Volume 336, number 2, 225–230 FEBS 13413 December 1993

The unusual *rps3*-like orf712 is functionally essential and structurally conserved in *Chlamydomonas*

Xiang-Qin Liu*, Changzhi Huang, Hui Xu

Biochemistry Department, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Received 19 October 1993; revised version received 9 November 1993

The *Chlamydomonas reinhardtii* chloroplast orf712 is a previously described open reading frame that lacks a detectable transcript but potentially encodes a polypeptide with sequence similarities to ribosomal protein Rps3 only at its N- and C-termini. Here we report that *orf712* is an essential gene, as demonstrated through gene disruption by particle gun-mediated chloroplast transformation. We also show that an *orf712* is present and structurally conserved in all of the two or three major *Chlamydomonas* lineages. Our results suggest that *orf712* is an unusual *rps3* gene that contains a large translated intervening sequence.

Chloroplast gene; *rps3*; Ribosomal protein; Gene disruption; Chloroplast transformation; *Chlamydomonas*

1. INTRODUCTION

Plastid ribosomes contain over 60 distinct ribosomal proteins, with their genes distributed between the nuclear and the plastid genomes (e.g. [1]). Plant chloroplast genomes encode about 20 ribosomal proteins, and some algal plastid genomes appear to encode a larger number of these proteins (see [2] for a review). Many of these plastid-encoded ribosomal protein genes are organized in clusters or operons that resemble ribosomal protein operons of *E. coli*, reflecting the bacterial endosymbiont ancestry of plastids. The *rps3* gene, encoding a protein of the 30S ribosomal subunit, is part of the *s10* operon consisting of 11 ribosomal protein genes in *E. coli* [3]. In all plant and algal plastids reported so far, an *rps3* gene is found in a gene cluster (or operon) resembling the bacterial *s10* operon. In chloroplasts of *Chlamydomonas reinhardtii*, however, an *rps3* gene is absent from the *s10*-like gene cluster (Liu et al. unpublished data). Instead, an *rps3*-like open reading frame (orf712) was found at a location more than 75 kbp away from the *s10*-like gene cluster [4].

The *C. reinhardtii* chloroplast orf712 potentially encodes a 712-residue polypeptide (Orf712) of unusual structure. The predicted amino acid sequence of Orf712 shows significant sequence similarity to bacterial and chloroplast ribosomal protein Rps3, but the similarity is limited to the N- and C-termini of Orf712, with the middle two thirds of the Orf712 sequence lacking similarity to all known proteins [4]. This structural oddity and the lack of a detectable transcript of *orf712* [4] raise questions about the identity and functionality of this open reading frame. Here we report that this chloroplast *orf712* is a functional gene required for cell growth, and that its unusual structural feature is conserved among distantly related *Chlamydomonas* species.

*Corresponding author. Fax: (1) (902) 494-1355.

2. MATERIALS AND METHODS

Wild-type strains of *Chlamydomonas reinhardtii* (137c mt^+), *Chlamydomonas peterfii*, *Chlamydomonas humicola*, and *Chlamydomonas frankii* were obtained from Dr. Robert Lee at Dalhousie University (Halifax, Nova Scotia). Cells were grown on TAP (Tris-acetate-phosphate) or HS (high salts) medium as described by Harris [5].

Chloroplast transformation was performed using a home-made particle inflow gun as described in [6]. After transformation, the cells were plated on TAP medium containing 125 μg/ml spectinomycin (TAP + spec) and incubated under dim light. Colonies were picked approximately 2 weeks later and plated on TAP medium containing 50 μg/ml streptomycin (TAP + strep). Transformants were identified by their ability to grow both in the presence of spectinomycin and in the presence of streptomycin. The initial growth on TAP + spec (instead of TAP + spec and strep) medium appears to allow heteroplasmic transformants to establish more easily, and the subsequent growth on TAP + strep medium removes spectinomycin-resistant mutants that arise spontaneously during the transformation procedure. Isolation and analysis of DNA from the transformants were carried out as described in [7].

Total DNA from various *Chlamydomonas* species was isolated as described in [7]. Oligonucleotide primers used in polymerase chain reactions included S-78 with a sequence of 5'ATGGG(TA)CAAAAAGT(TA)CATCC(TA)(TC)T(ACGT)GG(ACGT)TT3', and S-79 with a sequence of 5'CA(AT)ACTTT(AT)AC(AT)CCAAT(ACGT)A(CT)(ACGT)CC(AG)TA3'. Degenerate sites are enclosed in parenthesis. DNA cloning, sequencing, and sequence analysis were carried out as previously described [8].

