Volume 298, number 1, 93–96 FEBS 10714 February 1992

A *sec*Y homologue is found in the plastid genome of *Cryptomonas Φ*

Susan E. Douglas

Institute for Marine Biosciences, National Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada

Received 25 November 1991

An open reading frame with significant similarity to the *sec*Y gene of *Escherichia coli* has been found within a ribosomal protein operon on the plastid genome of the chlorophyll *c*-containing alga *Cryptomonas Φ*. The gene encodes a protein of 420 amino acids (molecular weight 46,906 daltons) and contains ten potential membrane-spanning domains, as in the *E. coli* homologue. This report of a *sec*Y homologue in a plastid genome provides preliminary evidence that a prokaryotic-like protein export system may be operating in plastids.

Algae; Plastid; Ribosomal protein operon; Secretion; Chloroplast endoplasmic reticulum

1. INTRODUCTION

In Gram-negative Eubacteria, at least six proteins (SecA, B, D, E, F and Y) are required for targeting and protein translocation across the plasma membrane (see [1]). Many of these proteins have been isolated and their genes cloned and sequenced. The cytosolic SecB protein functions as a chaperonin and as part of a receptor cascade [2], the membrane-associated SecA protein functions as an ATPase [3] and the integral membrane proteins SecY and SecE mediate translocation by interacting with SecA [4]. Integral membrane proteins SecD and SecF may be required for the later stages of protein export or for interaction of precursors with signal peptidases [5]. Recently, *sec* homologues have also been found in the Gram-positive Eubacterium, *Bacillus subtilis* [6–8].

In land plants and green algae, translocation across plastid membranes (see [9],[10–12]) and targeting of nuclear-encoded proteins to the plastid envelope membrane of [13,14] have been studied. Cleavage sites for plastid transit peptides [15] and thylakoid lumen leader peptides [16,17] have also been investigated. However, protein targeting and translocation has not been investigated in the rhodophyte, chromophyte or cryptophyte algae.

Chromophyte and cryptophyte algae present especially interesting systems for studying protein translocation since the plastids of these organisms are surrounded by an extra pair of membranes, termed the chloroplast endoplasmic reticulum, or CER [18]. The outermost membrane of the CER, which bears eukaryotic ribosomes and is usually continuous with the nuclear membrane, is probably derived from the endomembrane system of the host, while the innermost membrane is probably derived from the plasma membrane of the eukaryotic endosymbiont which gave rise to the plastid [19,20]. Cryptomonads are unique among the CER-containing algae in retaining a nucleomorph (the vestigial nucleus of the eukaryotic endosymbiont) in the periplastidal space between the CER and plastid envelope [21].

Presumably in an alga as complex as *Cryptomonas Φ*, gene products encoded by the nuclear or nucleomorph genomes that are targeted to the plastid, contain distinctive N-terminal extensions which serve as signals in intracellular sorting and assist these proteins to traverse the additional membranes of the CER. Those gene products destined for the thylakoid of the plastid may contain additional signals. However, sequence analysis of *Cryptomonas Φ* plastid *cpe*B gene, which encodes the *β* subunit of phycoerythrin (a lumenal protein), does not predict an N-terminal extension [22]. On the other hand, the sequences of genomic clones of *cpe*A from the cryptomonad *Chroomonas*, do predict a leader sequence [23]. The mechanism of protein translocation across either the thylakoid, plastid or CER membranes is unknown although Gibbs [24] has suggested that vesicles carrying nucleus-encoded proteins destined for the plastid bud off from the inner CER membrane and then fuse with the outer plastid envelope membrane.

In both *B. subtilis* and *E. coli*, the integral membrane protein SecY is encoded at the promoter distal region of the *spc* ribosomal protein operon. Although the *sec*Y gene is not present on any of the three land plant plastid genomes which have been completely sequenced [25–27], it is possible that it might reside on the plastid genome of a less advanced alga. It appears that gene transfer from the plastid progenitor to the host genome has not proceeded as far in some algae as in land plants

Correspondence address: S.E. Douglas, Institute for Marine Biosciences, National Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada. Fax: (1) (902) 426-9413.

[28,29] and in cryptomonads in particular, there appear to be many genes which are encoded on the plastid rather than the nuclear genome [30,31]. For example, sequence analysis of the *str* operon from *Cryptomonas Φ* shows the presence of ribosomal protein genes which are not found on plastid genomes of land plants [32]. In order to see if a *sec*Y homologue is present on the plastid genome of *Cryptomonas Φ*, the promoter distal region of the *spc* ribosomal protein operon was sequenced.

