAT GAT TAA AAT TAA A	Fragment of *psbD* 5′-UTR
T7-rbcL5′	GTA ATA CGA CTC ACT ATA GGG TAT GCT CGA CTG ATA AGA C	Fragment of *rbcL* 5′-UTR
rbcL3′	CTG CTT TAG TTT CTG TTT GTG GAA CC	Fragment of *rbcL* 5′-UTR
for_Rps18_NheI	CTG CTA GCA TGG GCT CTC TCG TCC A	*Rps18*
rev_Rps18_NheI	GAG CTA GCC TTC TTC TTG GCG ACA CC	*Rps18*
for_Rps18_EcoRI	AGG AAT TCT CTG CGT CTG CTG AAC ACT AAC	Fragment of exon 2 of *Rps18*
rev_Rps18_BamHI	CGG GAT CCA GGT CAA CCT CAG CCT TCT T	Fragment of exon 2 of *Rps18*
for_tscADII	AGG AAT TCA AAT TTT TAG TAA TGT TAA AG	Domain DII of *tscA* RNA
rev_tscADII	CGG GAT CCC CTT AAT TTT TAG AAA ATG	Domain DII of *tscA* RNA
for_tscADIII	AGG AAT TCT AAA CAA CAA GTA ATG CC	Domain DIII of *tscA* RNA
rev_tscADIII	CGG GAT CCT AAA CAA TAA AGT TTT TCA A	Domain DIII of *tscA* RNA
for_a1Tub_EcoRI	AGG AAT TCC GGC TTC AAG TGC GGT ATC AAC	Fragment of exon 3 of *Tba1A*
rev_a1Tub_BamHI	CGG GAT CCT CGC CGA TAG CAG TGC TGT T	Fragment of exon 3 of *Tba1A*

XL1-blue MRF′ served as host for general plasmid construction and maintenance (19). The *C.reinhardtii* cDNA library contains double-stranded cDNA fragments with EcoRI adaptors cloned into the EcoRI site of pGAD10 in DH10B (BD Biosciences Clontech, Heidelberg, Germany) (20). pGAD10 derivatives were used in the yeast three-hybrid screen. *C.reinhardtii* total RNA was prepared according to (4) and RNA blot experiments were carried out according to (8). Northern analyses were performed with a radioactively labelled 1690 bp pGAD6-19/EcoRI *cNapl*-specific probe and a 460 bp pGAD-Rps18/EcoRI–BamHI *Rps18*-specific probe. cDNA was prepared using Omniscript™ Reverse Trancriptase (Qiagen, Hilden, Germany) and purified with NucAway™ Spin Columns (Ambion, Austin, TX, USA) according to the manufacturer's protocols. For PCRs, *Taq* polymerase (Eppendorf, Hamburg, Germany), Hot Master™ *Taq* DNA Polymerase (Eppendorf) and the Expand Long Template PCR System (Roche, Mannheim, Germany) were used. Oligonucleotide primers applied for conventional PCR are listed in Table 2.

The yeast three-hybrid system

The yeast three-hybrid analysis was performed as described previously (8). Plasmids are listed in Table 1. Plasmids pADRevM10 and pRevR2 were generous gifts from Dr U. Putz (University of Hamburg) (21). Plasmid pRevtscA1 (8) was used to screen a *C.reinhardtii* cDNA library (Table 1).

Electrophoretic mobility shift assays (EMSAs)

Recombinant plasmid pQE70-cNAPL (see above) was used for synthesis of cNAPL in *E.coli*. Recombinant His-tagged cNAPL protein was purified from *E.coli* with Ni^{2+}-NTA-agarose resin according to the manufacturer's protocols (Qiagen). For RNA mobility shift assays, uniformly ^{32}P-UTP-labelled run-off transcripts served as substrate RNAs and were generated by *in vitro* transcription of plasmids as given in Table 1. *In vitro* transcription and EMSAs were performed as previously reported (22,23). Unlabelled competitor RNAs and non-specific competitor RNA, derived from plasmid pBIIKS+, were synthesized as described (22). DNA mobility shift assays were done essentially as described previously (24) with the following modifications. The probes were amplified by PCR with appropriate plasmid DNA or genomic DNA of strain arg-cw15 in the presence of [^{32}P]dATP. The labelled probes were purified by denaturing PAGE. Binding reactions were carried out by incubating 30 fmol of the probes with 15 μg of recombinant cNAPL. The reaction mixtures were separated on a 5% non-denaturing polyacrylamide gel.

Laser scanning confocal fluorescence microscopy

The fluorescence emissions of transformed *C.reinhardtii* cells were analysed by LSCFM, using a Zeiss LSM 510 META microscopy system (Carl Zeiss, Jena, Germany) based on an Axiovert inverted microscope. cGFP and plastids were excited with the 488 nm line of an

argon-ion laser. The fluorescence emission was selected by bandpass filter BP505-530 and longpass filter LP560, respectively, using beam splitters HFT UV/488/543/633 and NFT545.

Sequences, alignments and phylogenetic analysis

Sequences were retrieved from GenBank (NCBI) at http://www.ncbi.nlm.nih.gov, the *C.reinhardtii* JGI database v2.0 (DOE Joint Genome Institute) at http://genome.jgi-psf.org/chlre2/chlre2.home.html, the *Thalassiosira pseudonana* JGI database v1.0 at http://genome.jgi-psf.org/thaps1/thaps1.home.html, the *Porphyra yezoensis* EST index at http://www.kazusa.or.jp/en/plant/porphyra/EST/ (25,26) and the CyanoBase genome database at http://www.kazusa.or.jp/cyano/. Multiple sequence alignments of NAPs were performed using ClustalW software (27). Sequence alignments were displayed with GENEDOC (28) and refined manually. The phylogenetic tree was generated based on trimmed protein sequences with programs contained in the program package PHYLIP version 3.63 (29). An alignment with sequences from mature proteins was impossible since all proteins differ largely in length and some were only available from partial cDNAs (AU191247, BM448458) obtained from EST libraries. Analyses with the PHYLIP package were performed using PROTDIST/NEIGHBOR and the PROTPARS algorithms for distance matrix and maximum parsimony analysis, respectively. Support for the branches was estimated by bootstrap analysis of 2000 replicates generated with SEQBOOT, and a majority rule consensus tree was generated using the program CONSENSE. The resulting tree topology was then visualized in TREEVIEW (30). For analyses of putative targeting signal sequences, ChloroP (31), LOCtree (32), TargetP (33), GENOPLANTE^{TM} PREDOTAR (http://www.genoplante.com/) and PSORT (34) were used. Isoelectric point and sequence masses were calculated by the program WinPep 3.01 (35).

RESULTS

Identification of a *tscA*-specific RNA-binding protein using the yeast three-hybrid system

To identify RNA-binding proteins most probably involved in organellar group II intron splicing, we performed an *in vivo* screen using the yeast three-hybrid system. In this system, an RNA–protein interaction is detected by the reconstitution of the yeast transcriptional activator GAL4 and subsequent activation of reporter genes. The system is based on one hybrid RNA and two recombinant fusion proteins containing the GAL4 DNA-binding and GAL4 transcription activation domains, respectively. To isolate chloroplast intron RNA-binding proteins of *C.reinhardtii*, a cDNA library was screened using the yeast three-hybrid system developed by (21). In this system, the chimeric protein RevM10, fused to a gene fragment encoding the DNA-binding domain of GAL4 (GAL4-DB RevM10), binds to RRE, which is part of a chimeric RNA. The 3′ part of this RNA consists of the *tscA* RNA which functions as bait during the screen (RRE-*tscA*). Chimeric proteins serve as prey and are encoded by cDNA clones of a *C.reinhardtii* cDNA library. All cDNAs were fused to a gene fragment encoding the transcription activation domain of GAL4 (GAL4-AD). The reliability of this three-hybrid screen was tested with a recombinant RNA carrying two copies of RRE (RRE–RRE), co-expressed with the two RevM10 fusion proteins (AD-RevM10, DB-RevM10). RRE and RevM10 are known to interact with each other (21). As depicted in lane 1 of Figure 1, the interaction resulted in reconstitution of GAL4 activity.

For a global *C.reinhardtii* cDNA screen, the yeast strain CG1945 was co-transformed with pRevtscA1, expressing both the hybrid protein DB-RevM10 and the recombinant RNA. This strain served as host for transformation of the library of pGAD10-derived plasmids harbouring cDNA sequences of *C.reinhardtii*. A total of 2.1×10^7 clones were analysed and 14 clones were identified to bind to *tscA* RNA. Interestingly, of these 14 candidate cDNAs, 13 encoded

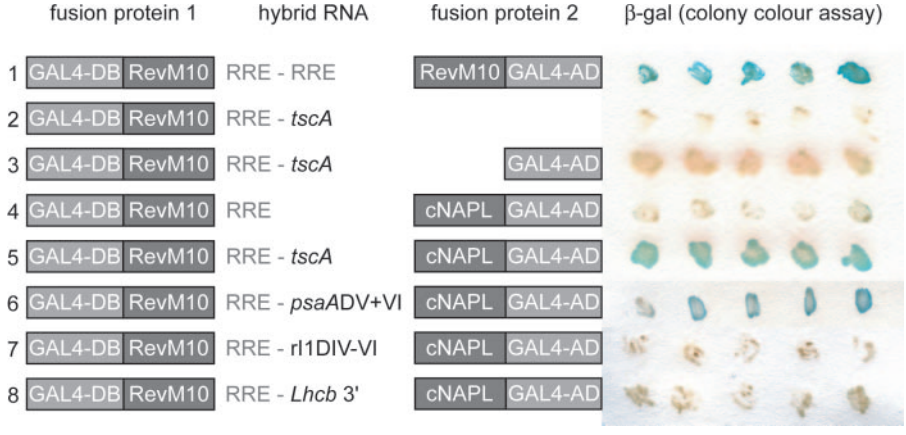

Figure 1. Schematic overview of selected constructs used in the yeast three-hybrid analyses. In each case, five representative yeast transformants were tested for their ability to activate β-galactosidase by a colony colour assay. Abbreviations: *Lhcb3′*, fragment of 3′-UTR of the *Lhcb1* transcript (37); cNAPL, chloroplast NAP-like; GAL4-AD, activation domain of GAL4 transcription factor; GAL4-DB, DNA-binding domain of GAL4 transcription factor; *psaA*DV+VI, domains DV and DVI of the first *psaA* intron (4); RevM10, RNA-binding protein RevM10 (21); rI1DIV-VI, domains DIV to DVI of intron rI1 (36); RRE, Rev responsive element that is specifically bound by RevM10 (21); *tscA*, processed *tscA* RNA (6).

identical polypeptides with significant homology to NAPs. Because of its binding to chloroplast intron RNA, the NAP-like polypeptide was designated as cNAPL.

