REVIEW ARTICLE
doi: 10.2306/scienceasia1513-1874.2022.155

ScienceAsia 49 (2023): **141–154**

Chaperonin paralogues in cyanobacteria: Their non-classical nature

Hitoshi Nakamoto[a,*], **Jittisak Senachak**[b], **Apiradee Hongsthong**[b]

[a] Department of Biochemistry and Molecular Biology, Graduate School of Science and Engineering, Saitama University, Saitama 338-8570, Japan

[b] Biosciences and System Biology Team, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency at King Mongkut's University of Technology Thonburi, Bangkok 10150 Thailand

*Corresponding author, e-mail: nakamoto@mail.saitama-u.ac.jp

Received 21 Oct 2022, Accepted 8 Nov 2022
Available online 16 Nov 2022

ABSTRACT: Chaperonin (GroEL, Hsp60) is a molecular chaperone involved in maintaining cellular protein homeostasis. It interacts with unfolded and misfolded proteins to assist in their folding. In this article, structures, functions, regulation, and significance of chaperonin in cyanobacteria will be reviewed. There are multiple kinds of *groEL* genes in cyanobacteria in contrast to the *Escherichia coli* chaperonin paradigm. The cyanobacterial *groEL1* gene forms an operon with the single *groES* gene, similar to the *E. coli groESL* operon. In contrast, the cyanobacterial *groEL2* gene is monocistronic. The regulation of expression and function of the *groEL1* and *groEL2* genes are mutually distinct. Transcription of the *groESL1* operon and *groEL2* gene is induced not only by heat but also by light. Unlike the *E. coli groESL* operon, the expression of the cyanobacterial *groESL1* operon and *groEL2* is not controlled by the transcription factor sigma32. *Cis*-regulatory elements such as K-box and CIRCE regulate it positively and/or negatively. Combinations of the elements in *groESL1* and *groEL2* are evolutionarily diversified. Functional studies suggested that GroEL1 is equivalent to *E. coli* GroEL, which is essential, whereas GroEL2 is nonessential but plays an important role under stress. The absence of GroEL2 affects proteome and phosphoproteome under stress conditions. Moreover, GroEL1 and GroEL2 are structurally different. We propose that the *groEL2* gene is an outcome of neofunctionalization after *groESL* operon duplication. The *groEL1* gene retains the original function essential for cellular activities under both normal and stress conditions, whereas *groEL2* acquires a novel, beneficial function required under stress.

KEYWORDS: heat shock protein, molecular chaperone, GroEL, neofunctionalization, cyanobacteria

INTRODUCTION ON MOLECULAR CHAPERONES

Molecular chaperones are involved in a wide variety of cellular processes including regulation of gene expression, *de novo* folding of proteins that are newly synthesized in ribosomes, assembly and disassembly of multimeric proteins and protein complexes, transport of proteins into mitochondria and chloroplasts, protein degradation in the ubiquitin-proteasome pathway and autophagy [1, 2].

Under normal and especially under stress conditions such as heat shock, proteins unfold, misfold, and aggregate. Molecular chaperones interact with unfolded, misfolded, and aggregated proteins and assist in the (re)folding of unfolded and misfolded polypeptides. They prevent aggregation and reactivate (disaggregate) aggregated proteins [1, 2].

These evolutionarily conserved molecular chaperones are classified into several families according to their molecular mass and/or protein structure/function. These families include the small heat shock proteins (sHsp), Hsp60, Hsp70, Hsp90, and Hsp100 [2]. Except for sHsp, they exhibit ATP-binding/hydrolysing activity, which is required for facilitating protein folding or protein disaggregation, in contrast to spontaneous protein folding. It is thought that a protein forms its three-dimensional structure that is defined by its amino acid sequence spontaneously [3]. This principle, or Anfinsen's dogma, is true when the protein is present in a dilute protein solution. However, total macromolecular concentrations in the cellular milieu are 0.3 to 0.4 g/ml [4]. Thus, macromolecular species occupy a significant fraction of the volume of the cellular medium, and such media are referred to as "crowded" [5]. Crowding affects the association equilibria of macromolecules. Macromolecular crowding in a cell is thought to increase the tendency of proteins to aggregate [6], which necessitates molecular chaperones in a cell as described below.