3. RESULTS

3.1. *Disruption of* C. reinhardtii *chloroplast orf712*

Gene disruption was carried out through chloroplast transformation, in order to determine whether the

Chlamydomonas chloroplast orf712 is functional. A transforming plasmid DNA (C-104) was constructed, in which *orf712* was disrupted by an AAD marker gene that encodes an aminoglycoside adenyl transferase (Fig. 1A). The AAD marker gene was a 1.9-kbp *Bam*HI-*Eco*RI fragment isolated from the plasmid pUC-atpX-AAD, which was constructed by Goldschmidt-Clermont to express resistance of spectinomycin and streptomycin in the chloroplast [9]. Upon transformation, the disrupted orf712 should replace the endogenous wild-type orf712 through homologous DNA recombination. As a result, the AAD marker gene will be integrated into the chloroplast genome, and will confer spectinomycin- and streptomycin-resistance to the transformed cell [6,9,10].

After transforming a wild-type *C. reinhardtii* strain (137c mt^+) with the C-104 plasmid DNA, six transformants were identified by their ability to grow both on TAP + spec medium and on TAP + strep medium. These transformants were subjected to further rounds of growth and screening in an attempt to isolate homoplasmic transformants. This is a common practice in transforming chloroplasts in this system, since there are approximately 80 copies of chloroplast DNA per cell. In each round of growth and screening, a single transformant from the previous round was plated and grown on TAP + spec medium, ten of the resulting colonies were subjected to DNA analysis, and the colony containing the highest proportion of the disrupted *orf712* was used for the next round of growth and screening. DNA analysis of transformants from the above process revealed the following (see Fig. 1B): (1) all the transformants were heteroplasmic mutants containing both the disrupted *orf712* and its wild-type copies; (2) the transformants persisted as heteroplasmic mutant after six rounds of growth and screening under selection pressure for the disrupted *orf712*; and (3) the heteroplasmic mutant stabilized at a level where approximately 70% of its *orf712* DNA copies were disrupted. This is in contrast to control transformations, in which the AAD marker gene was inserted at other parts of the chloroplast genome not affecting *orf712*, and homoplasmic transformants were obtained after just two rounds of growth and screening ([6], also Fig. 2 in this paper).

Growth phenotype of the *orf712* heteroplasmic mutant was analyzed and compared with untransformed wild-type cells and a control transformant (Fig. 2). Under mixotrophic (or photoheterotrophic) conditions (on TAP medium, under light), the mutant grew significantly slower than untransformed wild-type cells and the control transformant. Under phototrophic (or photoautotrophic) conditions (on HS medium, under light), the mutant grew extremely slow or not at all. As expected, untransformed wild-type cells did not grow in the presence of streptomycin, due to its sensitivity to this antibiotic. The control transformant grew under all