To assess the specificity of the three-hybrid interaction between *tscA* RNA and cNAPL, several controls were carried out. The specificity of the *tscA* interaction was tested against two further hybrid RNAs. Recombinant control RNAs containing sequences of domains DIV to DVI of the heterologous intron rI1 (36) (RRE-rI1DIV-VI) and a fragment of the 3′-untranslated region (3′-UTR) of the *Lhcb1* transcript (37) (RRE-*Lhcb*3′) were fused separately to the RRE sequence, respectively. As shown in lanes 5, 7 and 8 of Figure 1, activity of the reporter β-galactosidase is dependent on the presence of the RRE-*tscA* hybrid RNA and the cNAPL-AD fusion protein. These results indicate a specific interaction of cNAPL with the *tscA* transcript. In addition, we were able to demonstrate that cNAPL also binds to domains DV and DVI of the first *psaA* intron (RRE-*psaA*DV+VI) as seen in lane 6 (Figure 1). To exclude one- or two-hybrid interaction or unspecific RNA interaction, appropriate negative controls were also performed in the filter assay, including different combinations of the three-hybrid components, represented in lanes 2–4 (Figure 1). In the absence of any one of the hybrid components, transformants displayed no β-galactosidase activity. Results of the three-hybrid analysis, therefore, clearly demonstrate the specific interaction of cNAPL with *tscA* RNA and the 3′ part of the first intron of *psaA* RNA.

In vitro binding of cNAPL to chloroplast transcripts

To confirm the specific interaction between cNAPL and *tscA* RNA in the yeast three-hybrid system, we performed *in vitro* binding studies using EMSA. For synthesis of cNAPL, the *cNapl* gene was fused with six histidine residues for expression in *E.coli*. After purification of the recombinant His-tagged protein with Ni^{2+}-NTA chromatography, cNAPL was incubated with *in vitro* transcribed radioactively labelled RNAs corresponding to domains DII and DIII of *tscA* RNA. As shown in Figure 2a, domains DII and DIII range from nt 75 to 218 and from 220 to 400, respectively. Together, they cover nt 75–400. RNAs were incubated in the absence or presence of recombinant cNAPL and fractionated on native gels (Figure 2b–i). In Figure 2b–d, increasing amounts of cNAPL (2–20 μg) were incubated with 30 fmol of *in vitro* synthesized transcripts of *tscA* intron RNA fragments comprising either domain DII, domain DIII or both domains DII+III. Arrows indicate RNA–protein complexes as multiple shifted bands that can be observed when the RNA is incubated with increasing amounts of cNAPL polypeptide.

To investigate the binding specificity of cNAPL, competition analyses were performed. The above-mentioned radiolabelled RNAs were incubated with a constant amount of 15 μg of cNAPL protein in the presence of increasing amounts of unlabelled specific and non-specific competitor RNAs, respectively. As shown in Figure 2f–h, the presence of just a 2-fold molar excess of unlabelled specific competitor RNA almost completely abolished the formation of a complex between cNAPL and the labelled target *tscA* RNA. In contrast, the presence of the same amount of non-specific competitor RNA, namely pBIIKS+ (121 nt), had no significant effect on the formation of the complex with the labelled target RNAs. In addition, EMSA was performed with *in vitro* synthesized RNA of domains DV and DVI of the first *psaA* intron. As shown in Figure 2a, the two domains range from nt 108 to 192. EMSAs with increasing amounts of cNAPL protein (Figure 2e) and competition analysis (Figure 2i) indicate that cNAPL binds specifically to domains DV and DVI of the first *psaA* intron, albeit with lower affinity than to *tscA* RNA.

Chloroplast transcripts are rich in A and U residues. To study the sequence preferences, we compared the ability of A, C, G or U RNA homopolymers to compete for binding of cNAPL to the labelled domain DII of the *tscA* RNA and domain DV+VI of the first *psaA* intron, respectively. As depicted in Figure 3a, excess amounts of poly(U) abolished cNAPL binding, while the same amounts of poly(C) or poly(G) have no significant effect on the formation of the complex. Using poly(A) resulted in a weak reduction of the binding specificity. As depicted in Figure 3b, binding of cNAPL to domain DV and DVI of the first *psaA* intron was also competed by poly(U) and in addition by poly(A) and to a lesser extent by higher amounts of poly(G) (0.2 and 0.5 μg). The competition assays suggest that cNAPL preferentially interacts with U-rich sequences. To further investigate the specificity of cNAPL for the *tscA* RNA, total wild-type RNA was used as competitor in EMSAs. As shown in Figure 3a and b, 0.5 μg of total RNA competed the binding of cNAPL to the *tscA* and *psaA* intron RNA, respectively, suggesting that there are other transcripts bound by cNAPL. Simultaneously, we examined the binding of cNAPL to transcripts of four chloroplast genes and three nuclear genes. As given in Figure 3c, cNAPL binds to 5′-UTRs of chloroplast *psbA*, *psbD*, *rbcL* and *rps4* RNA and to a lesser extent to nuclear *Lhcb1* 5′-UTR and *Rps18*. No binding is observed to nuclear *Lhcb1* 3′-UTR and *Tba1A*. To investigate the binding specificity of cNAPL to the *tscA* RNA, unlabelled RNA of the above-mentioned chloroplast and nuclear transcripts was used in competition assays. As shown in Figure 3d, a 50-fold molar excess of *Lhcb1* 3′-UTR, *Tba1A* and a 15-fold molar excess of *Lhcb1* 5′-UTR had no significant effect on the formation of the complex, whereas a 50-fold molar excess of *psbA* 5′-UTR, *psbD* 5′-UTR and *Rps18* resulted in a slight decrease in *tscA* RNA–protein complex formation. Interestingly, similar results in competition experiments were obtained when binding of cNAPL to the labelled *tscA* RNA was competed by a 50-fold molar excess of 5′-UTRs of *rbcL* or *rps4*, or by a 2-fold molar excess of unlabelled *tscA* RNA. Thus, cNAPL shows specificity for the *tscA* RNA and to a lesser extent to other U-rich chloroplast transcripts as well. The overall comparison of all tested *in vitro* transcripts revealed that binding of cNAPL depends on the U-content and stretches of Us (data not shown).

We further tested whether the typical secondary structures of group II intron domains mediate cNAPL binding. In Figure 3e, only domains with U-rich sequences are able to shift cNAPL, indicating that the structure is not relevant for binding. For example, the U-rich domain DV+VI of the first *psaA* intron (Figure 2e) show cNAPL binding, while the corresponding domains of the heterologous intron rI1 of *Scenedesmus obliquus* do not. Furthermore, the experiment

in Figure 3f shows that cNAPL do not have DNA-binding activity.

Taken together, these results indicate that cNAPL binds specifically the first *psaA* group II intron and thus confirm data from the yeast three-hybrid analyses. Furthermore, cNAPL has affinities to U-rich chloroplast 5′-UTRs. Hence, the data imply that the RNA-binding protein cNAPL may play a more general role in the RNA metabolism of *C.reinhardtii*.

