The plastid-encoded psbA gene in the dinoflagellate *Gonyaulax* is not encoded on a minicircle

Yunling Wang, David Morse *

Département de Sciences Biologiques, Université de Montréal, 4101 Sherbrooke est, Montreal, Quebec, Canada H1X 2B2

Received 22 April 2005; received in revised form 28 November 2005; accepted 29 November 2005

Available online 15 February 2006

Received by W. Martin

Abstract

In all dinoflagellate species studied to date, the plastid genome is highly reduced, with many genes normally found in the plastid genome found instead encoded by the nucleus. Furthermore, those genes still remaining in the plastid are found as primarily single gene minicircles whose size is typically only 2–3 kb. We show here that the plastid genome architecture in the dinoflagellate *Gonyaulax polyedra* is unusual for this class of organism. In particular, the psbA gene is associated with DNA of roughly 50–150 kb and appears to have an unusually high complexity.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Plastid genome; Peridinin-containing dinoflagellate; Plastid evolution; psbA

1. Introduction

The plastid genomes of all dinoflagellates that have been studied to date are characterized by two peculiar features. First, plastid genes are found on small, plasmid-like minicircles. These minicircles have been observed in *Heterocapsa triqueta* (Zhang et al., 1999), *H. pygmeae* and *Protoceratium reticulatum* (Zhang et al., 2002), *Amphidinium operculatum* (Barbrook and Howe, 2000; Barbrook et al., 2001; Nisbet et al., 2004), *Amphidinium carterae* (Hiller, 2001) and *Symbiodinium* sp. (Moore et al., 2003). The minicircles from any given species have a common core conserved among the different genes within a species, and some motifs of the core are also found conserved between species, but no elements are conserved between genera. Typically, the minicircles contain a single gene (Zhang et al., 1999; Barbrook and Howe, 2000) although exceptions containing two or three genes and “empty” minicircles lacking any recognizable protein coding sequence (Barbrook et al., 2001; Hiller, 2001; Nisbet et al., 2004) have been documented. Those minicircles containing several genes do not appear to transcribe them as a polycistrionic message (Barbrook et al., 2001).

The second peculiar feature of the plastid genome is the paucity of genes. Indeed, very few of the genes normally found in plastids are found on the minicircles. While it might initially be thought that the remaining plastid genes might be found in an alternate plastid genome structure, recent EST sequencing projects have discovered that many of the expected proteins are actually encoded by the dinoflagellate nucleus. Indeed, a number of proteins encoded by the plastids of higher plants and other algae are nuclear encoded in *Alexandrium tamarense* (Hackett et al., 2004), *A. carterae* and *Gonyaulax polyedra* (Bachvaroff et al., 2004). The nuclear location for these plastid-directed proteins was confirmed by the presence of a unique N-terminal extension required for plastid targeting in dinoflagellates (Nassoury et al., 2003), the presence of a poly(A) tail, and the expected high GC-content of the sequences (Spector, 1984). The only genes remaining in the plastid are the minicircle genes encoding the 16S and 23S rRNA, eight photosystem components (*psaA*-B, *psbA*-E, *psbI*), two ATP synthase subunits (*atpA*-B), two cytochrome b$_6$f subunits (*petB, petD*) as well as two unidentified ORFs (*ycf16, ycf24*) and two ribosomal...
proteins (rpl28, rpl23). Dinoflagellate plastid genomes thus contain a highly reduced gene complement distributed amongst many small circular elements.

One possible caveat to the idea that these unusual features reflect a unique plastid genome is the suggestion that the minicircles may in fact be encoded by the nucleus (Laatsch et al., 2004). However, in Symbiodinium, psbA transcripts have been observed in the plastids by in situ hybridization (Takashita et al., 2003). Furthermore, none of the minicircle genes sequences to date encode the characteristic N-terminal leader sequence required for plastid targeting (Nassoury et al., 2003). Lastly, we have recently demonstrated that psbA synthesis in vivo is inhibited by chloramphenicol but not by cycloheximide in Gonyaulax and Amphidinium (Wang et al., 2005). Taken together, these studies provide strong presumptive evidence for the presence of minicircles in the plastid. However, since few dinoflagellate species have been examined, it is still a formal possibility that different dinoflagellates may utilize different strategies for encoding and expressing the plastid genome.

