Curr Genet (1992) 21:521–525

Current Genetics

Structure and expression of a plastid-encoded *groEL* homologous heat-shock gene in a thermophilic unicellular red alga

Udo Maid, Rolf Steinmüller, and Klaus Zetsche

Institut für Pflanzenphysiologie, Justus Liebig Universität, Heinrich-Buff-Ring 58, W-6300 Gießen, Federal Republic of Germany

Received November 25, 1991/January 7, 1992

Summary. A gene homologous to the *E. coli groEL* locus was identified on the plastid genome of the unicellular red alga *Cyanidium caldarium* strain 14-1-1 (synonym: *Galdieria sulphuraria*). The complete nucleotide sequence was determined and compared to bacterial- and nuclear-encoded counterparts of higher plants. At the amino-acid level the *C. caldarium* gene shows 70% homology to the corresponding gene of the cyanobacterium *Synechococcus* and 52% homology to nuclear-encoded counterparts of higher plants, respectively. Northern and Western blot experiments were used to investigate the dependence of the transcript- and protein-level on culture temperature and heat shock.

Key words: Chaperonin – Evolution – Heat shock – Plastid – Red algae – Rubisco-subunit binding protein

Introduction

Recent studies on chromophytic and rhodophytic algae have revealed a large body of striking differences concerning the organization of their plastid genomes compared to those of chlorophytic algae and higher plants, respectively. While not pretending to resolve the issue of whether the photosynthetic organells are descendants of single or multiple endosymbiosis events, these data point unmistakably to a long period of separated evolution of the different algal plastids. One has, therefore, to consider the meaning of the differences as well as of the similarities. All types of plastids investigated so far possess genes for the ribosomal RNAs, as well as a set of tRNAs, and several essential proteins of the photosynthetic apparatus (Palmer 1985; Douglas 1988; Markowicz et al. 1988; Delaney and Cattolico 1989; Hwang and Tabita 1989; Maid and Zetsche 1990, 1991; Winhauer et al. 1991). Rhodophytes show plastid genomes which are somewhat larger than the majority of chlorophyll a+b plastid genomes (Reith and Cattolico 1986; Shivji 1991). Together with the lack of introns within protein-coding genes this seems to enable the plastid genomes of Rhodophytes to encode more genes than their chlorophytic counterparts: rhodophytic plastids, like those of Chromophytes, encode for the gene of the small Rubisco subunit (Douglas and Durnford 1989; Hwang and Tabita 1989; Valentin and Zetsche 1989, 1990; Kostrzewa et al. 1990). Phycobiliprotein genes, as well as as the *secA* gene and the *ompR* gene, have been detected on the plastid genomes of red algae (Valentin et al., in preparation; Kessler et al., in press). These features may be considered as more primitive attributes and as relics of the former cyanobacterium-like endosymbiont. They also may indicate a long period since the separation of the evolutionary pathways of chloroplasts and rhodophytic and chromophytic plastids, respectively.

In our laboratory we study the plastid genome organization of rhodophytic and chromophytic algae in relation to their coding capacity, the occurrence of additional genes compared to chlorophytic plastids, and any unusual gene arrangements, qualities which may all be useful as taxonomic tools. In this paper we present the nucleotide sequence of a red algal plastid-encoded gene homologuous to the *E. coli groEL* gene. This protein, belonging to the chaperonin group (reviewed by Roy 1989; Gatenby and Ellis 1990; Ellis 1990), has its nuclear-encoded counterparts in the Rubisco subunit-binding protein of higher plants (Hemmingsen et al. 1988). In the case of *Cyanidium caldarium* the gene product is of double interest. First, both Rubisco subunit genes, *rbcL* and *rbcS*, are co-transcribed from the plastid genome as in other red algae (Valentin and K. Zetsche 1989, 1990); hence, together with the *C. caldarium groEL* gene, there may be the opportunity for expression-experiments to elucidate the *rbcL-rbcS/groEL* interactions. Second, the thermophilic eucaryote *C. caldarium* may provide insight into the plastid heat-shock response system.