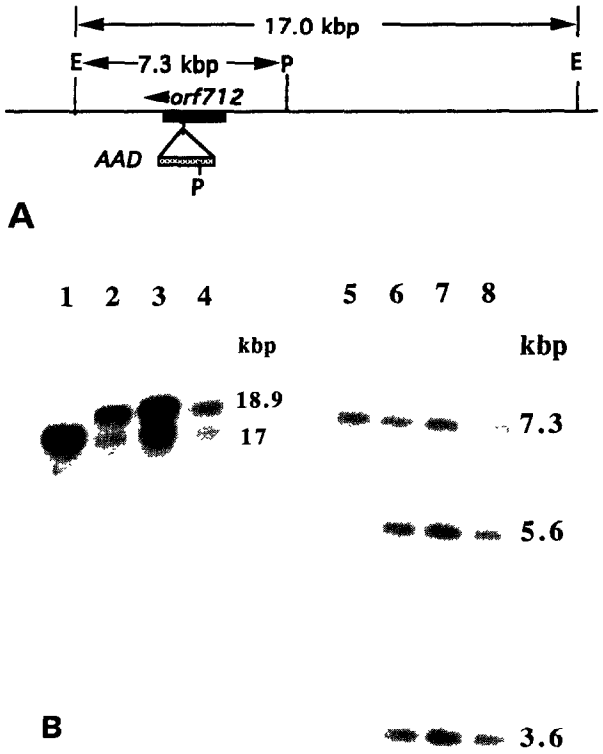


Fig. 1. Disruption of *C. reinhardtii orf712*. (A) Restriction map showing orf712 and its flanking regions. The letters E and P represent *Eco*RI and *Pst*I recognition sites, respectively. The transforming DNA (C-104) consists of the 2.1-kbp *orf712* coding sequence (black box), the 1.9-kbp AAD gene (gray box) inserted at a unique *Bgl*II site within *orf712*, and plasmid pUC118 as the cloning vector. (B) Analysis of DNA from transformants. Total cellular DNA was digested either with *Eco*RI alone (lanes 1–4) or with *Eco*RI and *Pst*I together (lanes 5–8), and the products were analyzed in Southern hybridizations using the *orf712* DNA as probe. Shown here are DNAs from untransformed wild-type cells (lanes 1 and 5), DNAs from the transformant after the 4th round (lanes 2 and 6), the 5th round (lanes 3 and 7), and the 6th round (lanes 4 and 8) of growth and screening. Wild-type DNA: a 17-kbp fragment in the *Eco*RI digest; a 7.3-kbp fragment in the *Eco*RI+*Pst*I digest. Mutant DNA (AAD inserted in *orf712*): a 18.9-kbp (i.e. 17+1.9 kbp) fragment in the *Eco*RI digest; two fragments of 5.6 and 3.6 kbp, respectively, in the *Eco*RI + *Pst*I digest.

conditions, because it contains an AAD gene that confers streptomycin-resistance without affecting any functional gene. The *orf712* heteroplasmic mutant grew slightly faster in the absence of streptomycin than it did in the presence of streptomycin. This was due to a progressive loss of the disrupted *orf712* gene in the absence of selection pressure to retain it, as was confirmed by DNA analysis (data not shown).

3.2. *Characterization of orf712 genes from other* Chlamydomonas *species*

Two oligonucleotides, S-78 and S79, were designed and synthesized, based on two stretches of protein sequences that are highly conserved between Orf712 of *C. reinhardtii* and Rps3 of other organisms. They were used as primers in a polymerase chain reaction (PCR)

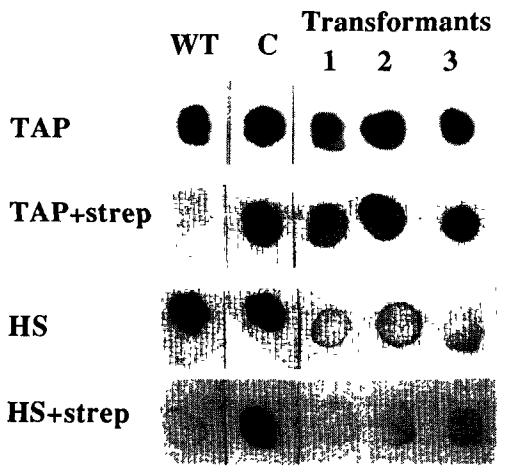


Fig. 2. Growth phenotype of the *orf712* heteroplasmic mutants. Equal numbers of cells were plated either on TAP medium for mixotrophic growth, or on HS medium for phototrophic growth, in the presence (+ strep) or absence of streptomycin, as indicated. Cells were photographed after 8 days of growth under light. They include untransformed wild-type cells (WT), a control transformant (C) in which the AAD marker gene was inserted at an intergenic site over 60 kbp away from *orf712*, and the *orf712* transformant after the 4th round (1), the 5th round (2), and the 6th round (3) of growth and screening. Darker patches indicate more growth.

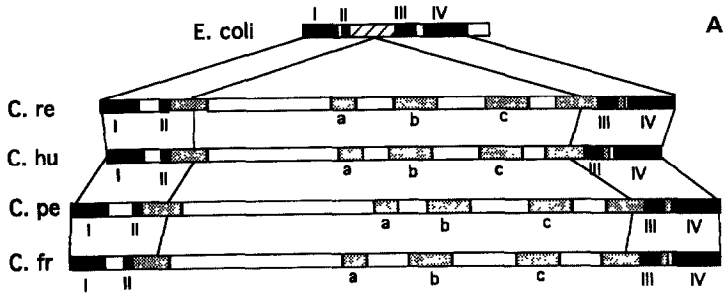
to amplify *orf712* DNA from *C. peterfii*, *C. frankii*, *C. pitschmannii*, *C. geitleri*, and *C. humicola*. These *Chlamydomonas* species represent all of the two or three major *Chlamydomonas* lineages [11]. From each of the *Chlamydomonas* species, a DNA fragment was specifically amplified that had a size similar to that of *C. reinhardtii orf712*. They were further identified as *orf712* by Southern hybridizations and by partial DNA sequencing (data not shown). The PCR-amplified *orf712* DNAs of *C. peterfii*, *C. frankii*, and *C. humicola*, were cloned, and their sequences determined completely.