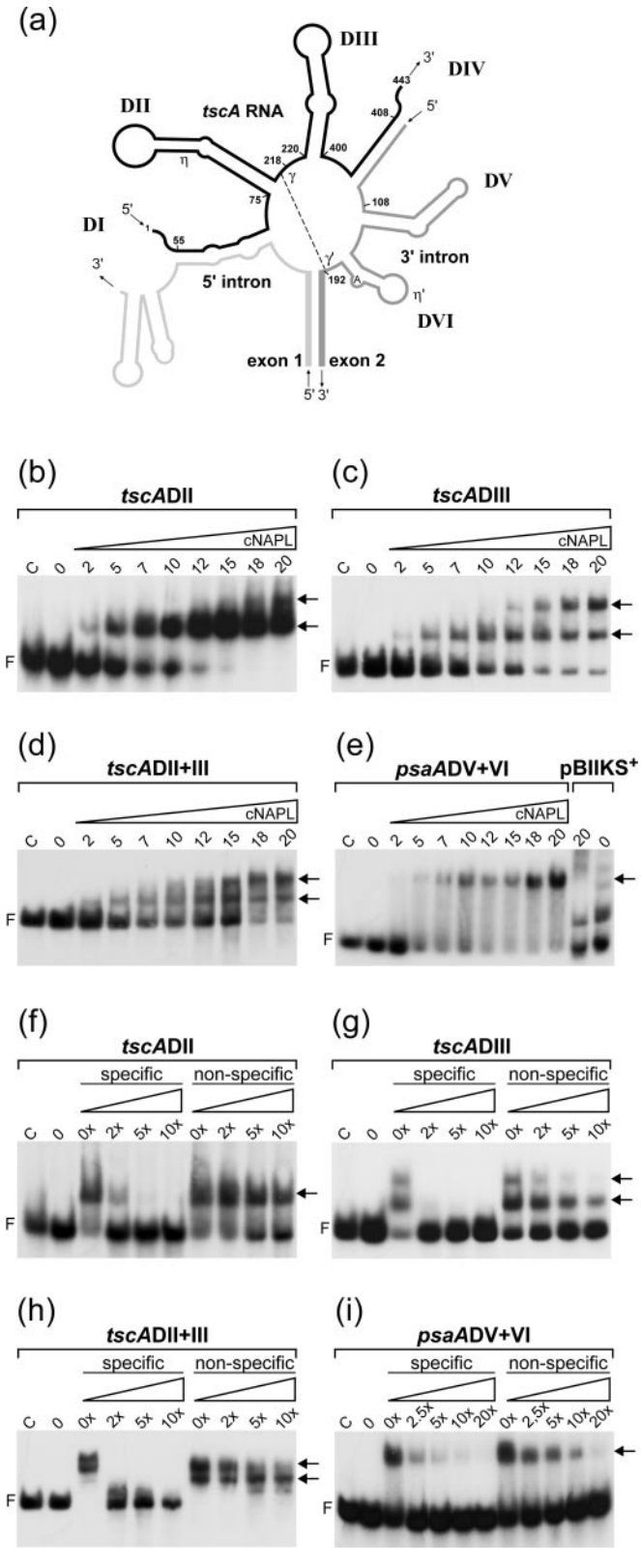

Figure 2. *In vitro* binding of cNAPL to representative intron domains of *tscA* and *psaA* RNA. (**a**) Secondary structure model of the first *psaA* intron (6). The tripartite group II intron consists of three separate RNA molecules (*tscA*, 5′ intron exon 1, 3′ intron exon 2). Arrows indicate 5′–3′ strand polarity of the three separate transcripts. Roman letters denote the six conserved domains of group II introns (DI-DVI). Black numerals correspond to nucleotides of the processed *tscA* RNA sequence and to the 3′ intron sequence of *psaA* exon 2. The γ–γ′ tertiary interaction is symbolized by a dashed line (79). The η–η′ interaction is indicated (72). The peripheral structures of domains DI and DIV could not be unambiguously determined and are not represented, modified according to (6). (**b**–**e**) RNA EMSAs of domain DII and DIII of *tscA* RNA and domain DV+VI of the first *psaA* intron with increasing amounts of cNAPL. EMSAs were performed by incubating 30 fmol internally labelled intron RNA (*tsc*ADII, *tsc*ADIII, *tsc*ADII+III, *psa*ADV+VI) with increasing amounts of affinity-purified cNAPL polypeptide (numbers indicate the amounts of protein in μg) or without protein (0). Affinity-purified CEFEF protein of *Acremonium chrysogenum* was used as a negative control (C). RNA–protein complexes were resolved from free RNA (F) by electrophoresis in 5 % native polyacrylamide gels. Arrows indicate shifted bands. (**f**–**i**) Competition assays of cNAPL using 30 fmol radioactive probes of internally labelled intron RNA (*tsc*ADII, *tsc*ADIII, *tsc*ADII+III, *psa*ADV+VI) and excess of cold specific (*tsc*ADII, *tsc*ADIII, *tsc*ADII+III, *psa*ADV+VI) and non-specific competitors (pBIIKS+). Lanes covered with triangle on top are 0-, 2-, 5- and 10-fold molar excess of competitor, respectively, in the case of *tscA* RNA and 0-, 2.5-, 5-, 10- and 20-fold molar excess of competitor, respectively, in the case of *psaA*. Free RNA, negative control and shifted bands are indicated as in (b–e).

Characterization of a chloroplast NAP-like polypeptide

The nucleotide sequence of the *cNapl* cDNA (pGAD6-19) consists of 1642 bp with a TGTTA polyadenylation site located 139 bp upstream of the poly(A) tail in the 5′-UTR. The 3′-UTR region is 588 bp in length, and a 5′ rapid amplification of cDNA ends (RACE) analysis revealed 22 bp for the 5′-UTR (data not shown). Sequence analysis identified an open reading frame of 1053 bp and 350 amino acids, encoding a polypeptide with a predicted size of 39.4 kDa. The amino acid composition of the protein revealed 20.6% acidic and 13.5% basic residues with an estimated pI of 4.33. The hydropathy profile calculated according to (38) indicates mostly hydrophilic regions, suggesting that cNAPL is a soluble protein. The neural-network based predictors for subcellular localization of proteins ChloroP (31) and LOCtree (32) revealed a putative transit sequence for import into the chloroplast at the N-terminal end of cNAPL, whereas TargetP proposes a mitochondrial localization. However, for chloroplast transit peptides of *Chlamydomonas*, it was shown that they are more similar to mitochondrial targeting peptides than to chloroplast transit peptides of higher plants (39). The N-terminal amino acid sequence of cNAPL possesses features typical of chloroplast transit peptides, such as an alanine as the second residue. The N-terminal end is an uncharged region with an abundance of arginine (13.6%), valine (13.6%) and alanine (18.2%) residues. This transit peptide could be 46 amino acids long since a possible cleavage site sequence Val-Gly-Ala is present (39). The cleavage would give rise to a predicted mature protein of 304 amino acids with a molecular mass of 34.6 kDa.

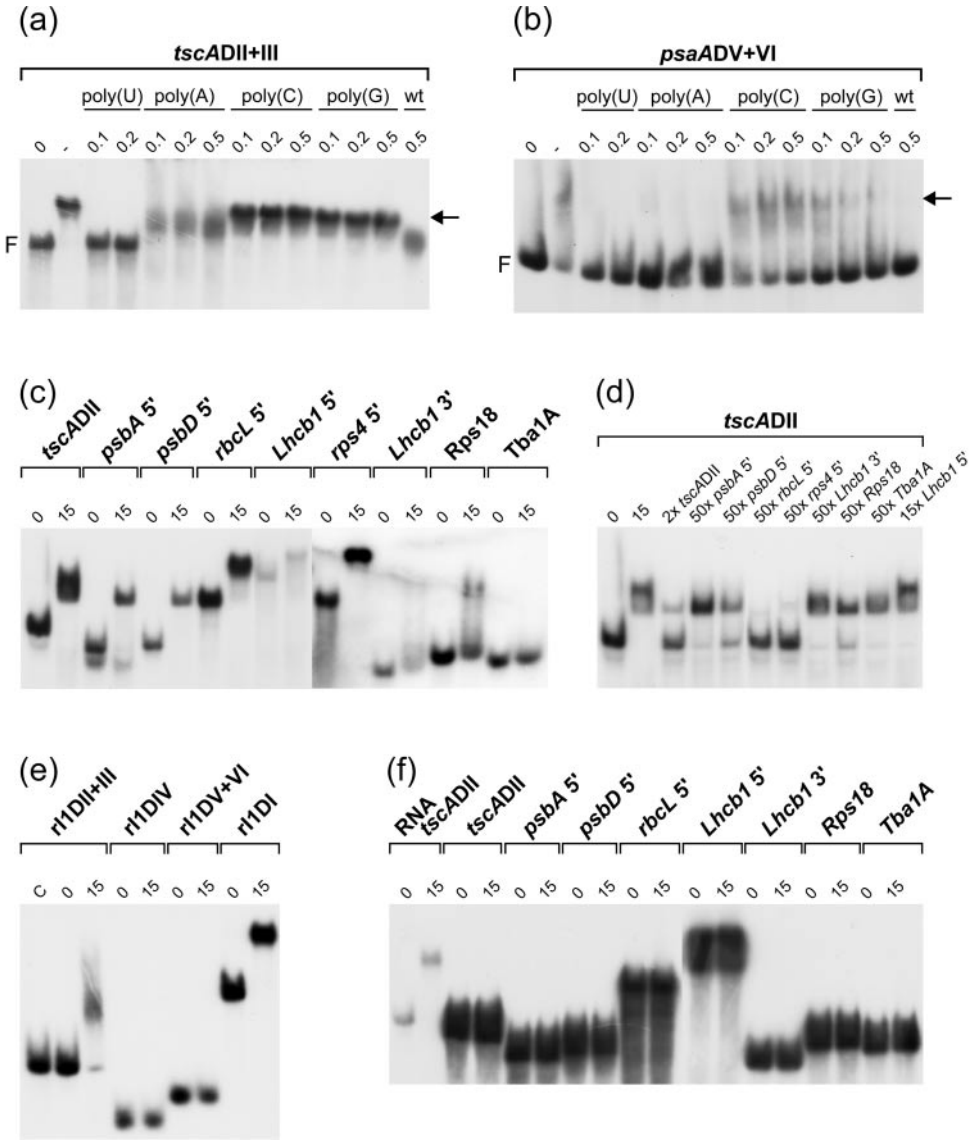

Figure 3. Binding specificity of cNAPL. (**a** and **b**) RNA EMSAs of domain DII+III of *tscA* RNA (30 fmol) and domain DV+VI of the first *psaA* intron (40 fmol) with cNAPL in the presence of no competitors (-), the unlabelled RNA homopolymer competitors poly(U), poly(A), poly(C), poly(G) or the unlabelled wt total RNA. Numbers indicate the amount of competitor in μg. (**c**) Chloroplast transcripts (30 fmol) corresponding to fragments of 5′-UTRs (5′), 3′-UTR (3′) and to regions of coding sequences were tested as indicated. (**d**) Competition assay of cNAPL using 30 fmol of labelled domain DII+III of *tscA* RNA and excess of cold specific or unspecific competitors as indicated. (**e**) RNA EMSA of domain DI (20 fmol), DII+III (40 fmol), DIV (60 fmol) and DV+VI (60 fmol) of heterologous intron rI1 of *S.obliquus* with cNAPL. (**f**) DNA EMSA of cNAPL using 30 fmol of labelled PCR fragments as indicated. DNA was incubated in the absence (0) or presence (15 μg) of recombinant cNAPL. RNA (30 fmol) serves as a control for shift conditions. RNA in all experiments was incubated in the absence (0) or presence (15 μg) of affinity-purified cNAPL. Free RNA, controls and shifted bands are denoted as in Figure 2b–e.

Comparison of the genomic and cDNA sequences of *cNapl* using the *C.reinhardtii* JGI database (DOE Joint Genome Institute) reveals the presence of 9 large introns and 10 exons (Figure 4). As mentioned above, database searches detected significant sequence similarity with NAPs. Using the *C.reinhardtii* JGI database, we could also identify a *Chlamydomonas* homologue of a NAP that does not possess a chloroplast targeting signal. Furthermore, another algal NAP homologue was detected in the marine centric diatom *T.pseudonana* using the *T.pseudonana* JGI database (DOE Joint Genome Institute). Figure 5 shows a multiple alignment of cNAPL with NAP polypeptides of algae and higher plants that displayed the highest homology using the BLASTP algorithm (40). Regions of sequence homology between these proteins were found over the entire length of the amino acid sequence, including several blocks of sequence identity. A total of 50 amino acids are identically positioned within all 8 sequences. Overall, there is ∼29% identity between cNAPL and the other NAP proteins, indicating that cNAPL is a member of the NAP family. With ∼46% identical amino acids, the *Chlamydomonas* NAP polypeptide shows a higher degree of homology to plant NAPs than cNAPL. However, *Chlamydomonas* cNAPL and *Thalassiosira* NAP share common features atypical for NAPs. They possess, for example, an N-terminal extension that is not found in plant NAPs. In cNAPL, this extension was

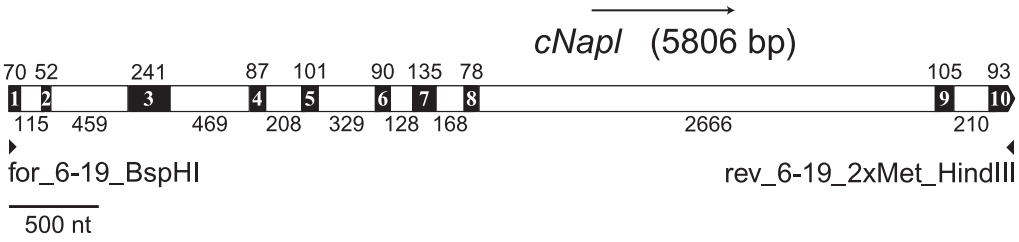

Figure 4. Physical map of the *cNapl* gene. Exons (1–10) and introns are represented by closed and open boxes, respectively, with their corresponding size in base pairs. The arrow indicates transcriptional direction. Oligoucleotides used for RT–PCR analysis are given as arrowheads.

shown *in silico* to correspond to a chloroplast targeting signal, whereas an unambiguously prediction is currently impossible for the extension in the sequence of *Thalassiosira* NAP, because of its complex plastids and multipartite plastid targeting signals (41). Interestingly, no chloroplast transit signal could be identified in all currently available NAP sequences of higher plants.