We have been interested for some time in the genes encoded by the plastid of the dinoflagellate Gonyaulax, but have so far been unsuccessful in cloning these genes based on sequence similarity to minicircle sequences of other dinoflagellate species. We report here a surprising structural difference in the plastid genome architecture of Gonyaulax. In particular, we find that psbA from this species is not found on a minicircle. The plastid genome organization in Gonyaulax thus appears quite different from that described for the other dinoflagellates.

2. Material and Methods

2.1. Cell culture

A. carterae (CCMP 1314) and G. polyedra (now Lingulodinium polyedrum; strain 70) were obtained from the Provasoli-Guillard Culture Center for Marine Phytoplankton (Boothbay Harbor, Maine) and grown in a modified seawater medium (F2) (Guillard and Ryther, 1962) at constant temperature (18 ± 1 °C) in 12-h light/12-h dark cycles using cool white fluorescent light at an intensity of 50 μmol photons m⁻² s⁻¹. The beginning of light period is defined as time 0 (LD 0), and the beginning of the dark period as LD 12. Cultures were grown to a cell density of 12,000 to 14,000 cells/mL (Gonyaulax) or 500,000 cells/mL (Amphidinium) as measured by counting using a hemocytometer. Both cultures were unialgal but only Amphidinium is axenic.

2.2. Nucleic acid isolation

Two grams wet weight cells, harvested by filtration on Whatman 541 paper, were ground in liquid nitrogen and resuspended in 10 mL lysis buffer (20 mM Tris pH 8 containing 50 mM EDTA and 100 mM NaCl). One milliliter 20% SDS and 0.5 mL 20 mg/mL proteinase K were added, mixed well and incubated at 50 °C. After 2 h, an additional 0.5 mL 20 mg/mL proteinase K was added and left overnight at room temperature. The mixture was extracted twice with an equal volume of phenol/chloroform and the nucleic acids precipitated from the aqueous fraction by addition of two volumes 100% ethanol. DNA was air-dried and resuspended in 2.5 mL TE, and 1.1 volumes of CTAB buffer (250 mM Tris HCl, pH 8, 100 mM EDTA, 55 mM hexadecylamine triammonium bromide, and 2.5 M NaCl) added and incubated for 45 min at 65 °C with vortexing every 5 min. The solution was extracted with one volume chloroform and 5.5 g CsCl and 5 μL 10 mg/mL bisbenzimide were added to the aqueous solution. Following centrifugation at 45,000 rpm in a Beckman 70.1 Ti rotor for 24 h at 20 °C, DNA bands were visualized by long wave UV, collected with a large bore needle, and precipitated by centrifugation after addition of 3 vol. TE.

2.3. Pulse field gel electrophoresis

Pulse field gel electrophoresis was carried out on 1% agarose gels in 0.5× TBE. Typically, samples were run at 6 V/cm for 15 h with a pulse ramp from 1 to 30 s. Samples included purified DNA obtained from CsCl gradients or total DNA from cells ground in liquid nitrogen. For the latter, powder from 0.5 g wet weight ground cells was resuspended in 5 mL of lysis buffer (10 mM Tris, pH 7.2, 50 mM NaCl, 0.2% sodium deoxycholate, 0.5% sodium lauryl sarcosine) and incubated at 50 °C for 10 min. The suspension was combined with plug molds at 50 °C using sterile transfer pipettes. The agarose plugs were solidified at 4 °C, incubated in lysis buffer for 1 h at 37 °C, rinsed with wash buffer (20 mM Tris, pH 8.0, 50 mM EDTA) and incubated with proteinase K at 50 °C overnight without agitation. After the protease treatment, the plugs were washed 4 times at room temperature using 50 mL of wash buffer for 1 h each. The washed DNA plugs were cut to a size 90% of the well height and gently pressed to the bottoms of the wells. Each sample well was filled with agarose at the same concentration as that in the gel, and the agarose allowed to set for 10 to 15 min before electrophoresis.

2.4. Cloning and sequencing

Restriction digests were performed on DNA isolated from each species and subcloned into pUC19 (Promega). DNA was isolated from colonies of several clones containing inserts of the desired size. DNA from colonies was amplified by PCR using AmpliTaq DNA polymerase (Perkin Elmer). The amplified DNA was cloned into the pUC19 vector and inserted into the E. coli host, DH10B. Plasmids were then isolated from transformed E. coli and amplified by PCR. The amplified DNA was sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer) with AmpliTaq FS DNA polymerase (Perkin Elmer). The amplified DNA was isolated from colonies of several clones containing inserts of the desired size.