Orf712 protein sequences predicted from the *orf712* DNA of the various *Chlamydomonas* species were compared among each other and with typical ribosomal protein Rps3 sequences from other organisms. The sources of Rps3 sequences were selected to represent distinct groups of organisms. At the overall structural level, the *Chlamydomonas* Orf712 protein has all of the four conserved sequence domains present in a Rps3 protein (Fig. 3A,B). The Orf712 proteins, however, contain a centrally located large stretch of sequence (termed IVS, for intervening sequence) that is absent from Rps3 proteins (Fig. 3A,C). The Orf712 proteins are around 700–800 amino acid residues in size, which are considerably larger than Rps3 proteins that have from 218 to 233 residues. This size difference, however, is due almost entirely to the presence in Orf712 proteins of the IVS that ranges in size from 457 amino acid residues in *C. peterfii* to 577 residues in *C. frankii*. The IVS contains three highly conserved sequence blocks that are internally located, in addition to two conserved blocks that are located at junctions with the upstream and downstream RPS3-like sequences.

4. DISCUSSION

Our gene disruption study clearly demonstrated that the *Chlamydomonas* chloroplast *orf712* is a functional gene, even with its unusual structural features and the lack of a detectable transcript of this gene. This conclusion is also supported by the strong conservation of this gene among the distantly related *Chlamydomonas* species. The failure to detect a transcript for this gene in Northern hybridizations by us and others [4] may be explained in two ways. First, an *orf712* transcript may exist at a very low level that is below the sensitivity of the Northern analysis but nevertheless sufficient for its function. In *Chlamydomonas* chloroplasts, gene regulation appears to be exerted largely at the translational level, with the protein synthesis rate not in proportion to the mRNA level. For example, transcript level for the ribosomal protein Rpl2 can be drastically reduced without affecting the synthesis rate of this protein [18]. Individual ribosomal protein genes in the same gene cluster produce very different transcript levels (Liu et al., unpublished data), although their protein products need to be made in equal amounts. An alternative explanation is that the *orf712* transcript may be heterogeneous in size and therefore does not migrate during electrophoresis as a single band detectable by Northern hybridization.

Although the size and location of the *orf712* gene product (Orf712) remain to be characterized, we have presented evidence suggesting that Orf712 is functionally equivalent to the ribosomal protein Rps3. First, all of the four conserved sequence domains of Rps3 are present in Orf712, suggesting that Orf712 has the structural base to function as Rps3. The Rps3-like domains in Orf712 are highly conserved among distantly related *Chlamydomonas* species, which further supports the suggestion of a Rps3-like function for the Orf712 protein. Second, Orf712 is functionally essential for cell growth, just as expected if it is a ribosomal protein. In *Chlamydomonas*, chloroplast ribosome function is essential for cell viability, although it can be reduced to a minimal level required for cell growth under heterotrophic or mixotrophic conditions [5,19,20]. This is presumably due to a necessity to make one or more chloroplast-encoded proteins that have essential cellular functions, as has been suggested for plants [21]. Third, the heteroplasmic *orf712* mutant has a growth phenotype similar to that of previously characterized mutants with decreased chloroplast ribosome functions. These mutants grow at or near wild-type rate under heterotrophic or mixotrophic conditions, but are unable to grow under phototrophic conditions [20,22]. This is not surprising, considering that chloroplast ribosomes synthe-