The GroEL/CHAPERONIN 60/Hsp60 FAMILY

GroEL, present in the bacterial cytosol, is a member of the Hsp60 family. It is also called chaperonin 60 and is classified as a Type I chaperonin. Chaperonin 60 in chloroplasts and Hsp60 in mitochondria are also Type I chaperonin. Type II chaperonin is found in the cytosol of eukaryotes and archaea [7, 8]. Both Type I and Type II form a barrel-like structure consisting of two rings arranged back to back. Each ring contains a central cavity in which a non-native protein is isolated for folding into the native state. Type I chaperonins are homo-oligomers with seven subunits per ring, whereas type II chaperonins are homo or hetero-oligomers with

eight or nine subunits per ring. Type I requires a co-chaperonin GroES but Type II has a protrusion extending from the apical domain of each subunit which acts as a built-in co-chaperonin [7, 8].

E. coli GroEL

The *E. coli* GroEL has been extensively studied, establishing a foundation for the development of chaperone research. GroEL, which is an essential protein [9], is a major heat shock protein (Hsp) of *E. coli*. It is heat-induced with other Hsps, including GroES, DnaJ, GrpE, DnaK, HtpG and ClpB [10]. These Hsps function as chaperones and co-chaperones. Their induction/expression is turned on or off in response to heat or other stress signals. The transcription factor sigma 32 (σ^{32}) is involved in this induction. σ^{32}, one of the sigma factors or a subunit of RNA polymerase, recognizes the heat shock promoters of genes encoding various Hsps. Upon heat shock, the synthesis and stabilization of σ^{32} are enhanced [11], leading to a rapid and transient increase in the cellular σ^{32} level. This σ^{32} increase results in an increase in the transcription and thus translation of heat shock genes. Among Hsps, there are proteases as well as molecular chaperones. These Hsps form a protein quality controlling system that prevents protein denaturation, disaggregates protein aggregates, and removes/digests denatured proteins. When the cellular protein quality is restored, molecular chaperones and proteases are freed from non-native/denatured proteins (their substrates). Then, they interact with σ^{32}, which is one of the substrates of these chaperones and proteases, resulting in its inactivation and degradation. Thus, these molecular chaperones and proteases exert negative feedback on the heat shock response [11].

The *E. coli* GroEL is a homotetradecamer of 57 kDa subunits made of two nearly sevenfold rotationally symmetrical rings stacked back-to-back, as shown in Fig. 1. Inside each ring, there is a cavity. The opening of the central cavity binds a non-native polypeptide and GroES. GroES, the ∼10 kDa co-chaperonin for GroEL, forms a homoheptameric 'dome'. GroES binding releases a bound polypeptide from GroEL and closes the cavity of the GroEL ring to generate a folding chamber. The folding chamber quarantines the non-native polypeptide so that it does not interact with other non-native polypeptides. This process is facilitated by ATP binding/hydrolysis to GroEL. The chamber can accommodate a protein up to ∼60 kDa in size [7, 8].

The captured non-native protein folds in the folding chamber during the time when GroEL hydrolyses its bound ATP to ADP (∼10 s). After hydrolysis of ATP in the GroES-bound ring (*cis* ring), ATP binds to the opposite ring (*trans* ring), which triggers discharge of the protein, ADP, and GroES from the *cis* ring (Fig. 2). If the released substrate is not correctly folded it can rebind for further cycles of chaperoning [7, 8].

The isolation of a non-native protein in the GroEL ring is important for the protein to avoid interaction with other non-native proteins in a cell. When a non-native protein is isolated individually in the GroEL cavity, whose inner wall is hydrophilic, it can escape from aggregating environments due to macromolecular crowding inside a cell. Non-native proteins are thermodynamically more prone to aggregate than to refold back to their native structure [1], and macromolecular crowding in a cell increases this tendency [6]. Molecular chaperones, including GroEL, interact with aggregation-prone proteins and lower the risk of protein aggregation.