2.5. Fluorescence microscopy

Aliquots of 10 μL were added to 50 μL water and diluted 50× with water. The diluted sample was examined with a Leica DMR microscope equipped with a Leica DMR filter set. A 100× dry objective was used for observation.

Table 1. Sequence similarities of psbA from various dinoflagellate species. Sequence identities are shown for four dinoflagellate psbA sequences at the DNA (upper right) and protein (lower left) levels.

<table>
<thead>
<tr>
<th>Species</th>
<th>DNA</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexandrium</td>
<td>66%</td>
<td>72%</td>
</tr>
<tr>
<td>Amphidinium</td>
<td>78%</td>
<td>79%</td>
</tr>
<tr>
<td>Heterocapsa</td>
<td>83%</td>
<td>82%</td>
</tr>
<tr>
<td>Gonyaulax</td>
<td>86%</td>
<td>82%</td>
</tr>
</tbody>
</table>

Fig. 1. Sequence similarities of psbA from various dinoflagellate species. Sequence identities are shown for four dinoflagellate psbA sequences at the DNA (upper right) and protein (lower left) levels.
2.4. Nucleic acid analysis

Gene probes for psbA and 23S rDNA were amplified from *Gonyaulax* DNA using PCR (Wang et al., 2005). All sequences were analyzed using MacVector software (Accelrys). For high stringency hybridization of nucleic acids, high temperature (65 °C) in 0.5 M Phosphate/5% SDS/2% BSA method was used (Church and Gilbert, 1984), while for low stringency (<65 °C) in 0.5 M Phosphate/5% SDS/2% BSA method was used (Church and Gilbert, 1984), while for low stringency hybridization, the temperature was reduced to 55 °C. Other standard molecular techniques were performed as described (Sambrook et al., 1989).

3. Results

3.1. *Gonyaulax* psbA is not found on minicircular DNA

DNA extracted from *G. polyedra* contains a psbA gene (encoding the photosystem D1) similar to that found in the dinoflagellate *A. carterae* and other dinoflagellates (Fig. 1). Curiously, PCR reactions using the reverse complement of the oligonucleotides used to amplify the psbA coding sequence did not amplify a minicircle sequence, as expected from the plastid genome structure known for other dinoflagellates. To test directly for the presence of a psbA minicircle in *Gonyaulax*, we probed undigested genomic DNA on Southern blots with the psbA sequence. The *Gonyaulax* DNA hybridizing with psbA appears significantly larger (>20 kb) than the roughly 3 kb DNA fragment observed using undigested DNA from *Amphidinium* (Fig. 2). The *Amphidinium* psbA is found only as a minicircle when probed with a homologous probe at high stringency, and the slight hybridization signal to larger DNA fragments observed here with is due to the low stringency hybridization with our heterologous probe. *E. coli* DNA, chosen as a negative control lacking psbA, also hybridizes weakly under these conditions. We conclude that the psbA in *Gonyaulax* is not encoded by minicircular DNA.

3.2. *Gonyaulax* psbA is found in an AT-rich context

The psbA sequence is AT-rich, unlike all nuclear-encoded genes from *Gonyaulax* studied to date. To determine if the psbA was found in an AT-rich context, we centrifuged total DNA extracts from *Gonyaulax* on CsCl gradients in the presence of bisbenzimide. Of the two bands resolved using this technique, only the upper AT-rich DNA was found to hybridize to our psbA probe (Fig. 3). In contrast, the lower GC-rich DNA was found to hybridize preferentially to a probe for the nuclear-encoded plastid-directed peridinin-chlorophyll a-protein (PCP) gene (Fig. 3). The small degree of PCP hybridization to DNA in the upper band may reflect difficulty in completely purifying away the large amount of PCP genes, found in roughly 5000 copies per *Gonyaulax* genome (Le et al., 1997) and the large amount of nuclear DNA, estimated at 200 pg per nucleus (Spector, 1984). We conclude that our psbA sequence is found in an AT-rich context.