B

```
          10        20        30        40        50        60        70        80        90
Chlre   MGQKVHPLGFRVGITKKHQSQWFARFQKYAYSQSVFEDHMLRTTLVNLFSNLEKESALATKQSKNR------GATQPKAPKITQIKIERGLIPYEIGIQI
Chlpe   .........................N....A...V..F...K..MKM.PD.LALPLRGK.TQ.SSGGAGQPNQETDVL.....................
Chlhu   .........................N.NK.A...L..R...E..T...PQ.LN----PVLKK.K.------DDASAIV.R.............S......
Chlfr   .......F...............K.H..N.....L..RI..N..LKI.PE.LNPVL--K.VQ.R-------EDDTTIT...........F.........
Euggr   ..........L....S.S.F.YV--ERRH.ASF.K..IVI.NFMNKELLE---------------------------TL.SL.....IYEFS.QRNNT
Maize   ....IN.....L.T.QN.H.F...--.PKN..EGLQ..KKI.NCIK.YIQKNR.------.G.NRK------MESDSSSEV..H.E.QKEIDTTHVI.H.
Spiol   ....IN.....L.T.QS.Y.L..S--.PKN.AEGLQ..QKI.DCIK.YVQKNT.--------------------.SSGVEG.AR.E.QKRIDLTQVI.HM
Marpo   ....IN.....L...QN.R.Y...--N.KY-.KVFE..KKI.DCIELYVQKHI.-------------------NSSNYGG.ARVE.K.KTDLTQVE.YT
Cyapa   V...I..I....L....Q..R.C...--NPKQ.PTLLQ...LI.QYIEKNL..-----------------------AG.A..Y.Q.KADRT.LELRT
Mycca   .....S.NVL.L..VRDWENR.Y.--E.DQ.VKWLDQ.IKI..A.FK.LKD-------------------------AAVSK.D...TTKDLTLF.KT
Ecoli   .......N.I.L..V.PWN.T...--NTKEFADNLDS.FKV.QY.TKELAK-----------------------ASVSR.V...PAKSTRVT.HT
          10        20        30        40                                  50        60        70

            100       110   118-580   590       600       610       620       630       640       650
Chlre   HSNDCLSITKAIDNIKVSKDLVT  SQRKMYKANLALTPKI-SIKFFSVKTTNLLEKASTVADSIVDALEKRKAFRGVIKKAKEDLMLRSRVTRV---KV
Chlpe   .AEN.EQ.QA....L.GNGTDGG  N..............-..Q.....QD..F....V....V......C..............R....-------.G
Chlhu   .AEN.DL.KSS.E.L.INR...S  K..TI.TD.IS.....-....Y..NAKSVKA...V....V........G..K....D...E..---.TSG.---.G
Chlfr   .AEN.EL.KS...KL.IDR..AC  K.KIVS.E..K.....-C...Y...QK..EA...I.....I..........K...D...N.TSN...NGIKLV.G
Euggr   IVYIHVARPERVIGRDGQGLSRIRDILIDRMNYL.GK..R.ITC.VVG.TSP..--D.RLL...VRRE....TP.IRAM.TV-----..QAMKAGA---EG
Maize   GFPNL.KKKG..EE--LE...QK----------EVNSVNQRLN.AIEK..EPYR--QPNIL.EY.AFQ.KN.VS..KAM...----IELTKKADI---.G
Spiol   GFPKL.IENRPQGVEDLKINVQK----------E.NCVNRKLN.AITRIAKPYG--DPNIL.EF.AGQ.KS.VS..KAM...----IELTEQADT---.G
Marpo   GFPAL.VESRGQGIEQLKLNVQN----------I.SSEDRRLRMTLIEIAKPYG--EPKIL.KK.ALK..S.V...RTM...----IELAKKGNI---.G
Cyapa   ARPGVVVGRGGRGIEVLR.G.KD----------L.GEQKQK-R.NVIE..QIDA--E.ALIGEF.TQQ..R.V...RIVR..----IT.AQRRGI---EG
Mycca   ARPAIVLGQEGKNIE.IVLAVRK----------TVKNKKLKVNVRVIEI.SPDA--RSTL..RW.GEQISN.AS..T.Q.L.----IKKALKAGA---.G
Ecoli   ARPGIVIGKKGEDVE.LR.VVAD----------IAGVPAQK-N.A--E.RKPE.--D.KL.....TSQ..R.VM..RAM.R.----VQNAMRLGA---.G
          80        90                  100       110       120       130       140

          660       670       680       690       700       710
Chlre   VKIQVAGRLNGAEIARSEWVRAGRVPLQTLRANIDYAYRTANTIYGIIGVKVWIFKGYSKI      712
Chlpe   ..................................S....S...........                 >807
Chlhu   .....S.....................L..S.K..S...........                    >682
Chlfr   .....S..........T.....................S.K.............              >809
Euggr   I.V..S.....I....T..F.E.....H....D...FNDI.H....VL.I...VY.V           218
Maize   I...I....A.K....A.CIKK..L....I..K...CCYPIR....VL...I...VEEE         224
Spiol   IQ..I...ID.K....I..I.E.......I..K....CAY.VR....VL.I.I...M.EE        218
Marpo   I...I...........V..A.E.......I..R.N.C.YA.Q....VL.I.....QDEE         217
Cyapa   I...IS.............S.E.........E...S.KR.Q....VL.........EVIPGNPTEISE     219
Mycca   I.TA.S...G.V.M..T.GYLECS...S...N.....LYE.P.T..Q.......NH.EVFK.ERMNNSQIMAKPRTNKGGKR          233
Ecoli   I.VE.S...G......T..Y.E.....H....D...NTSE.H.T..V..........EILGGMAAVEQPEKPAAQPKKQQRKGRK      233
        150       160       170       180       190       200       210       220       230
```