2. MATERIALS AND METHODS

Isolation of plastid DNA from *Cryptomonas Φ* and construction of a clone bank has been described [33]. Ribosomal protein operons were mapped on the plastid genome by hybridisation with heterologous probes and by sequencing [32]. The promoter distal region of the *spc* operon which potentially would encode *sec*Y was located on two adjacent 1.9 kb *Sal*I fragments. Both strands of the DNA from this region were completely sequenced using the dideoxy chain termination method [34] and synthetic oligonucleotide primers. Analysis of the DNA sequence and the derived amino acid sequence was performed using the DNA Strider version 1.1 program [35] and databank searching was performed using the FastP program [36] on a Macintosh II computer (Apple Computers).

3. RESULTS AND DISCUSSION

A physical map of the plastid genome of *Cryptomonas Φ* is shown in Fig. 1a. The positions of the *S10*, *spc*, *alpha* and *str* ribosomal protein operons are indicated by solid bars. The promoter distal region of the *spc* operon is shown in Fig. 1b and the locations of coding regions are indicated by stippled bars. An open reading frame of 1,260 nucleotides with similarity to *sec*Y from *E. coli* was found which spanned the two 1.9 kb *Sal*I fragments. The nucleotide sequence of *sec*Y and the flanking upstream and downstream regions, including the 5'-terminus of *rpl*36 (the gene encoding ribosomal protein L36), are shown in Fig. 2. The derived amino acid sequence (420 residues, 46,906 Da) is given below the nucleotide sequence. As with other *Cryptomonas Φ* plastid genes [33,35], there is very little space between the two genes (only 27 nucleotides).

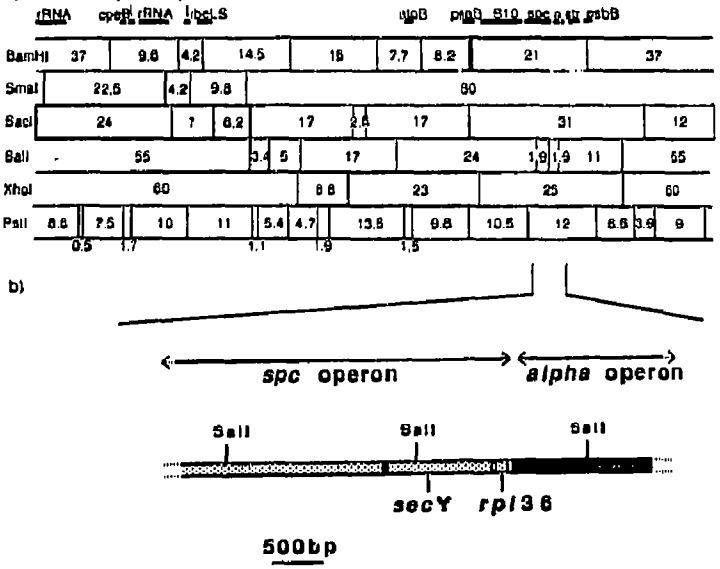

Fig. 1. (a) Physical map of the plastid genome of *Cryptomonas Φ*. The locations of the *S10*, *spc*, *alpha* and *str* ribosomal protein operons and other coding regions are indicated by solid bars. (b) Region of the plastid genome containing the *sec*Y gene. Coding regions within the *spc* and *alpha* ribosomal protein operons are denoted by light and dark stippling, respectively.

An alignment of the *sec*Y gene products from *Cryptomonas Φ*, *B. subtilis* and *E. coli* is shown in Fig. 3. The two proteins are essentially colinear, with the exception of N- and C-terminal extensions in the *E. coli* sequence relative to those from *Cryptomonas Φ* and *B. subtilis*. There are several other places where small insertions or deletions occur. Excluding the N- and C-terminal extensions, the amino acid sequences from *Cryptomonas Φ* and *E. coli* are 38.7% identical and 60.1% similar (when conservative replacements are included). The similarity to the SecY sequence from *B. subtilis* is only slightly lower (35.3 and 60.7%, respectively).

In *E. coli*, SecY is an integral membrane protein which contains ten transmembrane segments, five periplasmically exposed regions and six cytoplasmically exposed regions [37]. It is part of the membrane-bound translocase which also contains the SecE [4] and possibly SecD and SecF proteins [5]. The hydropathic profiles of the SecY proteins from *E. coli*, *B. subtilis* and *Cryptomonas Φ* are virtually superimposable (Fig. 4) and the ten putative membrane-spanning domains are present in all three molecules.