When we compared the protein sequences of NAPs from different sources, several evolutionarily conserved domains, such as stretches of acidic residues, as well as nuclear transport signal sequences, were detected (12,42). Recently, Dong and co-worker (42) identified a 16 amino acid sequence in the N-terminal regions of NAP-like sequences from rice and tobacco that is similar to the nuclear export signal (NES) of the yeast NAP1 protein. From the comparison in Figure 5, it is evident that this 16 amino acid stretch is absent in the *C.reinhardtii* cNAPL sequence. However, there are stretches of amino acid residues in the central and C-terminal part of the sequences of Arabidopsis, maize, pea, rice, soybean and *Chlamydomonas* NAP that have characteristics of nuclear localization signals (43). The NLS sequences are similar to PKKKKRKV, the NLS found in SV40 T antigen (44). In contrast, cNAPL and *Thalassiosira* NAP do not possess any NLS sequences. The alignment in Figure 5 reveals two clusters of highly acidic amino acids at the C-terminus of higher plants and *Chlamydomonas* NAP; both these clusters can also be found in similar positions in the amino acid sequences of yeast NAP1 and HeLa hNRP. Interestingly, however, neither of these clusters is present in cNAPL and *Thalassiosira* NAP.

Furthermore, Arabidopsis, maize, pea, rice, soybean and *Chlamydomonas* NAP contain the CKQQ sequence at their C-terminal ends, a motif suggested as a potential prenylation signal. This signal is supposed to promote membrane interactions and appears to play a major role in several protein–protein interactions. The CKQQ sequence was also identified in NAP polypeptides of rice and tobacco (45). The prenylation signal is missing in cNAPL and *Thalassiosira* NAP.

To analyse the expression of *cNapl* in the wild-type strain, RT–PCR analysis was carried out with specific primers designed to amplify *cNapl* from the start (ATG) to the stop (TAG) codon (Figure 4). As shown in Figure 6a, a single band of around the expected size of 1 kb was obtained. This result indicates the specific expression of *cNapl* in *C.reinhardtii* wild-type strain. To examine *cNapl* expression, total RNA was prepared from wild-type strains CC406 cw15 and arg$^-$cw15 and photosystem I mutants T11-1 59 and T11-3 39 (S. Glanz, unpublished data). Northern blot analysis was carried out using the ^{32}P-labelled 1.7 kb fragment of the *cNapl* cDNA as probe. As shown in Figure 6b, the probe detected a specific transcript of ~1.6 kb, which is consistent with the expected size of the *cNapl* transcript. An equal load check using *Rps18* (46) revealed no difference between transcript accumulation in wild-type versus photosystem I mutant strains. Thus, pleiotropic phenotypes that are often associated with photosynthetic defects do not affect transcriptional expression of the *cNapl* gene.

Cellular localization of the cNAPL

The N-terminal 46 amino acids of cNAPL were predicted to be a signal sequence because of certain structural features. To determine the cellular localization of the cNAPL polypeptide, we performed LSCFM. A map of the constructs is shown in Figure 7a. As can be seen in Figure 7b, non-transformed arg$^-$cw15 cells (wt) do not show any fluorescence when green fluorescent protein (GFP) was excited. The complete N-terminal signal sequence (46 amino acids) and the sequence for the first 38 amino acids of the mature cNAPL were fused to a synthetic GFP sequence (47) (N84-cGFP) in frame, keeping the putative N-terminal transit peptide functional. Previous work on diatom protein localization established that the complete pre-sequence of plastid preproteins is sufficient to drive the import of GFP into plastids (41). A similar strategy was successfully applied recently for localization of the phage-type organellar RNA polymerases of *Arabidopsis thaliana* and *Chenopodium album* (48). To determine the cellular localization of cNAPL in the algal cell, the *cNaplN84-cgfp* fusion gene was expressed in *C.reinhardtii* cells under the control of the *HSP70A*/*RBCS2* promoter (49). As shown in Figure 7b, LSCFM confirmed a plastidic localization of cNAPL. The fluorescent images reveal a co-localization of the chimeric N84-cGFP polypeptide (green fluorescence) with the chloroplast (red autofluorescence of plastids). Right panels show the merged images of cGFP fluorescence and chlorophyll autofluorescence of the same cell. To rule out the possibility that use of the non-native promoter *HSP70A*/*RBCS2* leads to mis-targeting of cNAPL, a gene for a fusion protein with a 31 amino acid deletion in the chloroplast targeting sequence of cNAPLN84 was constructed (N84Δ8-39-cGFP). As shown in Figure 7b, this protein clearly localizes in the cytoplasm and to spot-like structures outside of the autofluorescing chloroplast. As further control for nuclear localization, we used cGFP in fusion with the Ble polypeptide (Ble-cGFP). The bacterial *ble* gene has been successfully expressed in *C.reinhardtii* by fusing the coding sequence to the 5′ and 3′ non-coding sequence of the *RBCS2* gene, containing additional intron sequences (50). For the Ble polypeptide, it has

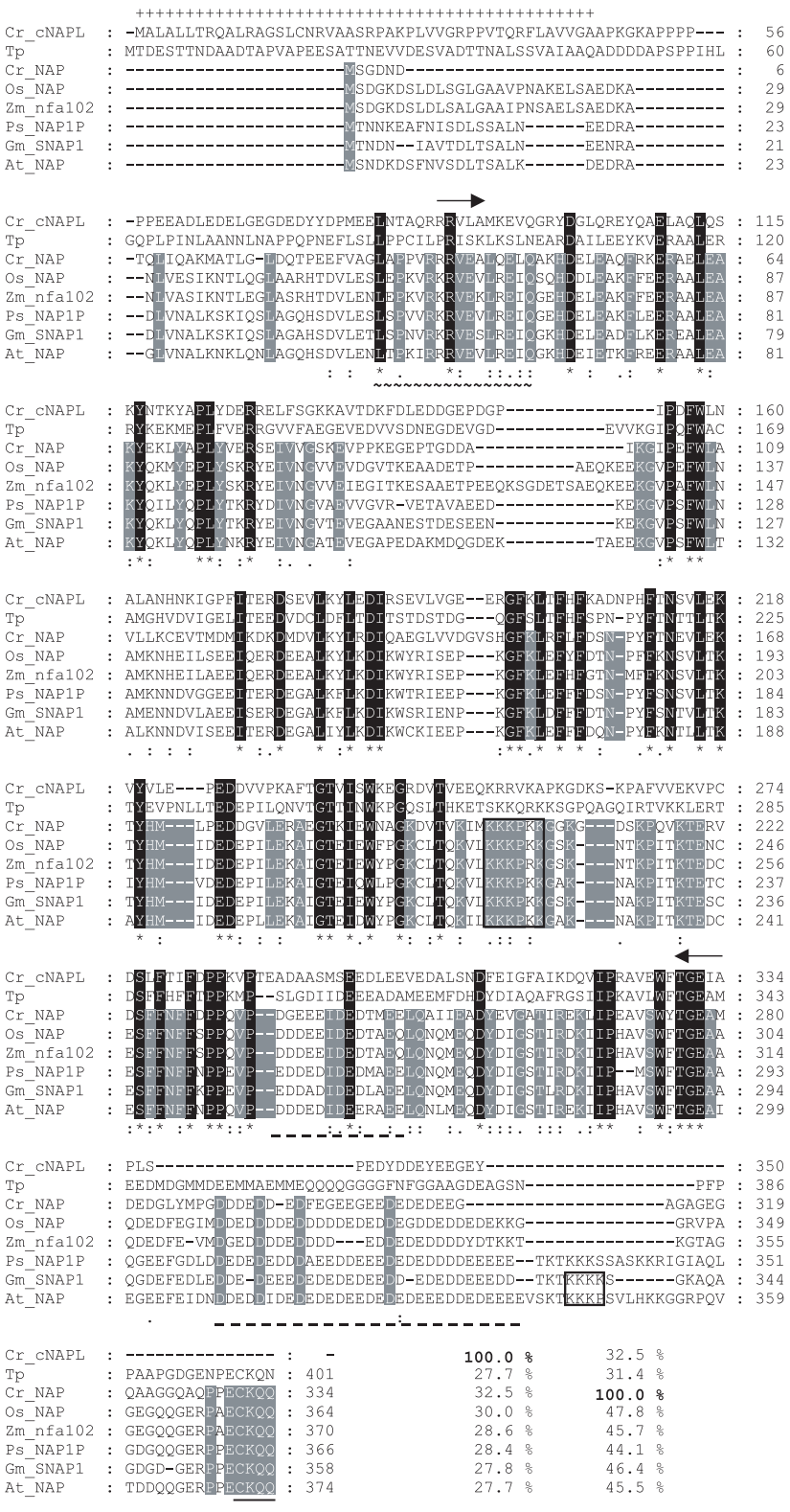

Figure 5. Amino acid sequence comparison of *C.reinhardtii* cNAPL, NAP and NAPs of higher plants and *T.pseudonana*. Black shading denotes distinct similarities of the *tscA*-binding protein cNAPL to NAPs. Grey shading shows the higher degree of homology of the NAP protein from *C.reinhardtii* to NAP proteins of higher plants in contrast to cNAPL. Arrows delimit the NAP domain. A putative prenylation motif is underlined, putative nuclear localization signals are framed and a putative NES is underlined with '∼'. A putative chloroplast targeting signal sequence is represented by '+' on top of the alignment. The two acidic amino acid regions are underlined with dashed lines. An asterisk indicates identical residues. Similar residues are indicated by one or two dots, depending on the degree of similarity of the relevant amino acids. The following protein sequences were used for comparison: At, *A.thaliana* AAL47354; Cr, *C.reinhardtii* C_1660026 (JGI); Gm, *Glycine max* AAA88792; Os, *Oryza sativa* XP_475795; Ps, *Pisum sativum* T06807; Tp, *T.pseudonana* 115707 (JGI); Zm, *Zea mays* AY100480.