C

```
                          120       130                   140       150
Chlre                   -----NLQKTRKYLFKAGTQLK-----------------NASMQVSDNLNVAEGENSTMLKT----------------------
Chlpe                   LRFFE......Q..S.VVSPITAYGDGKKKSKMDTVKILRQTLVNH.HTVSSLSADATNTVSKKTFIKSLSQKKTERSLNLDKKWT
Chlhu                   -----K.....R..QHLRIK.N-----------------DLNT-M..S.QKK.SNPTSEV.K----------------------
Chlfr                   -----H...A.Q...NL.LKT.TNVNLNKENSLEN----S.QLLGLPNG.PGK.N.I..NAAK--LSPVLKNKKKKGQKLKRV--

                                  160       170       180       190       200       210
Chlre   --------------------------KTSASGTSVLKKR-----TFKIQTQKPTKGKPVFKG-KRKQ--KKKLSKAVFMRL------KNIKRRPKKR----QTI
Chlpe   GSAGKAGAASGLDKKFQSTRGYAGAGANFAKADNQKGFRSGKNPKNT.TGKMSLRTNTTPV.RF.KR.SIRNRYFQLLSKG.FIRKNGNQ.M.KVSLK-GFIKLS
Chlhu   --------------------------DSLNNDKKKRFRP-----KKRLTK.QLIRRRL.L.RL..R.SIRRRHKQL.YK..FLAKKGNKVVQKLTMP----TFG
Chlfr   ------------KKVNIA--------SNFINKKNKKRFIS-----RP.LTK.EYK.K.L.QQSLRNR.IIRRWYRQFIAKG.LINKTG.K...KIATKKGKTCFG

        220       230       240       250       260       270       280                 290       300       310
Chlre   KKRYLNIISKGLLI-RKKGNLIIRNVKIKRKNKTARLTSRKSQPTLRSGSNLRDNLAPVK-SKMQRFN----------NRMSKKFANLFLTKLNKQFLVRLKAIM
Chlpe   P-EK.KYN.GMVSAGSRFNKKNGLAGQNLTTTSQ.K.NIKGARNQAF-KPG.LGQRPTTSRLGTSVLIKKLS------.KL.....T..IN....K...S.....
Chlhu   R.GFSGKNK.T.FKGSRTASA.D.EASSILFSRGYSAKGERP.RSVKTS.KSIAKAK.Q.YTSSNLLK----------L.IK...VS..INRM.NK..HA...SL
Chlfr   RFKKF.KVNSFSKTKNLTNK.KKKKQTPLSVAT.NFSENKTNKKSSFIKNPILNKQTKS.QK.GGL.VKNLGVKSAVYLKIKR..VT.Y.N.V..K..KN..EL.

                320       330       340       350       360       370       380       390       400       409
Chlre   KFW-----HNQNVTKAPLGYNKKWSLAKSYALIN-NLKAKYLAKDILSLGSLRVQKLRKLISILEKKSLVKMETLRKDFITFGTLSKTRAFGYYQMITFLKQLKE
Chlpe   .Y.-----.K..TNTI--------FANYNKKWTAWQDSMQ---.KQNLANPQSIT.WN...DV......Q.L.Y.....MN..AI....S.....L..L...V..
Chlhu   NET-----QKNQNSGVR---TQQ..KKWNF..VE----------kR..QKSLTQ.TNF.TL..E...K...S.........SI..AAS.R...IL.......Q
Chlfr   .Y.FFVFEEKNSDS.ININTSISSLTSAKINFAPFGYNKNWDI.KLRV.KNQPLA..K..FEV.....FI.LQA.KQYY.G..S....Q.YSFF...V......K

                          420       430       440       450       460       470       480       490
Chlre   LVTK--------------IK-KQTIANVTTKQELCCYNKLA------LNKTKIQNLIRAKSKQTKSITQKVVNNFVKLVDDNQAMANESRKIKWISYLKDLVNKHR
Chlpe   F.MRTTKLRLLFGRSTAQ.NQNKTTGSQQTLNKAYLA.AT------TPL...SGVTLQ..VT.ALQENIFNTIQANINNGGSN-NK....L.L....QNRIQTY.
Chlhu   F.S.-------------L.T..KFNYTLS.LNHSETQQST------GSGSALSDSVNF..ARLTEASNTIST.L--.ANKLNNID..C....F.E...QI.K...
Chlfr   .IKRRYYAVLQKYKNKNFVF..SLNTERGSLSSKFLQQTKNASGLET.LQ.SESG.NKSI..KTLANYSYIQTQKA.KKIVSNVDD.C....F.E...EI.K...

        500       510       520       530       540                 550       560       570       580
Chlre   TENIFYYLATIATARKDLNALKRYTKQHANFLFGVNVENAKENPNA----------------LLQRVTKTLTQYSKNPLV-NNDFENSEGLTKLQTAFLTQIE
Chlpe   ...L....SSM.N.S..FERT.....RTLRISYL.LRRMLSYRKKPRGGLYTINNKT------F.QEKIE.V..KSNQK-----V..DKN------..D...D.V.
Chlhu   S...YL......E.....RK..QF..R..S....L.LNQTS.-AS.----------------IQE.LKSV.ENS.TK-----P...KT------..DV..E...
Chlfr   ...L....S.L.E.....KK..KF.QN..E....LDLKL..TGNKEGSDYSENNQNSSDQMQLVVKNQ...I.E.AN.KE.SM.L.L.R.------..DS.IE...
```