MULTIPLE GroELs IN CYANOBACTERIA

groEL1 and *groEL2* genes in cyanobacterial genomes

The *groEL* gene in *E. coli* forms an operon with the *groES* gene, the *groESL* operon. Thus, the two genes are transcribed together under the control of the heat shock promoter and the sigma factor σ^{32} described above. It is reasonable for these genes coding for chaperonin and co-chaperonin to be clustered and regulated together because the co-chaperonin GroES cooperates with GroEL (Fig. 2). The *groESL* (designated as *groESL1*) operon is also conserved in cyanobacteria. Interestingly, some cyanobacteria, such as *Chlorogloeopsis fritschii* PCC 6912, contain multiple *groESL* operons. In contrast to *E. coli*, cyanobacteria have a *groEL* (designated as *groEL2*) gene that is not associated with a *groES* gene [12, 13]. Our transcriptional analysis indicates that the *groEL2* gene is not co-transcribed with any other genes and thus monocistronic [13]. The amino acid sequences of the translation products of the two *groEL* genes are ∼60% identical. Amino acid sequence comparison shows that *E. coli* GroEL is equally similar to GroEL1 and GroEL2. *E. coli* GroEL, GroEL1 and GroEL2 are all acidic proteins of ∼58 kDa in size. Interestingly, chloroplasts of algae such as *Chlamydomonas reinhardtii* and higher plants also possess multiple chaperonin 60s [14–16].

Among 141 cyanobacterial genomes, 115 genomes encode a single *groESL* operon and a monocistronic *groEL* (without *groES* closely located) [17]. Hereafter, a *groEL* gene with the operonic gene structure and a monocistronic *groEL* are called as *groEL1* and *groEL2*, respectively. Most of the other cyanobacterial genomes encode two *groESL1* operons and a single *groEL2* gene. For example, *C. fritschii* PCC 6912, as well as other filamentous-type cyanobacteria, has two *groESL* operons (*groESL1*, *groESL1.2*) and a single *groEL2*. A phylogenetic tree constructed from all cyanobacterial *groEL* DNA sequences reveals two main clades that correspond to the *groESL1* operon and *groEL2*, indicating that the duplication of *groESL1* and *groEL2* is ancient and occurred at the base of the cyanobacterial