The fact that the *sec*Y gene is found within the transcriptionally active *spc* ribosomal operon (Wang, Liu and Douglas, in preparation) indicates that it is not an inactive pseudogene and that the SecY polypeptide is functional in the cell. Also, the conservation in primary amino acid sequence and, more importantly, the hydropathy profile, argues for a functional gene product. The discovery of a gene for a component of a prokaryotic protein export system in a plastid genome is intriguing and may shed light on possible mechanisms of protein translocation in plant cells. Thus far, no *sec* genes have been reported in any plant system, possibly because they are present in nuclear genomes rather than the extensively studied plastid genomes. The plastid genome of *Cryptomonas Φ* contains genes not found on other plant plastid genomes and therefore is a useful system for studies of plastid functions which are normally dependent on the products of nuclear-encoded genes. The possibility that other *sec* genes are also plastid-encoded in *Cryptomonas Φ* is currently under investigation.

The subcellular location of the SecY from *Cryptomonas Φ* is unknown. The inner membrane of the plastid envelope is the homologue of the plasma membrane of *E. coli* and it is possible that SecY is involved

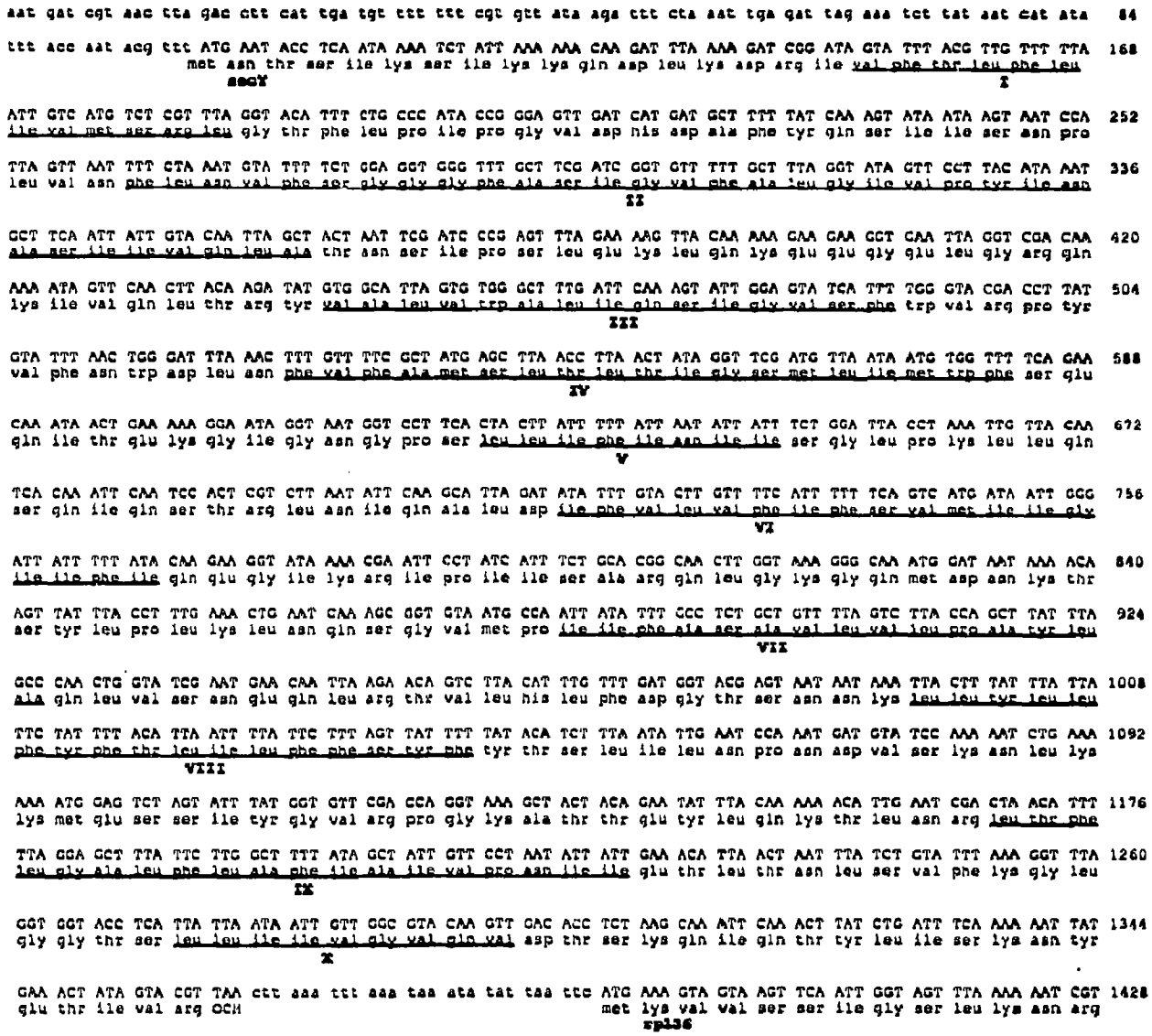

Fig. 2. Nucleotide sequence of the *sec*Y gene and flanking regions from *Cryptomonas Φ*. The deduced amino acid sequence is shown beneath the nucleotide sequence. The positions of putative transmembrane segments (see Fig. 4b) are indicated by roman numerals I–X.