Fig. 3. Protein sequence comparisons. Proteins compared include: *Chlamydomonas* chloroplast Orf712 from *C. reinhardtii* (Chlre), *C. peterfii* (Chlpe), *C. humicola* (Chlhu), and *C. frankii* (Chlfr); plastid Rps3 from *Euglena gracilis* [12] (Euggr, a euglenoid), *Zea mays* [13] (maize, a monocot plant), *Spinacia oleracea* [14] (Spiol, a dicot plant), *Marchantia polymorpha* [15] (Marpo, a lower plant), and *Cyanophora paradoxa* [16] (Cyapa, an organism containing plastid-like cyanelles); bacterial Rps3 from *Mycoplasma capricolum* [17] (Mycca, Gram-positive) and *E. coli* [3] (*E. coli*, Gram-negative). (A) Schematical comparison among Orf712 of *Chlamydomonas* and Rps3 of other organisms (represented by *E. coli*). Black boxes represent sequences conserved among all the Orf712 and Rps3 proteins, gray boxes represent sequences conserved only among Orf712 proteins, the slashed box in *E. coli* Rps3 represents sequences that are marginally conserved among all the Orf712 and Rps3 proteins, open boxes represent sequences that are not conserved. (B) Sequence comparison of Orf712 and Rps3 proteins. Dots represent residues that are identical to corresponding residues in the first sequence (Chlre). Dashes represent gaps introduced to maximize the sequence alignment. Residues are numbered for the first (Chlre) and the last (*E. coli*) sequences. The total number of residues in each sequence is given at the end of that sequence. Positions of the Orf712 intervening sequences (IVS), located between residues 118–580 in the Chlre sequence, are indicated. The sequences from Chlpe, Chlhu, and Chlfr are not complete at their C-termini. (C) Comparison of Orf712 intervening sequences. All symbols are the same as in (B). The Chlre sequence is numbered continuously from (B).

←

size a large number of chloroplast-encoded proteins that are required in photosynthesis (e.g. [15,23]). In *Chlamydomonas* mutants with reduced chloroplast ribosomal functions, the residual chloroplast protein synthesis activity appears to preferentially synthesize ribosomal proteins and possibly other essential proteins at the expense of photosynthetic proteins [24].

The *Chlamydomonas* chloroplast *orf712* gene, in contrast to conventional *rps3* genes of other chloroplasts and bacteria, contains a large intervening sequence (IVS) in the middle. Our results suggest that the *orf712* IVS is most likely translated continuously with its flanking *rps3*-like sequences, based on the following considerations. (1) The IVS does not have recognizable structural features associated with any of the known RNA introns, and is therefore unlikely to be spliced out at the RNA level. (2) The entire IVS is an open reading frame with a codon usage that is strongly biased in the same way as known protein-coding genes in this system, while the other two reading frames of the IVS DNA contain numerous translation termination codons because of the high AT-content of this DNA. (3) The IVS is preserved as an open reading frame in all the distantly related *Chlamydomonas* species examined, despite an extensive divergence of its primary sequence. (4) It is also unlikely for the *orf712* gene to be translated as three separate open reading frames corresponding to the N-terminal Rps3-like sequence domain (ND), the IVS, and the C-terminal Rps3-like sequence domain (CD), respectively, due to the lack of appropriately placed translation initiation and/or termination codons.