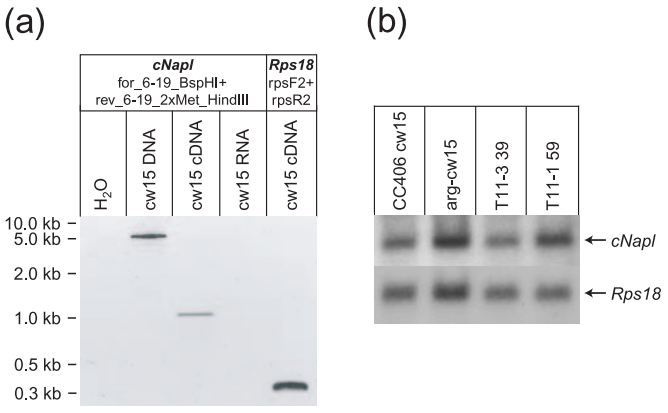

Figure 6. Analyses of *cNapl* transcript. (**a**) RT–PCR analysis of *cNapl* transcripts in wild-type arg$^-$cw15. RT–PCR was carried out by using either DNase treated poly(A) RNA as a control (cw15 RNA), cDNA (cw15 cDNA) or genomic DNA (cw15 DNA) as template. Oligonucleotides used for RT–PCR are for_6-19_BspHI and rev_6-19_2xMet_HindIII. Their position is shown in Figure 4. Oligonucleotides rpsF2 and rpsR2 specifically amplify a fragment of the *Rps18* transcript and serve as a control for cDNA preparation. Abbreviation: *Rps18*, cytoplasmic ribosomal protein S18 of *C.reinhardtii* (46). (**b**) RNA blot analysis of *cNapl* transcripts in wild-type and non-photosynthetic mutant strains. Total RNA (20 µg) of wild-type strains CC406 cw15 and arg$^-$cw15 and photosystem I mutants T11-3 39 and T11-1 59 (S. Glanz, unpublished data) were separated on a 1% agarose/formaldehyde gel, transferred onto a nylon membrane and hybridised separately with a *cNapl*- and *Rps18*-specific probe as a loading control.

been shown previously that the bacterial protein is localized in the nucleus (51). This localization was also confirmed by our experiments as demonstrated in Figure 7b and there was no congruence of cGFP fluorescence with the position of the plastid as observed by fluorescence microscopy. In addition, we used the C-terminal cGFP-fused Rps18 polypeptide (Rps18-cGFP) as a marker for cytoplasmic localization. The corresponding cGFP fluorescence signals appeared to accumulate in spots within the cytoplasm. To the best of our knowledge, this is the first demonstration of a cytoplasmic localization of a *C.reinhardtii* protein using GFP. In conclusion, our results from microscopic investigations clearly demonstrate that the N-terminal extension of cNAPL functions as a chloroplast targeting signal that is able to direct cNAPL to the organelle.

Phylogenetic analyses

The members of the NAP family are considered as histone chaperones, which are conserved from yeast to man. To examine the evolutionary relationships between cNAPL and the above-mentioned homologues, we performed a phylogenetic analysis using the neighbour-joining method. A tree was constructed from multiple alignments of the NAP domains of cNAPL and of NAP polypeptides from various eukaryotes. In addition to the NAP homologue in the diatom *T.pseudonana*, a partial sequence of a predicted protein with homology to NAPs was found in the EST database of the red alga *Porphyra yezoenis* (Kazusa DNA Research Institute) (25,26). Similarly, we detected another partial EST homologue (GenBank) from the green alga *Dunaliella salina*. So far, genomic sequences of these algae are not available yet. For prokaryotes, however, we were unable to detect any NAP or cNAPL homologues in the publicly available genome sequences of bacteria (GenBank) and cyanobacteria (CyanoBase).

The unrooted tree in Figure 8 shows four main clusters: animals, algae, fungi and higher plants. Trees generated with maximum parsimony and Bayesian analysis (52) had an overall similar topography (data not shown). From Figure 8, it becomes evident that the *Chlamydomonas* NAP polypeptide is related to the group of higher plant NAP proteins, which all lack a chloroplast targeting sequence, whereas cNAPL clusters together with the algal NAP homologues. For the NAP domain, the amino acid sequence of *Chlamydomonas* NAP shows ~54% identity to higher plant NAPs, whereas the NAP domain of cNAPL exhibits ~35% identity to higher plant NAPs. The corresponding values for the algal NAPs are ~35% for *Thalassiosira* NAP, ~26% for *Dunaliella* NAP and ~15% for *Porphyra* NAP. Overall, we propose that cNAPL with its chloroplast transit peptide represents a new type of NAP-like polypeptide in photoautotrophic organisms.

DISCUSSION

cNAPL is an *tscA*-specific RNA-binding protein

The splicing chemistry and RNA structures of self-splicing group II introns and nuclear pre-mRNA introns are strikingly similar. Thereby, suggesting some evolutionary relationship between group II introns and the nuclear spliceosomal intron (53). Moreover, there is compelling evidence to suggest that chloroplast spliceosomes catalyse intron splicing using a similar mechanism to that reported for nuclear spliceosomes. This suggestion is the result of data obtained form a combination of different experimental approaches that have identified components of a putative chloroplast spliceosome in *C.reinhardtii* (7–11) and higher plants (54). Additional support comes from studies with tobacco chloroplasts. Nakamura and co-workers (55) identified chloroplast ribonucleoproteins (cpRNPs) that are associated *in vivo* with various species of chloroplast mRNAs and intron-containing precursor tRNAs. They suggested that these stromal RNA–protein complexes promote, for example, RNA splicing, processing or editing. Previously, the yeast three-hybrid system was used to demonstrate specific binding of proteins to organellar intron RNA (8,56). In this work, we used the yeast three-hybrid system successfully to screen a *C.reinhardtii* cDNA library and we were able to detect the novel organellar protein cNAPL that specifically binds to the first *psaA* group II intron. The binding of cNAPL to domains DII+III and DV+VI of the first *psaA* intron was further shown in electromobility shift assays.

In several other studies using defined intron domains, separated molecules were found to fold into a structure corresponding to their conformation in the complete active intronic RNA (57,58). Recent studies have demonstrated that fragmented intron molecules are appropriate targets for gel retardation assays (22). In EMSA, we observed multiple shifted bands after incubating domains DII+III of *tscA* RNA with increasing amounts of cNAPL. As was shown for example for the mutated SxlN1 protein of *Drosophila*, multiple shifted bands probably correspond to the sequential filling of binding sites when protein concentration is

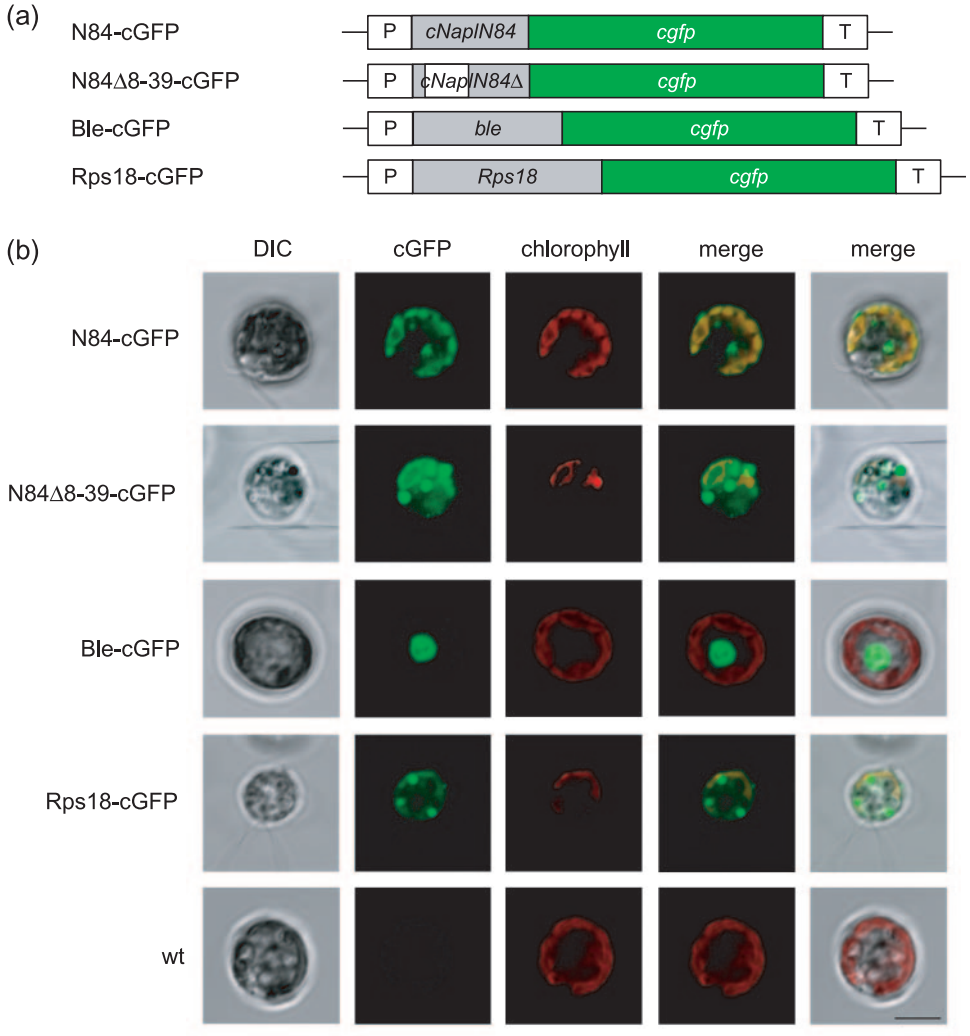

Figure 7. Confocal fluorescence microscopy to demonstrate chloroplast localization of the cNAPL polypeptide. (**a**) Schematic representation of fusion protein constructs used for cGFP-assays. (**b**) Laser scanning confocal microscopy of *C.reinhardtii* arg$^-$cw15 transformants. Transformants were assayed by differential interference contrast microscopy (DIC) or by confocal fluorescence microscopy. DIC, cGFP fluorescence (green) and chlorophyll autofluorescence (red) were merged as indicated. The DIC and cGFP images are representatives of five independent experiments. Scale bar represents 5 µm. Abbreviations: *ble*, phleomycin resistance gene of *Streptoalloteichus hindustanus* (50); *cgfp*, synthetic GFP (47); N84, the N-terminal 84 amino acids of cNAPL; N84Δ8-39, the N-terminal 84 amino acids of cNAPL lacking 31 amino acids in the chloroplast targeting sequence; P, *HSP70A/RBCS2* promoter (49); *Rps18*, cytoplasmic ribosomal protein S18 of *C.reinhardtii*; T, 3′-UTR of *Lhcb1* or of *RBCS2* gene.