Offprint requests to: U. Maid and K. Zetsche

Materials and methods

Strain and culture conditions. C. caldarium Geitler, strain 14-1-1 was isolated by Allen; we used strain 107.79 from the collection of algae maintained by the Institute of Plant Physiology of the University of Göttingen. This strain is identical with *Galdiera sulphuraria* (Merola et al. 1981). Cells were grown axenically under constant light (7 500 lux) in 1 or 5 litre flasks as described (Steinmüller et al. 1983) at 45 °C. Cells were heat-shocked at 60 °C for 40 min with RNA isolation following immediately. For protein extraction cells were heat-shocked for 1 and 2 h, respectively, at 60 °C. Cells were harvested by centrifugation for 10 min at 5 000 g.

Isolation of rhodoplast DNA. Was as described by Maid et al. (1990).

Mapping and cloning of rhodoplast DNA fragments. Restriction endonucleases and T4-DNA Ligase, purchased from Boehringer Mannheim Biochemicals FRG and Gibco-BRL FRG, were used according to the manufacturer's recommendations. The cloning vector was pUC18. Starting with the rhodoplast rRNA operon (Maid et al. 1990; Maid and Zetsche 1991) overlapping plastid DNA fragments were identified from *Eco*RI, *Bgl*II, *Hin*dII and *Hin*dIII libraries by colony hybridization.

Labelling of DNA probes. Hexanucleotide primers (Boehringer) were used to label DNA fragments in low melting point agarose slices (FMC) either with digoxygenin-UTP for Southern hybridizations or with ^{32}P-α-dATP for Northern hybridizations.

Northern analysis. Total cellular RNA was isolated as described (Ausubel et al. 1987).

Protein techniques. We obtained antibodies raised against a LacZ(alpha peptide)-*groEL*(*E. coli*) fusion protein from R. Webb (Purdue University, West Lafayette, Indiana) and against the *E. coli groEL* protein from G. Lorimer (Molecular Biology Division, Central Research and Development Department, E.I. DuPont de Nemours and Company, FRG). Polyacrylamid gel electrophoresis and Western blotting were done as described (Towbin et al. 1979) and were developed with alkaline phosphatase-conjugated secondary antibodies.

Sequence analysis. Subclones were directly sequenced in plasmid pUC 18 with the T7 polymerase kit (Pharmacia). Sequence data were computed with the "Kroeger-Menu" (Kröger and Kröger-Block 1984) and the PROSIS software (Pharmacia) using the EMBL data library and the CLUSTAL-program (Higgins and Sharp 1988).

Results and discussion

The *C. caldarium groEL* gene was located about 5 kb upstream of the 16S rRNA gene and the inverted repeat border, respectively. It is separated from the 16S rRNA gene by the *psbD*/*psbC* operon, the *rps*16 gene (in preparation), and a gene homologous to the *E. coli ompR* (Kessler et al. 1992) which is transcribed from the opposite strand. Figure 1 shows a physical map of the identified genes. Due to the small amount of enriched plastid DNA of *C. caldarium* we have not yet established a physical map of the total rhodoplast genome, so there is no possibility of assigning this gene to either the small or the large single-copy region. Higher plant plastids, and those of the Chromophytes *Olisthodiscus luteus* and *Cryptomonas* Φ, show the inverted repeats oriented with the 23*s* rRNA genes towards the small single-copy region,

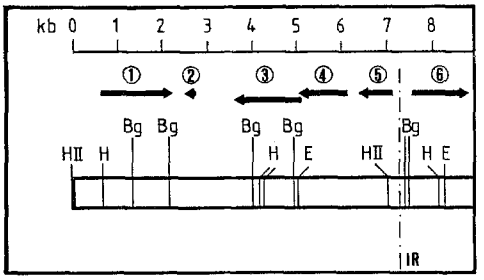

Fig. 1. Physical map of one plastid single-copy area upstream to the 16*s* rDNA. *Arrows* represent the sizes and localizations of the following genes: *1*, *groEL*; *2*, *rps16*; *3*, *psbC*; *4*, *psbD*; *5*, *ompR*; *6*, 16S rRNA (2–4, submitted; 5, Kessler et al. 1992; 6, Maid and Zetsche 1990). Abbreviations: *BG*, *Bgl*II; *E*, *Eco*RI; *H*, *Hin*dIII; *HII*, *Hin*dII; *IR*, inverted repeat; *kb*, kilobases