If indeed the *orf712* gene product is made as a single polypeptide, this would have important implications on the structure and function of this protein. Several scenarios can be suggested. The first and simplest scenario is that the polypeptide assembles and functions in the ribosome without further processing. The ND and CD may still fold together to form a conventional Rps3 protein, with the IVS looping out as spacer sequence of no obvious function. It is more difficult, however, to explain why parts of the IVS are strongly conserved among the distantly related *Chlamydomonas* species, and how the ribosome tolerates this long extra sequence. In a second scenario, the precursor polypeptide may be processed into three parts (ND, IVS, CD), with the ND and CD assembling and functioning in the ribosome as two polypeptides. There is no known example of a split Rps3 protein. In a third scenario, the IVS sequences may be a protein intron that is removed from the precursor protein through protein splicing, while the ND and CD are consequently rejoined to form a conventional RPS3 protein. There have been examples of protein splicing in a prokaryote, an archaea, and a eukaryote (see [25] for a review). A characterization of the *orf712* gene product(s) would be necessary to prove or disprove the above scenarios.

Acknowledgements: We thank Dr. Michel Goldschmidt-Clermont for the plasmid pUC-atpX-AAD, and Dr. Robert W. Lee for the *Chlamydomonas* cultures. This work was supported by an operating grant from the Medical Research Council of Canada to X.-Q.L. who is also a scholar in the Evolutionary Biology Program of the Canadian Institute for Advanced Research.

REFERENCES

[1] Schmidt, R.J., Richardson, C.B., Gillham, N.W. and Boynton, J.E. (1983) J. Cell Biol. 96, 1451–1463.
[2] Subramanian, A.R. (1993) Trends Biochem. Sci. 18, 177–181.
[3] Zurawski, G. and Zurawski S.M. (1985) Nucleic Acids Res. 13, 4521–4526.
[4] Fong, S.E. and Surzycki, S.J. (1992) Curr. Genet. 21, 527–530.
[5] Harris, E.H. (1989) The Chlamydomonas Source book: A Comprehensive Guide to Biology and Laboratory Use, Academic Press, San Deigo.
[6] Liu, X.-Q., Xu, H. and Huang, C. (1993) Plant Mol. Biol. (in press).
[7] Newman, S.M., Boynton, J.E., Gillham, N.W., Randolph-Anderson, B.L., Johnson, A.M. and Harris, E.H. (1990) Genetics 126, 875–888.
[8] Wang, S. and Liu, X.-Q. (1991) Proc. Natl. Acad. Sci. USA 88, 10783–10787.
[9] Goldschmidt-Clermont, M. (1991) Nucleic Acids Res. 19, 4083–4089.
[10] Takahashi, Y., Goldschmidt-Clermont, M., Soen, S.-Y., Franzén, L.G. and Rochaix, J.-D. (1991) EMBO J. 10, 2033–2040.
[11] Turmel, M., Gutell, R.R., Mercier, J.-P., Otis, C. and Lemieux, C. (1993) J. Mol. Biol. 232, 446–467.

[12] Christopher, D.A., Cushman, J.C., Price, C.A. and Hallick, R.B. (1988) Curr. Genet. 14, 275–286.
[13] McLaughlin, W.E. and Larrinua, I.M. (1987) Nucleic Acids Res. 15, 4689.
[14] Zhou, D.X., Quigley, F., Massenet, O. and Mache, R. (1989) Mol. Gen. Genet. 216, 439–445.
[15] 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.
[16] Michalowski, C.B., Pfanzagl, B., Löffelhardt, W. and Bohnert, H.J. (1990) Mol. Gen. Genet. 224, 222–231.
[17] Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F. and Osawa, S. (1987) Mol. Gen. Genet. 210, 314–322.
[18] Hosler, J.P., Wurtz, E.A., Harris, E.H., Gillham, N.W. and Boynton, J.E. (1989) Plant Physiol. 91, 648–655.
[19] Hanson, M. and Bogorad, L. (1978) J. Gen. Microbiol. 105, 253–262.
[20] Harris, E.H., Boynton, J.E. and Gillham, N.W. (1974) J. Cell Biol. 63, 160–179.
[21] Wolfe, K.H., Morden, C.W. and Palmer, J.D. (1992) Proc. Natl. Acad. Sci. USA 89, 10648–10652.
[22] Shepherd, H.S., Boynton, J.E. and Gillham, N.W. (1979) Proc. Natl. Acad. Sci. USA 76, 1353–1357.
[23] Hallick, R.B., Hong, L., Drager, R.G., Favreau, M.R., Monfort, A., Orsat, B., Spielmann, A. and Stutz, E. (1993) Nucleic Acids Res. 21, 3537–3544.
[24] Liu, X.-Q., Hosler, J.P., Boynton, J.E. and Gillham, N.W. (1989) Plant Mol. Biol. 12, 385–394.
[25] Shub, D.A. and Goodrich-Blair, H. (1992) Cell 71, 183–186.