increased (59). Hence, the two shifted bands observed in Figure 2b–d shows non-cooperative binding of cNAPL to domains DII and DIII, thus indicating at least two binding sites. However, we were not able to detect a known RNA-binding motif [http://www.els.net/, (60)] in cNAPL, such as the RNP motif, the K homology motif (KH) or the RGG (Arg-Gly-Gly) box. Further binding studies with truncated cNAPL constructs could help to determine the exact RNA-binding domain.

cNAPL is phylogenetically related to nuclear localized NAPs and most probably has not been acquired from a cyanobacterial endosymbiont

cNAPL is a member of the multifunctional family of NAPs that are involved in processes which normally take place in the nucleus or the cytoplasm of various eukaryotes (12,42,61). Besides their distinct functions in mitotic events and cytokinesis, NAPs act as histone chaperones, binding to all core histones with a preference for H2A and H2B and working as a deposition factor by transferring NAP1-bound histones to double-stranded DNA. In addition to its nucleosome assembly and histone-binding activity, NAP1 as well as plant NAP1-like proteins may be involved in regulating gene expression and therefore cellular differentiation (45,62,63). NAP1 has also been shown to facilitate transcription factor binding by disruption of the histone octamer through the binding of H2A and H2B (13,64). NAP1 also shuttles histones H2A and H2B as well as a complex of H2A, H2B, H3 and H4 between the template DNA and nascent RNA during transcription (65). To date, however, it is still unknown whether NAP1 interacts directly also with RNA.

Phylogenetic analysis indicates that cNAPL is related to nuclear localized NAP polypeptides. The clustering of

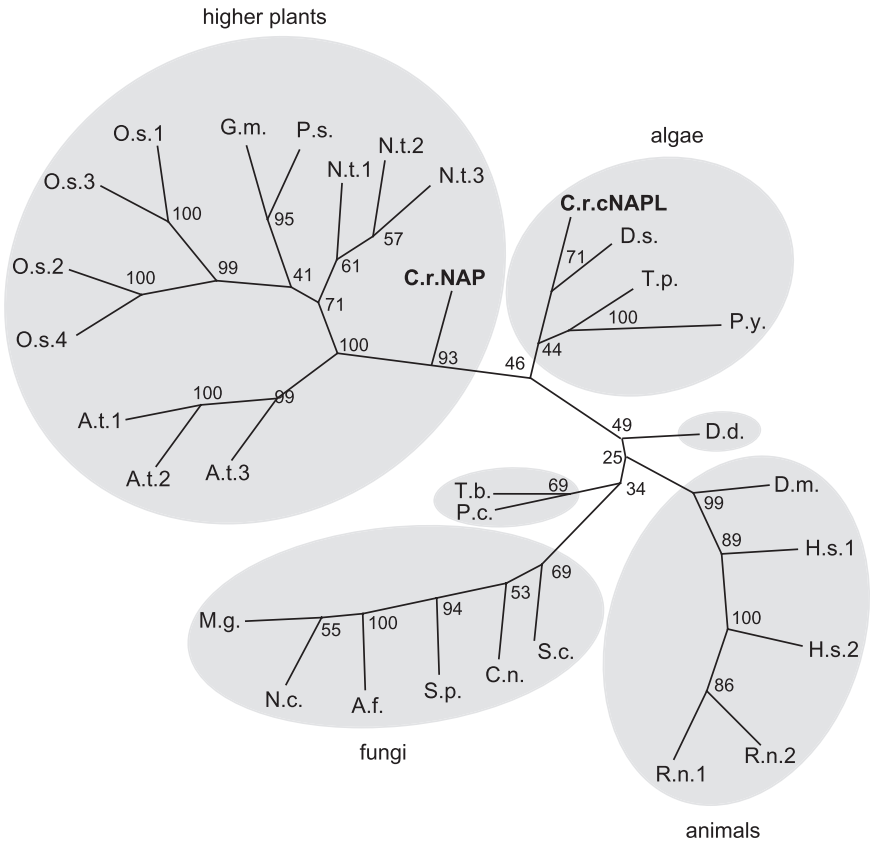

Figure 8. Phylogenetic tree of 31 NAP/NAP-like proteins. The tree was generated using PHYLIP (29), as indicated in Material and Methods. Numbers at branches indicate bootstrap support (2000 bootstrap replicates) in percent. Abbreviations: A.f., *Aspergillus fumigatus* XP_754077; A.t.1, *A.thaliana* AAL47354; A.t.2, *A.thaliana* BAA97025; A.t.3, *A.thaliana* NP_194341; C.n., *Cryptococus neoformans* AAW43807; D.d., *Dictyostelium discoideum* EAL71995; D.m., *Drosophila melanogaster* AAB07898; D.s., *D.salina* BM448458; G.m., *G.max* AAA88792; H.s.1, *Homo sapiens* BAA08904; H.s.2, *H.sapiens* NP_004528; M.g., *Magnaporthe grisea* AAX076; N.c., *Neurospora crassa* CAD70974; N.t.1, *Nicotiana tabacum* CAD27460; N.t.2, *N.tabacum* CAD27461; N.t.3, *N.tabacum* CAD27462; O.s.1, *O.sativa* AAV88624; O.s.2, *O.sativa* CAD27459; O.s.3, *O.sativa* P0456F08; O.s.4, *O.sativa* XP_475795; P.c., *Paramecium caudatum* BAB79235; P.s., *P.sativum* T06807; P.y., *P.yezoensis* AU191247; R.n.1, *Rattus norvegicus* AAC67388; R.n.2, *R.norvegicus* NP_446013; S.c., *S.cerevisiae* P25293; S.p., *Schizzosaccharomyces pombe* T41330; T.b., *Trypanosoma brucei* XP_826968; T.p., *T.pseudonana* 115707 (JGI).

cNAPL with algal NAPs and not with NAPs of higher plants could be explained by its chloroplast targeting signal. In the case of *Thalassiosira* NAP, TargetP predicted a 69 amino acid signal sequence, which could not be assigned clearly to any organelle. Diatoms, such as *T.pseudonana*, possess so-called 'complex plastids' delineated by four distinct membranes. For a nuclear-encoded plastid polypeptide to be directed to the plastids, multiple targeting signals are necessary. Thus, determining a protein's localization in diatoms using today's prediction programs is difficult.

It is now generally accepted that double-membrane-bound chloroplasts are the result of an endosymbiotic event involving a cyanobacterial-like organism early on in evolution. During evolution, the cyanobacterial endosymbiont has lost its autonomy, and this has been accompanied by significant changes of the chloroplast proteome (14,66). A key factor in this process of adoption was the loss of genetic material resulting from gene transfer to the cell nucleus. The vast majority of proteins in present day chloroplasts are encoded by the host nucleus and require N-terminal presequences that target them back to the chloroplast. Therefore, in order to establish a functional organelle–nucleus interaction, chloroplasts have had to import a set of new proteins to adapt their metabolism to the new conditions. Interestingly, a proportion of these proteins do not seem to have been acquired from cyanobacteria (14). Phylogenetic analysis and database research revealed that no NAP homologues could be identified in prokaryotic organisms. We therefore suggest that cNAPL does not originate from an endosymbiotic gene transfer event and that the chloroplast of *C.reinhardtii* has gained cNAPL as a novel nuclear-encoded factor. The phylogenetic analysis suggests that *cNapl* originate from a nuclear *Nap* gene, and is probably the result of a gene duplication event that occurred after separation of higher plants from green algae. We thus propose that cNAPL with its chloroplast transit peptide represents a new type of NAP-like polypeptide.

Is cNAPL a multifunctional protein?

Many group II introns of plants have lost their self-splicing activity during evolution and thus, recruited host-encoded proteins as splicing factors (1,67). The participation of nuclear-encoded proteins in chloroplast splicing could provide a means to control the biogenesis of the photosynthetic apparatus. It has been suggested that these recruited

host-encoded proteins probably had other functions in the cell. Many of these proteins show some degree of homology to proteins involved in RNA metabolism. For example, Raa2 of *C.reinhardtii* shows homology to pseudouridine synthetase; however, such enzymatic activity is not required for *trans*-splicing of *psaA* RNA (7). Another example is Mss116p of *Saccharomyces cerevisiae*, which influences the splicing of all nine group I and all four group II introns in mitochondria (68). Mss116p is related to DEAD-box proteins with RNA chaperone function and may facilitate splicing by resolving misfolded introns. However, it is generally discussed that proteins involved in splicing and interacting with intron RNAs seem to support RNA folding or to stabilize the active conformation, whereas the catalytic potential is clearly located in the RNA itself (69). Furthermore, Mss116p is a bifunctional protein that influences, in addition to splicing of group I and group II introns, some RNA end-processing reactions and translation of a subset of mitochondrial mRNAs. Indeed, for several nuclear-encoded host proteins, it is known that they have additional cellular functions and seem to have been recruited for splicing during evolution (1). Therefore, it has been questioned whether cNAPL is also a multifunctional protein. The analysed interactions of cNAPL with *tscA* RNA and domains DV and DVI of the first *psaA* intron implicate a role for cNAPL as a potential *tscA* RNA processing factor and also as a splicing factor of the first *psaA* intron. An example for such a bifunctional protein comes from recent work where Rat1 is required for both processing of *tscA* RNA as well as for splicing of the first *psaA* intron (8). Another factor involved in *tscA* RNA processing and splicing of both introns is Raa1-L137H (9), allelic to HN31 (70).

cNAPL binds specifically to domains DII and DVI of the first *psaA* intron. Domain DII is a phylogenetically less-conserved region and was assigned as functionally unimportant because of its high degree of sequence and structure variations (71). However, Chanfreau and Jacquier (72) could show that the so-called tertiary η–η' interaction (see also Figure 2a) between DII and DVI of group IIA introns is responsible for a conformational change between the two catalytic transesterification steps. Interestingly, only a few of the known group IIB introns can potentially form this interaction. It is discussed that the majority of the group IIB introns interact with protein factors to change the conformation between the first and the second step (69). The fact that cNAPL interacts with both domains, DII and DVI, implies a role for cNAPL in this conformational change.