whereas the inverted repeat orientation is reversed on the cyanelle genome of *Cyanophora paradoxa* (Bohnert et al. 1985) and the plastid genome of the red alga *Porphyra yezoensis*. However, the existing physical maps of red algal plastomes do not agree with our findings. In the case of *Griffithsia pacifica* (Li and Cattolico 1987) no inverted repeat could be detected and the map of *P. yezoensis* shows hybridization signals to the *ppcBA* gene upstream of one 16S rRNA gene (Shivji 1991). In our laboratory we found structural differences within and around the ribosomal operons of the unicellular red algae *C. caldarium* and *Porhyridium aerugineum*, as well as the multicellular red alga *Antithamnion* sp. (in preparation), so the plastid genomes of red algae seem to exhibit large structural differences. According to the chaperon concept (for a review see Gatenby and Ellis 1990) proteins which are involved in the correct folding of other polypeptides and the correct assembly of multimeric proteins, respectively, without being a part of the resulting structure, are called chaperonines. The *C. caldarium* chaperonin is the first one described to be encoded on the plastid genome. Figure 2 shows the nucleotide sequence with the deduced amino-acid sequence, Fig. 3 an alignment with corresponding sequences from bacteria and a higher plant. The amino-acid homologies of the compared gene products to each other are listed in Table 1.

Western blotting experiments with antibodies against the *E. coli* 65 kDa Chaperonin (obtained from G. Lorimer, see Materials and methods) detect two protein bands of nearly equal size (Fig. 4) which is in good agreement with the molecular weight deduced from the amino-acid sequence of 57.8 kDa. Two forms of the nucleus-encoded higher plant Rubisco-binding protein with an Mr of 61 000 (α) and 60 000 (β) have been described; nevertheless the two Western-blot signals in the case of *C. caldarium* may point to different *groEL*-protein processing products as only one *groEL* gene could be detected by homologous Southern hybridization against total *C. caldarium* DNA (data not shown). However, there are considerable cross reactions of the antibody with other *E. coli* proteins. Therefore, further studies with antibodies against a cyanobacterial *groEL* protein, together with expression-experiments, may help to characterize the *C. caldarium groEL* homologue.

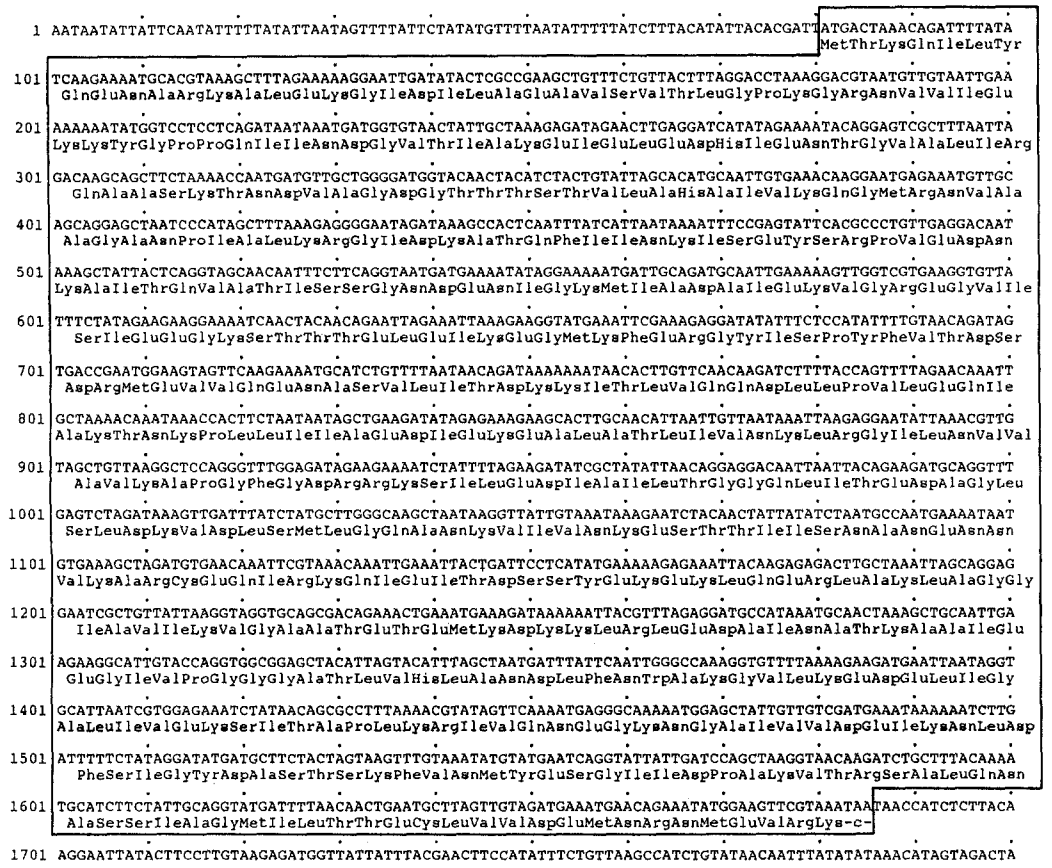

Fig. 2. Nucleotide sequence and deduced amino-acid sequence of te *C. caldarium groEL* gene