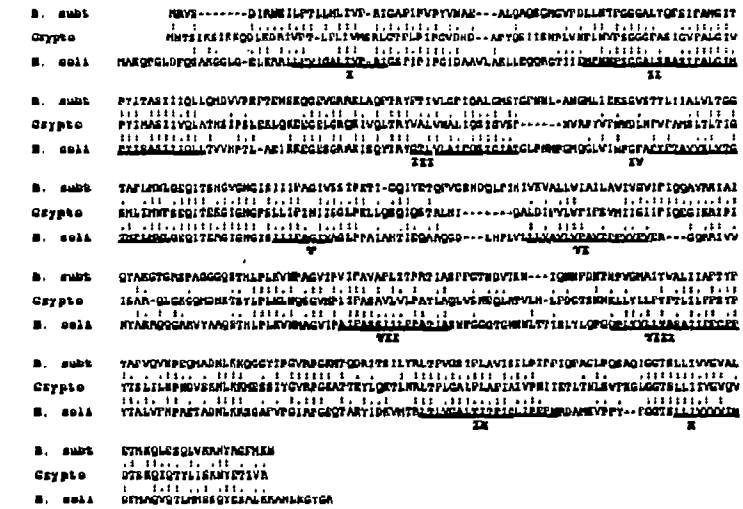

Fig. 3. Alignment of SecY proteins from *Cryptomonas Φ*, *Bacillus subtilis* and *Escherichia coli*. Colons indicate identities and dots indicate conservative replacements according to Schwartz and Dayhoff [38].

in protein translocation across this membrane. However, most of the protein translocation in plastids is in the reverse direction, i.e. from the cytoplasm into the plastid. SecY may be involved in protein translocation across the thylakoid membranes or even across the CER, although the latter is unlikely given that the outer membrane of this pair is derived from the host endomembrane system rather than the endosymbiont. Possible mechanisms of protein targeting and translocation in *Cryptomonas Φ* are currently under investigation.

Acknowledgements: The author wishes to thank Colleen Murphy for technical assistance and Ron MacKay for critical reading of this manuscript. This is NRCC publication number 33009.

REFERENCES

[1] Bieker, K.L., Phillips, G.J. and Silhavy, T.J. (1990) J. Bioenerg. Biomemb. 22, 291–310.

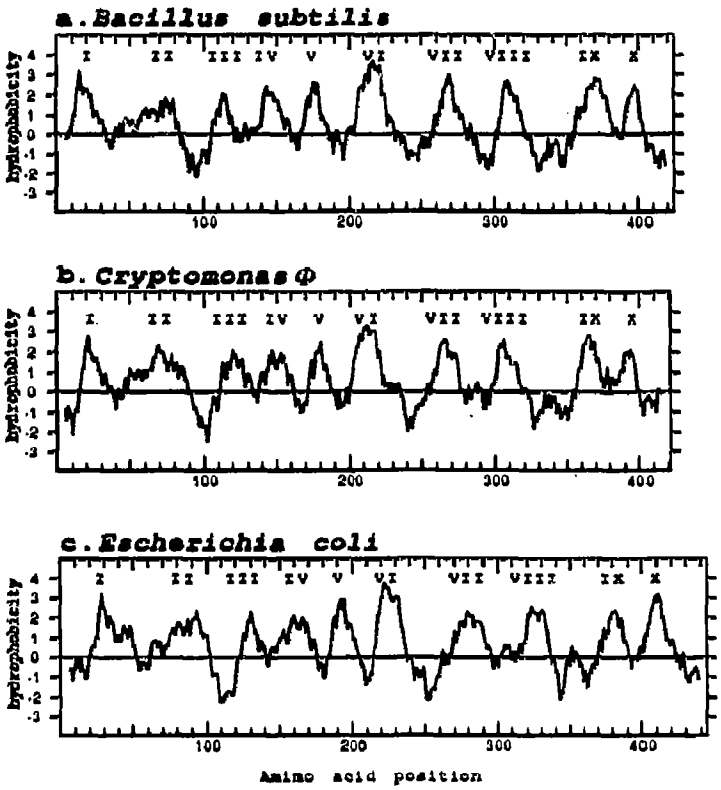

Fig. 4. Hydropathy plots of SecY proteins from *Cryptomonas Φ*, *Bacillus subtilis* and *Escherichia coli*. Plots were obtained using the Kyte-Doolittle option of DNA Strider [35]. The positions of putative transmembrane segments are indicated by roman numerals I–X.