We were curious as to the physical size of the DNA encoding the psbA gene, since agarose gels do not allow fragments greater than 20 kb to be resolved. In a first approach, we attempted to estimate the size of the plastid genome by restriction fragment analysis of the AT-rich psbA-containing DNA. For typical plastids, this approach estimates plastid genome sizes of between 100 to 200 kb. However, no distinct bands were ever observed by HindIII digestion of AT-rich DNA (Fig. 3 panel iii). DNA extracted from *Gonyaulax* is always difficult to digest completely, and thus the restriction of the DNA was verified by hybridization to psbA (the sequence contains a single internal HindIII site). In the samples shown, at

Fig. 2. The psbA gene in *Gonyaulax* is not a minicircle. Low stringency Southern blot analysis of total DNA extracted from *Gonyaulax polyedra* (lane 2), *Amphidinium carterae* (lane 3) and *Escherichia coli* (lane 4) electrophoresed on a standard 1.2% agarose gel. Transfers were hybridized to either a 900 bp fragment of *Gonyaulax* 23S RNA or to a 700 bp fragment of *Gonyaulax* psbA. Molecular weight markers (lane 1, three sizes in kb shown at left) are a BstE II digest of lambda DNA. The open triangles (right) indicate ~20 kb and the closed triangle 3 kb. The ethidium bromide stained gel is at top, and the two Southern blots with the indicated probes below.

Fig. 3. The psbA gene in *Gonyaulax* is in an AT-rich context. (i) Total DNA extracts were resolved into GC-rich (lane 2) and AT-rich fractions (lane 3) by CsCl gradient centrifugation in the presence of bisbenzimide and hybridized to a 700 bp fragment of *Gonyaulax* psbA. Molecular weight markers (lane 1) are in Fig. 2. Ethidium bromide staining is at left and the corresponding Southern blot at right. (ii) As in panel i, except that the nuclear-encoded PCP gene was used a probe. (iii) AT-rich DNA alone (lane 2) or following digestion with HindIII (lane 3) was electrophoresed and hybridized to a 0.7 kb fragment of *Gonyaulax* psbA. Molecular weight markers (lane 1) are in Fig. 2.
least half the psbA sequences are digested. Despite this, no distinct DNA bands are visible by ethidium bromide staining. We conclude from this that the AT-rich DNA from *Gonyaulax* may be more complex than that usually found in plastids.

### 3.3. *Gonyaulax* psbA is associated with DNA of 50–150 kb

In a second approach to determining the physical size of the DNA encoding the psbA gene, DNA from whole cells ground in liquid nitrogen and embedded in agarose plugs was analyzed using pulse field gel electrophoresis. Control DNA fragments ranging in size from 50 to 1000 kb are easily resolved using these conditions. *Gonyaulax* DNA (Fig. 4) is found in two size classes, one ≥1000 kb (upper triangle) and one of 50–150 kb (lower triangle). The large DNA presumably reflects nuclear DNA not resolved into individual chromosomes under these experimental conditions. *Gonyaulax* contains roughly $10^{11}$ bp of DNA in a reported 220 chromosomes (Spector, 1984), suggesting that individual chromosomes may be as large as ~500,000 kb. In any event, a clear hybridization signal with the psbA gene probe is found in the region of 50–150 kb (Fig. 4).

The *Gonyaulax* psbA gene is thus associated with a DNA molecule whose size is significantly larger than the minicircles reported for other dinoflagellate species. It is noteworthy that the AT-rich DNA isolated from the bisbenzimide-CsCl gradients is of similar size (Fig. 4), although some breakage of the DNA in these samples is likely to have occurred during the purification procedure.

### 4. Discussion

#### 4.1. The plastid genome in *Gonyaulax* appears large and complex

It is now generally accepted that dinoflagellate plastid genomes are highly unusual, and are comprised of a limited number of genes encoded on small plasmid-like minicircles. We find that the plastid genome architecture in *Gonyaulax* is quite different from that typical of dinoflagellates. Instead of the small 2–3 kb minicircles found in other species, we find that the psbA gene in *Gonyaulax* associated with DNA of roughly 50–150 kb. While this amount of DNA appears on the surface to be similar to the genome size of higher plant plastids, several lines of evidence suggest to us that the *Gonyaulax* plastid genome may in fact be even larger. First, hybridization of our pulse field gels with a psbA probe shows a smear with undigested DNA (Fig. 4) but two small bands after HindIII digestion (Fig. 3). The hybridization to the digested DNA is consistent with the presence of a single HindIII in the psbA sequence, but the smeared hybridization signal observed with undigested DNA might be interpreted either as several conformations of the same molecule or as several different molecules. These characteristics of hybridization to a psbA probe are reminiscent of the hybridization of a cox1 probe to mitochondrial DNA from the dinoflagellate *Cryptothecodinium*: a smear between 6 and 9 kb is observed when undigested DNA is probed, and discrete bands are observed when a restriction enzyme digest is probed (Norman and Gray, 2001). These observations have led to the suggestion that the mitochondria may contain many different molecules, and an analogous situation may thus also occur in the *Gonyaulax* plastid genome.