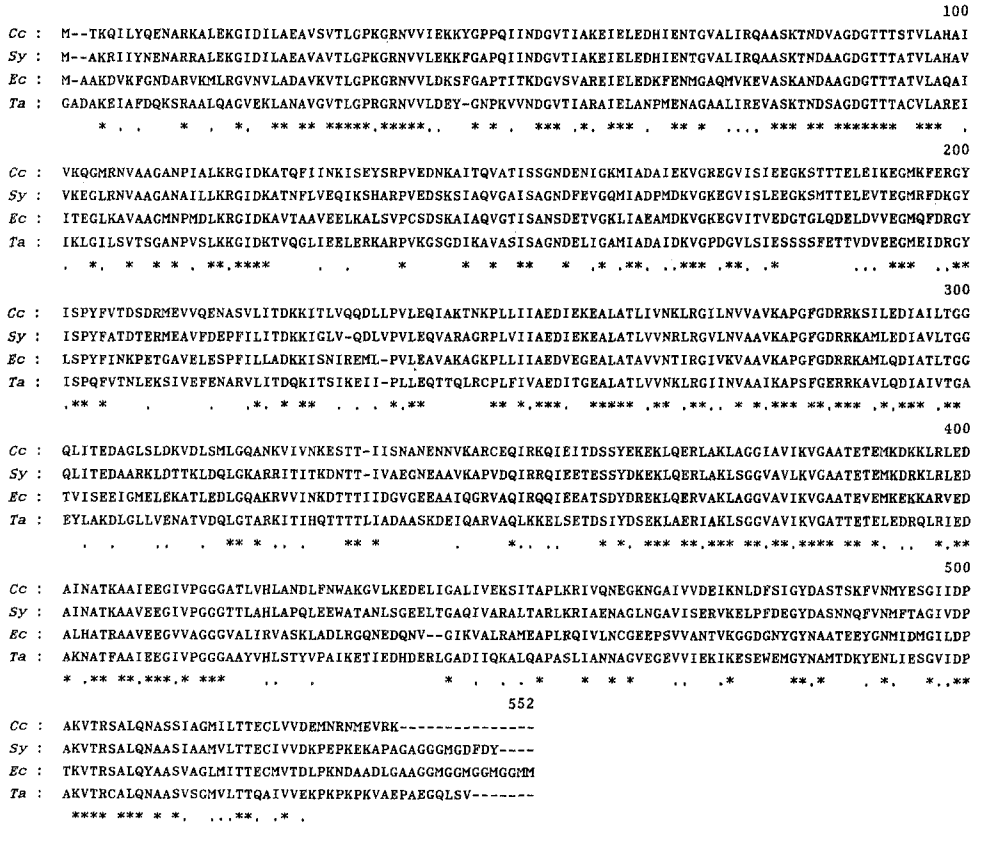

Fig. 3. Amino-acid alignment of *groEL*-homologous proteins from *Cyanidium caldarium* (*Cc*), *Synechococcus* sp. Strain PCC 7942 (*Sy*), *Escherichia coli* (*Ec*) and *Triticum aestivum* (*Ta*). The latter sequence is deduced from a c-DNA clone from the nuclear-encoded α-subunit gene. For references see Table 1. (*) represents exact matches and (.) conservative matches across all positions