[2] Hartl, F.-U., Lecker, S., Schiebel, E., Hendrick, J.P. and Wickner, W. (1990) Cell 63, 269–279.
[3] Lill, R., Dowham, W. and Wickner, W. (1990) Cell 60, 271–280.
[4] Brundage, L., Hendrick, J.P., Schiebel, E., Dreissen, A.J.M. and Wickner, W. (1990) Cell 62, 649–657.
[5] Gardel, C., Johnson, K., Jacq, A. and Beckwith, J. (1990) EMBO J. 9, 3209–3216.
[6] Nakamura, K., Nakamura, A., Takamatsu, H., Yoshikawa, H. and Yamane, K. (1990) J. Biochem. 107, 603–607.
[7] Suh, J.W., Boylan, S.A., Thomas, S.M., Dolan, K.M., Oliver, D.B. and Price, C.W. (1990) Mol. Microbiol. 4, 305–314.
[8] Overhoff, G., Klein, M., Speis, M. and Freudl, R. (1991) Mol. Gen. Genet. 228, 417–423.
[9] Schmidt, G.W. and Mishkind, M.L. (1986) Annu. Rev. Biochem. 55, 879–912.
[10] Keegstra, K. and Olsen, L.J. (1989) Annu. Rev. Plant Physiol. Plant Mol. Biol. 40, 471–501.
[11] Schnell, D.J., Blobel, G. and Pain, D. (1990) J. Cell Biol. 111, 1825–1838.
[12] Franzen, L.-G., Rochaix, J.-D. and von Heijne, G. (1990) FEBS Lett. 260, 165–168.
[13] von Heijne, G. and Nishikawa, K. (1991) FEBS Lett. 278, 1–3.
[14] Li, H.-M., Moore, T. and Keegstra, K. (1991) Plant Cell 3, 709–717.
[15] Gavel, Y. and von Heijne, G. (1990) FEBS Lett. 261, 455–458.
[16] Last, D.I. and Gray, J.C. (1989) Plant Mol. Biol. 12, 655–666.
[17] Howe, C.J. and Wallace, T.P. (1990) Nucleic Acids Res. 18, 3417.
[18] Gibbs, S.P. (1981) Int. Rev. Cytol. 72, 49–99.
[19] Gibbs, S.P. (1981) Annu. N.Y. Acad. Sci. 361, 193–208.
[20] Douglas, S.E., Murphy, C.A., Spencer, D.F. and Gray, M.W. (1991) Nature 350, 148–151.
[21] Greenwood, A.D., Griffiths, H.B. and Santore, U.J. (1977) Brit. Phycol. J. 12, 119.
[22] Reith, M. and Douglas, S.E. (1990) Plant Mol. Biol. 15, 585–592.
[23] Jenkins, J., Hiller, R.G., Speirs, J. and Godovac-Zimmermann, J. (1990) FEBS Lett. 273, 191–194.
[24] Gibbs, S.P. (1979) J. Cell Sci. 35, 253–266.
[25] Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
[26] Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.-Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) EMBO J. 5, 2043–2049.
[27] Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.-R., Meng, B.-Y., Li, Y.-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989) Mol. Gen. Genet. 217, 185–194.
[28] Baldauf, S.L. and Palmer, J.D. (1990) Nature 344, 262–265.
[29] Gantt, J.S., Baldauf, S.L., Calie, P.J., Weeden, N.F. and Palmer, J.D. (1991) EMBO J. 10, 3073–3078.
[30] Douglas, S.E. and Durnford, D.G. (1989) Plant Mol. Biol. 13, 13–20.
[31] Wang, S. and Liu, X.-Q. (1992) Proc. Natl. Acad. Sci. USA (in press).
[32] Douglas, S.E. (1991) Curr. Genet. 19, 289–294.
[33] Douglas, S.E. (1988) Curr. Genet. 14, 591–598.
[34] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[35] Marck, C. (1988) Nucleic Acids Res. 16, 1829–1836.
[36] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[37] Akiyama, Y. and Ito, K. (1987) EMBO J. 6, 3465–3470.
[38] Schwartz, R.M. and Dayhoff, M.O. (1978) in: Atlas of Protein Sequence and Structure, vol. 5, suppl. 3 (M.O. Dayhoff, ed.) pp. 353–358, National Biomedical Research Foundation, Washington.