A second line of evidence is based on a potentially large amount of non-coding DNA in the plastid genome. For example, random sequencing of mechanically sheared DNA fragments recovered no identifiable genes in ~100 kb of sequence. This suggests that either large amounts of non-coding DNA are present or that the plastid DNA is heavily edited. While editing in the dinoflagellate mitochondrial genome appears widespread and extensive (Lin et al., 2002), and editing has been reported for the minicircles in the dinoflagellate *Ceratium* (Zauner et al., 2004) there is no evidence of editing in *Amphidinium* minicircles (Barbrook and Howe, 2000). While we cannot conclusively rule out in editing in *Gonyaulax*, our psbA sequence amplified from DNA was almost identical to the sequence amplified from RNA, suggesting this possibility is unlikely.

A third line of evidence supporting a high complexity plastid genome is that no discrete bands are observed by ethidium bromide staining of restriction enzyme digests of the AT-rich DNA. Plastid genomes in general, and curiously enough, the first reports of dinoflagellate plastid genomes (using one species of *Glenodinium* and two species of *Protogonyaulax*), generally produce a limited number of gene fragments after restriction enzyme digestion (Koczar et al., 1991). While neither *Glenodinium* nor *Protogonyaulax* have been reexamined for the presence of a minicircular plastid genome, our results with digestion of *Gonyaulax* AT-rich DNA argues for a high complexity DNA sample.

Lastly, although we have so far only identified one other gene fragment belonging to the *Gonyaulax* plastid genome (a 0.9 kb region with homology to dinoflagellate 23 rRNA), long range PCR combining primers from both psbA and 23S genes did not amplify any DNA. This suggests that either the two sequences are too far apart to be amplified (>40 kb) or are located on different molecules. Taken together, then, our results point for
an unusual organization of the plastid genome in the dinoflagellate *Gonyaulax*: two plastid genes (psbA and 23S) are not encoded by minicircles, and it appears as though a large proportion of the genome may be non-coding. At this point, it is impossible to determine whether the *Gonyaulax* plastid genome will be ancestral or derived. Clearly, it will be necessary to isolate many more plastid genes from this species before this aspect to the evolution of dinoflagellate plastids can be resolved.

### 4.2. Dinoflagellate plastids have many unique features whose provenance is unknown

The reddish-colored peridinin-containing dinoflagellate plastids are thought to have originated from a secondary endosymbiotic event, and are distinct from the so-called "green" dinoflagellates whose plastids contain fucoxanthin (Yoon et al., 2005). However, the identity of the ancestor to the peridinin-type plastid is still enigmatic. *Gonyaulax*, for example, contains a form II ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (Morse et al., 1995), an enzyme found previously only in anaerobic proteobacteria. Furthermore, the major light harvesting protein in this organism, the soluble peridinin-chlorophyll a-protein (Lee et al., 1993) is found in no other extant class of organisms, and interesting, has a three-dimensional structure unrelated to any other light harvesting protein (Hofmann et al., 1996). Lastly, the nuclear-encoded plastid-directed glyceraldehyde-3-phosphate dehydrogenase (GAP), plastid gene appear to be derived from lateral gene transfer and cryptomonads (Fagan et al., 1998). Although much further work will be required to fully understand the complex evolution of the dinoflagellate plastid, the *Gonyaulax* plastid genome in particular deserves attention because of its dissimilarity with the plasmid-like circles found in other species.

### Acknowledgements

We thank B. Green (University of British Columbia) for critical analysis of our initial experiments and helpful suggestions for further work. We gratefully acknowledge the technical expertise of B.F. Lang (University of Montreal) in cloning and random sequencing the AT-rich DNA. This work has been supported by a grant from the National Science and Engineering Research Council of Canada.

### References