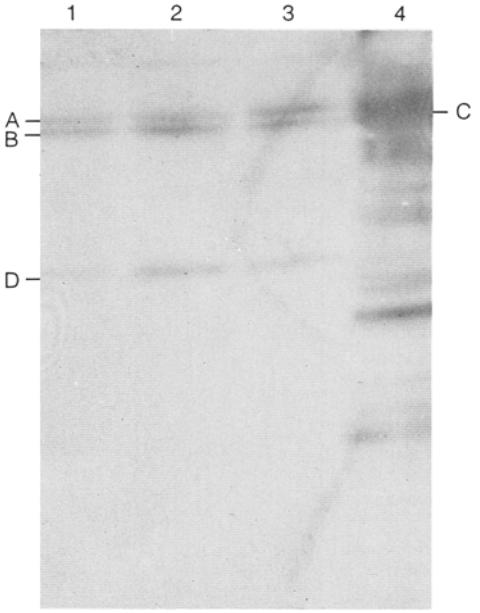

Fig. 4. Western blot of *C. caldarium* proteins. *groE*-related proteins were detected using polyclonal antibodies against the *E. coli groEL*-protein. *Lane 1*, *C. caldarium* grown at 45 °C, than heat-shocked at 60 °C for 2 h. *Lane 2*, cells grown at 45 °C, than heat-shocked at 60 °C for 1 h. *Lane 3*, cells grown at 45 °C, no heat-shock. *Lane 4*, *E. coli* grown at 37 °C, protein extraction. *A+B*, *C. caldarium* groEL-protein; *C*, *E. coli groEL*-protein. Additional signals (e.g., *D*) may be due to cross reactions of the polyclonal antibody with conserved epitopes of other heat-shock related proteins

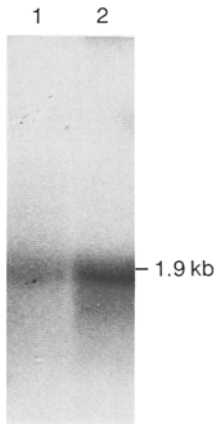

Fig. 5. Northern blot of *C. caldarium* total cellular RNA hybridized against a 0.6 kb ^{32}P-labeled plastid-DNA probe from the central groEL-coding region. Equal amounts of RNA were loaded on each lane. *Lane 1*, cells grown at 45 °C. *Lane 2*, Cells grown at 45 °C, than heat-shocked for 40 min at 60 °C with RNA isolation following immediately

There is a high degree of conservation between the *E. coli groEL* gene product, which is one of the main proteins of the bacterial heat-shock response system (Hemmingsen et al. 1988), the 65 kDa protein gene of *Mycobacterium tuberculosis* (Shinnick 1987) and *M. leprae* (Mehra 1986), and the Rubisco subunit-binding proteins of higher plants (about 50 to 65% at amino-acid level, see Table 1). The latter are encoded in the nuclear DNA as larger precursors and are transported into the chloroplasts (Hemmingson et al. 1988). In this organelle they manage the correct assembly of the nuclear-encoded small Rubisco subunits and the plastid-encoded large subunits. For this reason *C. caldarium* is an organism of dual interest. On one hand it is an extremely thermophilic unicellular eucaryote (Allen 1959); on the other hand, like all red algae so far investigated, both of the Rubisco subunits are encoded on the plastid genome. The highest homology (70%) is between the deduced *C. caldarium* protein and its counterpart in Cyanobacterium *Synechococcus* sp. strain PCC 7942 (Webb et al. 1990). A close relationship between red algal plastids and cyanobacteria was also found in the case of the *psbA* gene and the ribosomal RNA operon (Maid et al. 1990; Winhauer et al. 1991; Maid and Zetsche 1991). Further studies are required to determine the distribution of the *groEL* gene throughout the diverse algal taxa.

Table 1. Identical amino-acid residues in percent between *groEL*-homologous proteins of the following organisms: *C. cal*, *Cyanidium caldarium* (plastid-encoded); *E. col*, *Escherichia coli* (Hemmingsen et al. 1988); *M. tub*, *Mycobacterium tuberculosis* (Shinnick 1987); *T. aes*, *Triticum aestivum* (cDNA-Fragment of nucleus-encoded α-subunit gene, Hemmingsen et al. 1988); *Syn*, *Synechococcus* sp. Strain PCC 7942 (Webb et al. 1990)

Species	Syn	E. col	M. tub	T. aes
C. cal	70	52	53	52
Syn		55	60	52
E. col			59	47
M. tub				49

Whereas the *E. coli* and cyanobacterial *groE* loci consist of two components, *groES* coding for a 15 kDa protein and *groEL* coding for a protein of about 65 kDa, no *groES* homologue has yet been detected in the closer upstream area (700 bp) of the red algal *groEL* gene. The same is true for the *Mycobacterium groEL* gene and no eukaryotic *groES* homologue has been found to-date (Hemmingsen et al. 1988). The transcript size of about 1.9 kb (Fig. 5) points to a monocistronic transcription of the *C. caldarium groEL* gene. As the existence of a functional *groEL* gene copy in the nucleus could not be excluded, we hybridized total cellular DNA against a plastid DNA fragment containing *groEL*. Only the fragments corresponding to the plastid area of Fig. 1 could be detected (data not shown). For this reason, and because of the high degree of conservation compared to other *groEL*-genes, the plastid gene seems to be functional and is not a pseudogene.