Our experiments further demonstrated the binding of cNAPL to domain DV of the first *psaA* intron. Domain DV is the phylogenetically most conserved sequence and essential for the catalytic splicing reaction. DV shows numerous important tertiary interactions and contains important conserved sequence motifs (69). Overall, we assume that cNAPL presumably has a function in the stabilization and correct folding of the catalytic core of the first *psaA* intron by interacting with *tscA* RNA and domains DV and DVI.

In addition, our analyses showed that the binding of cNAPL to *tscA* RNA was competed efficiently by poly(U) in contrast to other RNA homopolymers. Moreover, cNAPL binds the 5′-UTR of U-rich chloroplast transcripts *psbA*, *psbD*, *rbcL* and *rps4* considerably better than the nuclear transcripts *Lhcb1* 5′-UTR, *Lhcb1* 3′-UTR, *Rps18* and *Tba1A*. It is proven that the 5′-UTRs of chloroplast transcripts contain the *cis*-acting determinants for the stabilization of plastid transcripts and are essential for initiation of chloroplast translation by binding *trans*-acting factors (mostly encoded by nuclear genes) and by interaction with the ribosomal preinitiation complex (73). For example, the stability of the *psbD* mRNA depends on a nucleus-encoded TPR protein, which is part of a high-molecular weight complex mediating its function via the *psbD* 5′-UTR (74). For the *psbA* mRNA it is assumed that a multiprotein complex specifically regulates translation of this mRNA through interaction with a U-rich sequence in the 5′-UTR (75). Furthermore, there are several proteins in *Chlamydomonas* binding to 5′-UTRs of chloroplast mRNAs, e.g. to 5′-UTRs of *atpB*, *rbcL*, *rps7* and *rps12*, that so far have not been characterized and are discussed to be required for translational regulation (76,77). However, the cpRNP complexes seem to be not only involved in regulation of translation, but also in distinct steps of post-transcriptional gene expression, e.g. by facilitating proper RNA folding for intron-containing RNA (78). Recently it was demonstrated that there are two RNP complexes in *C.reinhardtii* that are involved in *trans*-splicing of *psaA*, one in the soluble fraction of the chloroplast and the other one associated with chloroplast membranes, which could provide a link to translation (11).

In conclusion, we have isolated a novel type of group II intron RNA-binding protein using the yeast three-hybrid system. Phylogenetic analysis indicates that the chloroplast located cNAPL protein is encoded by a nuclear gene that probably has evolved from duplication of an ancestral *Nap* gene. Moreover, it is proposed that cpRNP complexes are global mediators of chloroplast RNA metabolism, which connect transcription and translation in the chloroplast (78). Because of cNAPL's binding to U-rich sequences of chloroplast 5′-UTRs, a more general role for this RNA-binding protein is conceivable, e.g. in translation initiation of several chloroplast transcripts. Thus, cNAPL might represent a component of the cpRNP complex.

ACKNOWLEDGEMENTS

We wish to express our thanks to Profs. Peter Hegemann (Berlin), Christoph Beck (Freiburg) and Dr Ulrich Putz (Hamburg) for the gift of plasmids and Drs Birgit Hoff and Stephan Pollmann for advice in LSCFM and Patricia Cebula and Franziska Thorwirth for help with some experiments. This work was financially supported by the Deutsche Forschungsgemeinschaft (SFB480 B3). Funding to pay the Open Access publication charges for this article was provided by DFG (Bonn, Bad Godesberg, Germany).

Conflict of interest statement. None declared.

REFERENCES

1. Lambowitz,A.M. and Zimmerly,S. (2004) Mobile group II introns. *Annu. Rev. Genet.*, **38**, 1–35.
2. Barkan,A. and Goldschmidt-Clermont,M. (2000) Participation of nuclear genes in chloroplast gene expression. *Biochimie*, **82**, 559–572.

3. Nickelsen,J. (2003) Chloroplast RNA-binding proteins. *Curr. Genet.*, **43**, 392–399.
4. Kück,U., Choquet,Y., Schneider,M., Dron,M. and Bennoun,P. (1987) Structural and transcriptional analysis of two homologous genes for the P700 chlorophyll α-apoproteins in *Chlamydomonas reinhardtii*: evidence for *in vivo trans*-splicing. *EMBO J.*, **6**, 2185–2195.
5. Choquet,Y., Goldschmidt-Clermont,M., Girard-Bascou,J., Kück,U., Bennoun,P. and Rochaix,J.D. (1988) Mutant phenotypes support a *trans*-splicing mechanism for the expression of the tripartite *psaA* gene in the *C. reinhardtii* chloroplast. *Cell*, **52**, 903–913.
6. Goldschmidt-Clermont,M., Choquet,Y., Girard-Bascou,J., Michel,F., Schirmer-Rahire,M. and Rochaix,J.D. (1991) A small chloroplast RNA may be required for *trans*-splicing in *Chlamydomonas reinhardtii*. *Cell*, **65**, 135–143.
7. Perron,K., Goldschmidt-Clermont,M. and Rochaix,J.D. (1999) A factor related to pseudouridine synthases is required for chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. *EMBO J.*, **18**, 6481–6490.
8. Balczun,C., Bunse,A., Hahn,D., Bennoun,P., Nickelsen,J. and Kück,U. (2005) Two adjacent nuclear genes are required for functional complementation of a chloroplast *trans*-splicing mutant from *Chlamydomonas reinhardtii*. *Plant J.*, **43**, 636–648.
9. Merendino,L., Perron,K., Rahire,M., Howald,I., Rochaix,J.D. and Goldschmidt-Clermont,M. (2006) A novel multifunctional factor involved in *trans*-splicing of chloroplast introns in *Chlamydomonas*. *Nucleic Acids Res.*, **34**, 262–274.
10. Rivier,C., Goldschmidt-Clermont,M. and Rochaix,J.D. (2001) Identification of an RNA–protein complex involved in chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. *EMBO J.*, **20**, 1765–1773.
11. Perron,K., Goldschmidt-Clermont,M. and Rochaix,J.D. (2004) A multiprotein complex involved in chloroplast group II intron splicing. *RNA*, **10**, 704–711.
12. Krude,T. and Keller,C. (2001) Chromatin assembly during S phase: contributions from histone deposition, DNA replication and the cell division cycle. *Cell. Mol. Life Sci.*, **58**, 665–672.
13. Ito,T., Ikehara,T., Nakagawa,T., Kraus,W.L. and Muramatsu,M. (2000) p300-Mediated acetylation facilitates the transfer of histone H2A-H2B dimers from nucleosomes to a histone chaperone. *Gene Dev.*, **14**, 1899–1907.
14. Timmis,J.N., Ayliffe,M.A., Huang,C.Y. and Martin,W. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Rev. Genet.*, **5**, 123–U116.
15. Harris,E.H. (1989) *The Chlamydomonas Sourcebook: A Comprehensive Guide to Biology and Laboratory Use*. Academic Press, San Diego, CA.
16. Kindle,K.L. (1990) High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA*, **87**, 1228–1232.
17. Zorin,B., Hegemann,P. and Sizova,I. (2005) Nuclear gene targeting by using single-stranded DNA avoids illegitimate DNA integration in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*, **4**, 1264–1272.
18. Sambrook,J. and Russel,D.W. (2001) *Molecular Cloning - A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
19. Jerpseth,B., Greener,A., Short,J.M., Viola,J. and Kretz,P.L. (1992) XL1-Blue MRF' *E. coli* cells: McrA-, McrCB-, McrF-, Mrr-, HsdR- derivative of XL1-Blue cells. *Strateg. Mol. Biol.*, **5**, 81–83.
20. Balczun,C., Bunse,A., Nowrousian,M., Korbel,A., Glanz,S. and Kück,U. (2005) DNA macroarray and real-time PCR analysis of two nuclear photosystem I mutants from *Chlamydomonas reinhardtii* reveal downregulation of *Lhcb* genes but different regulation of *Lhca* genes. *Biochim. Biophys. Acta*, **1732**, 62–68.
21. Putz,U., Skehel,P. and Kuhl,D. (1996) A tri-hybrid system for the analysis and detection of RNA–protein interactions. *Nucleic Acids Res.*, **24**, 4838–4840.
22. Bunse,A.A., Nickelsen,J. and Kück,U. (2001) Intron-specific RNA binding proteins in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta*, **1519**, 46–54.
23. Balczun,C., Bunse,A., Schwarz,C., Piotrowski,M. and Kück,U. (2006) Chloroplast heat shock protein Cpn60 from *Chlamydomonas reinhadrtii* exhibits a novel function as a group II intron-specific RNA-binding protein. *FEBS lett.*, **580**, 4527–4532.
24. Kucho,K., Yoshioka,S., Taniguchi,F., Ohyama,K. and Fukuzawa,H. (2003) *Cis*-acting elements and DNA-binding proteins involved in CO_2-responsive transcriptional activation of *Cah1* encoding a periplasmic carbonic anhydrase *Chlamydomonas reinhardtii*. *Plant Physiol.*, **133**, 783–793.
25. Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S. (2003) Comparison of RNA expression profiles between the two generations of *Porphyra yezoensis* (Rhodophyta), based on expressed sequence tag frequency analysis. *J. Phycol.*, **39**, 923–930.
26. Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and Tabata,S. (2000) Generation of 10 154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis*. *DNA Res.*, **7**, 223–227.
27. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) ClustalW—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
28. Nicholas,K.B., Nicholas,H.B.J. and Deerfield,D.W. (1997) GeneDoc: analysis and visualization of genetic variation. *EMBNET News*, **4**, 1–4.
29. Felsenstein,J. (1989) PHYLIP: phylogeny inference package. *Cladistics*, **5**, 164–166.
30. Page,R.D.M. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.*, **12**, 357–358.
31. Emanuelsson,O., Nielsen,H. and Von Heijne,G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.*, **8**, 978–984.
32. Nair,R. and Rost,B. (2005) Mimicking cellular sorting improves prediction of subcellular localization. *J. Mol. Biol.*, **348**, 85–100.
33. Emanuelsson,O., Nielsen,H., Brunak,S. and von Heijne,G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.*, **300**, 1005–1016.
34. Nakai,K. and Kanehisa,M. (1991) Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins*, **11**, 95–110.
35. Hennig,L. (1999) WinGene/WinPep: user-friendly software for the analysis of amino acid sequences. *Biotechniques*, **26**, 1170–1172.
36. Herdenberger,F., Holländer,V. and Kück,U. (1994) Correct *in vivo* RNA splicing of a mitochondrial intron in algal chloroplasts. *Nucleic Acids Res.*, **22**, 2869–2875.
37. Jasper,F., Quednau,B., Kortenjann,M. and Johanningmeier,U. (1991) Control of *cab* gene expression in synchronized *Chlamydomonas reinhardtii* cells. *J. Photochem. Photobiol. B Biol.*, **11**, 139–150.
38. Kyte,J. and Doolittle,R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–132.
39. Franzen,L.G., Rochaix,J.D. and von Heijne,G. (1990) Chloroplast transit peptides from the green alga *Chlamydomonas reinhardtii* share features with both mitochondrial and higher plant chloroplast presequences. *FEBS Lett.*, **260**, 165–168.
40. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
41. Apt,K.E., Zaslavkaia,L., Lippmeier,J.C., Lang,M., Kilian,O., Wetherbee,R., Grossman,A.R. and Kroth,P.G. (2002) *In vivo* characterization of diatom multipartite plastid targeting signals. *J. Cell. Sci.*, **115**, 4061–4069.
42. Dong,A.W., Liu,Z.Q., Zhu,Y., Yu,F., Li,Z.Y., Cao,K.M. and Shen,W.H. (2005) Interacting proteins and differences in nuclear transport reveal specific functions for the NAP1 family proteins in plants. *Plant Physiol.*, **138**, 1446–1456.
43. Boulikas,T. (1993) Nuclear localization signals (NLS). *Crit. Rev. Eukaryot. Gene Expr.*, **3**, 193–227.
44. Kalderon,D., Roberts,B.L., Richardson,W.D. and Smith,A.E. (1984) A short amino acid sequence able to specify nuclear location. *Cell*, **39**, 499–509.
45. Dong,A.W., Zhu,Y., Yu,Y., Cao,K.M., Sun,C.R. and Shen,W.H. (2003) Regulation of biosynthesis and intracellular localization of rice and tobacco homologues of nucleosome assembly protein 1. *Planta*, **216**, 561–570.
46. Hahn,D. and Kück,U. (1995) cDNA nucleotide sequences and expression of the genes encoding the cytoplasmic ribosomal proteins S18 and S27 from the green alga *Chlamydomonas reinhardtii*. *Plant Sci.*, **111**, 73–79.
47. Fuhrmann,M., Oertel,W. and Hegemann,P. (1999) A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. *Plant J.*, **19**, 353–361.