The Northern blot analyses with total RNA of *C. caldarium* show higher amounts of the *groEL* transcript when the cells were heat-shocked at 60 °C for 40 min in contrast to cells grown at 45 °C (Fig. 5). The protein level, however, seems to remain almost constant (Fig. 4), indicating transcriptional as well as translational regulation of *groEL* expression. Further investigations in our laboratory will focus on the plastid stress-response of the thermophilic *C. caldarium* and its mode of regulation.

References

Allen MB (1959) Arch Microbiol 32:270–277

Ausubel MB, Brenz R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) In: Ausubel FM (ed) Current protocols in molecular biology. Greene Publishing Associate and Wiley-Interscience, John Wiley and Sons, New York, chapter 4.3–4.4

Bohnert HJ, Michalowski C, Bevacqua S, Mucke H, Löffelhardt W (1985) Mol Gen Genet 201:565–574

Delaney TP, Cattolico RA (1989) Curr Genet 15:221–229

Douglas SE (1988) Curr Genet 14:591–598

Douglas SE, Durnford DG (1989) Plant Mol Biol 13:13–20

Ellis RJ (1990) Science 250:954–959

Gatenby AA, Ellis RJ (1990) Annu Rev Cell Biol 6:125–149

Hemmingsen SM, Woolford C, van der View SM, Tilly K, Dennis DT, Georgopoulos CP, Hendrix RW, Ellis RJ (1988) Nature 333:330–334

Higgins DG, Sharp PM (1988) Gene 73:237–244

Hwang SH, Tabita FR (1989) Plant Mol Biol 13:69–79

Kessler U, Maid U, Zetsche K (1992) Plant Mol Biol (in press)

Kostrzewa M, Valentin K, Maid U, Radetzky R, Zetsche K (1990) Curr Genet 18:465–469

Kröger M, Kröger-Block A (1984) Nucleic Acids Res 12:113–123

Li N, Cattolico RA (1987) Mol Gen Genet 209:343–351

Maid U, Zetsche K (1990) Nucleic Acids Res 18:3996

Maid U, Zetsche K (1991) Plant Mol Biol 16:537–546

Maid U, Valentin K, Zetsche K (1990) Curr Genet 17:255–259

Markowicz Y, Loiseaux-de Goer S, Mache R (1988) Curr Genet 14:599–608

Mehra V, Sweetser D, Young RA (1986) Proc Natl Acad Sci USA 83:7013–7017

Merola A, Castaldo R, De Luca P, Gambardella R, Musacchio A, Taddei R (1981) Giornale Bot Ital 115:189–195

Palmer JD (1985) Annu Rev Genet 19:325–354

Reith M, Cattolico RA (1986) Proc Natl Acad Sci USA 83:8599–8603

Roy H (1989) Plant Cell 1:1035–1042

Shinnick TM (1987) J Bacteriol 169:1080–1088

Shivji MS (1991) Curr Genet 19:49–54

Steinmüller K, Kaling M, Zetsche K (1983) Planta 159:308–313

Towbin H, Staehelin T, Gordon J (1979) Proc Natl Acad Sci USA 76:4350–4354

Valentin K, Zetsche K (1989) Curr Genet 16:203–209

Valentin K, Zetsche K (1990) Plant Mol Biol 15:575–584

Winhauer T, Jäger S, Valentin K, Zetsche K (1991) Curr Genet 20:177–180

Webb R, Reddy KJ, Sherman LA (1990) J Bacteriol 172:5079–5088

Communicated by H. Kössel

Note added in proof

Further analysis of the *C. caldarium* plastid genome revealed the small singlecopy region located between the 23S rRNA genes of the inverted repeat. Hence the *groEL* gene is located within the large singlecopy region