48. Hedtke,B., Meixner,M., Gillandt,S., Richter,E., Börner,T. and Weihe,A. (1999) Green fluorescent protein as a marker to investigate targeting of organellar RNA polymerases of higher plants *in vivo*. *Plant J.*, **17**, 557–561.
49. Schroda,M., Blocker,D. and Beck,C.F. (2000) The *HSP70A* promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. *Plant J.*, **21**, 121–131.
50. Lumbreras,V., Stevens,D.R. and Purton,S. (1998) Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron. *Plant J.*, **14**, 441–447.
51. Calmels,T.P.G., Mistry,J.S., Watkins,S.C., Robbins,P.D., Mcguire,R. and Lazo,J.S. (1993) Nuclear localization of bacterial *Streptoalloteichus hindustanus* bleomycin resistance protein in mammalian cells. *Mol. Pharmacol.*, **44**, 1135–1141.
52. Ronquist,F. and Huelsenbeck,J.P. (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.
53. Valadkhan,S. (2005) snRNAs as the catalysts of pre-mRNA splicing. *Curr. Opin. Chem. Biol.*, **9**, 603–608.
54. Ostheimer,G.J., Rojas,M., Hadjivassiliou,H. and Barkan,A. (2006) Formation of the CRS2-CAF2 group II intron splicing complex is mediated by a 22-amino acid motif in the COOH-terminal region of CAF2. *J. Biol. Chem.*, **281**, 4732–4738.
55. Nakamura,T., Ohta,M., Sugiura,M. and Sugita,M. (1999) Chloroplast ribonucleoproteins are associated with both mRNAs and intron-containing precursor tRNAs. *FEBS Lett.*, **460**, 437–441.
56. Jaeger,S., Eriani,G. and Martin,F. (2004) Results and prospect's of the yeast three-hybrid system. *FEBS Lett.*, **556**, 7–12.
57. Franzen,J.S., Zhang,M.C. and Peebles,C.L. (1993) Kinetic analysis of the 5′ splice junction hydrolysis of a group II intron promoted by domain 5. *Nucleic Acids Res.*, **21**, 627–634.
58. Pyle,A.M. and Green,J.B. (1994) Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry*, **33**, 2716–2725.
59. Black,D.L., Chan,R., Min,H., Wang,J. and Bell,L. (1998) The electrophoretic mobility shift assay for RNA binding proteins. In Smith,C.W.J. (ed.), *RNA: Protein Interactions: A Practical Approach. The Practical Approach Series*. IRL Press, Oxford, UK, Vol. **192**, pp. 109–136.
60. Houser-Scott,F. and Engelke,D.R. (2001) Protein-RNA Interactions. *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd, Chichester, UK.
61. Steer,W.M., Abu-Daya,A., Brickwood,S.J., Mumford,K.L., Jordanaires,N., Mitchell,J., Robinson,C., Thome,A.W. and Guille,M.J. (2003) *Xenopus* nucleosome assembly protein becomes tissue-restricted during development and can alter the expression of specific genes. *Mech. Develop.*, **120**, 1045–1057.
62. Kawase,H., Okuwaki,M., Miyaji,M., Ohba,R., Handa,H., Ishimi,Y., FujiiNakata,T., Kikuchi,A. and Nagata,K. (1996) NAP-I is a functional homologue of TAF-I that is required for replication and transcription of the adenovirus genome in a chromatin-like structure. *Genes Cells*, **1**, 1045–1056.
63. Shikama,N., Chan,H.M., Krstic-Demonacos,M., Smith,L., Lee,C.W., Cairns,W. and La Thangue,N.B. (2000) Functional interaction between nucleosome assembly proteins and p300/CREB-binding protein family coactivators. *Mol. Cell. Biol.*, **20**, 8933–8943.
64. Walter,P.P., Owenhughes,T.A., Cote,J. and Workman,J.L. (1995) Stimulation of transcription factor-binding and histone displacement by nucleosome assembly protein-1 and nucleoplasmin requires disruption of the histone octamer. *Mol. Cell. Biol.*, **15**, 6178–6187.
65. Levchenko,V. and Jackson,V. (2004) Histone release during transcription: NAP1 forms a complex with H2A and H2B and facilitates a topologically dependent release of H3 and H4 from the nucleosome. *Biochemistry*, **43**, 2359–2372.
66. Leister,D. (2003) Chloroplast research in the genomic age. *Trends Genet.*, **19**, 47–56.
67. Barkan,A. (2004) Intron splicing in plant organelles. In Daniell,H. and Chase,C. (eds), *Molecular Biology and Biotechnology of Plant Organelles*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 281–308.
68. Mohr,S., Matsuura,M., Perlman,P.S. and Lambowitz,A.M. (2006) A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc. Natl Acad. Sci. USA*, **103**, 3569–3574.
69. Lehmann,K. and Schmidt,U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol.*, **38**, 249–303.
70. Hahn,D., Nickelsen,J., Hackert,A. and Kück,U. (1998) A single nuclear locus is involved in both chloroplast RNA *trans*-splicing and 3′ end processing. *Plant J.*, **15**, 575–581.
71. Michel,F., Umesono,K. and Ozeki,H. (1989) Comparative and functional anatomy of group II catalytic introns—a review. *Gene*, **82**, 5–30.
72. Chanfreau,G. and Jacquier,A. (1996) An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.*, **15**, 3466–3476.
73. Hauser,C.R., Gillham,N.W. and Boynton,J.E. (1998) Regulation of chloroplast translation. In Rochaix,J.D., Goldschmidt-Clermont,M. and Merchant,S. (eds), *The Molecular Biology of Chloroplasts and Mitochondria in Chlamydomonas*. Kluwer Academic Publishers, Dordrecht, The Netherlands, Vol. 7, pp. 197–217.
74. Boudreau,E., Nickelsen,J., Lemaire,S.D., Ossenbuhl,F. and Rochaix,J.D. (2000) The *Nac2* gene of *Chlamydomonas* encodes a chloroplast TPR-like protein involved in *psbD* mRNA stability. *EMBO J.*, **19**, 3366–3376.
75. Barnes,D., Cohen,A., Bruick,R.K., Kantardjieff,K., Fowler,S., Efuet,E. and Mayfield,S.P. (2004) Identification and characterization of a novel RNA binding protein that associates with the 5′-untranslated region of the chloroplast *psbA* mRNA. *Biochemistry*, **43**, 8541–8550.
76. Fargo,D.C., Boynton,J.E. and Gillham,N.W. (2001) Chloroplast ribosomal protein S7 of *Chlamydomonas* binds to chloroplast mRNA leader sequences and may be involved in translation initiation. *Plant Cell*, **13**, 207–218.
77. Hauser,C.R., Gillham,N.W. and Boynton,J.E. (1996) Translational regulation of chloroplast genes—proteins binding to the 5′-untranslated regions of chloroplast mRNAs in *Chlamydomonas reinhardtii*. *J. Biol. Chem.*, **271**, 1486–1497.
78. Nakamura,T., Schuster,G., Sugiura,M. and Sugita,M. (2004) Chloroplast RNA-binding and pentatricopeptide repeat proteins. *Biochem. Soc. T*, **32**, 571–574.
79. Jacquier,A. and Michel,F. (1990) Base-pairing interactions involving the 5′ and 3′-terminal nucleotides of group II self-splicing introns. *J. Mol. Biol.*, **213**, 437–447.