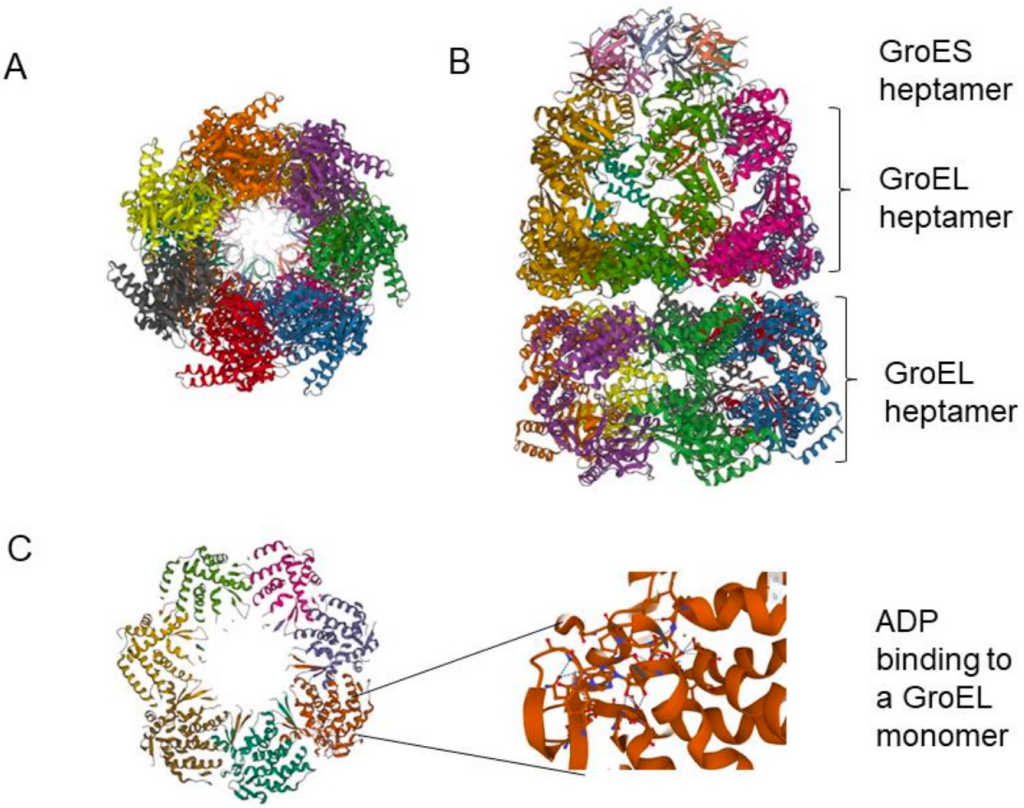

Fig. 1 Structure of the *E. coli* GroEL heptamer ring, the GroEL-GroES-ADP complex, and a GroEL monomer bound by ADP in the GroEL heptamer. (A) GroEL heptamer ring viewed from above down the axis of sevenfold symmetry. (B) Double rings of the GroEL heptamer capped with the GroES heptamer viewed from the side of the GroES-GroEL-ADP complex. (C) Illustration of ADP binding to a GroEL monomer. Constructed from PDBID: 1SVT [55].

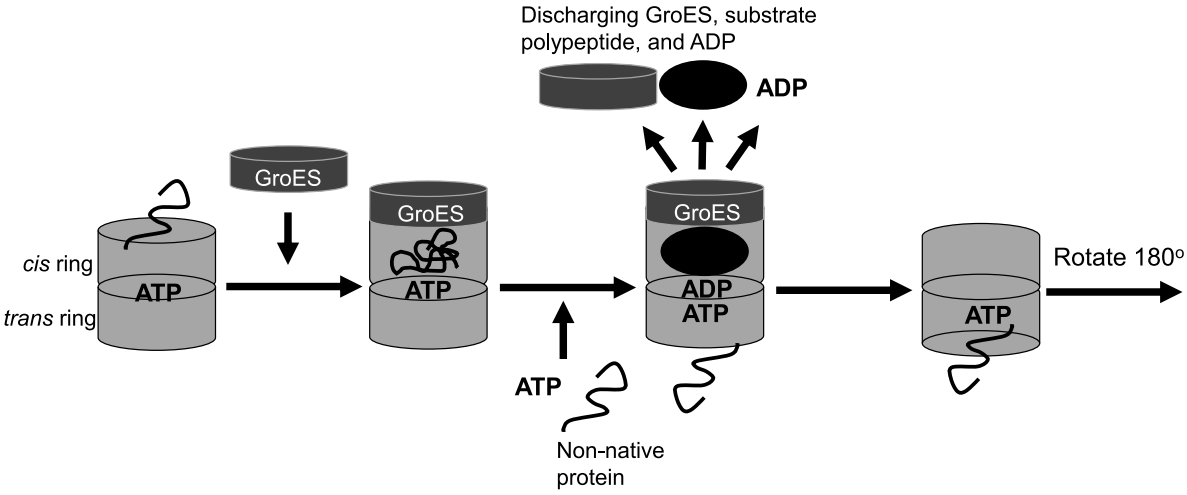

Fig. 2 Chaperone cycle of *E. coli* GroEL. ATP and polypeptide binding to one of the two GroEL heptamer rings is followed by binding of GroES. Concomitant with GroES binding, the polypeptide is ejected into the central cavity, where it begins to fold in this chamber. Folding continues during ATP hydrolysis in the cis ring (∼10 s). ATP hydrolysis weakens the cis complex and permits ATP (and non-native polypeptide) binding to the trans ring. This discharges ADP, substrate polypeptide, and GroES from the cis ring, regardless of the folding state of the substrate polypeptide. The released polypeptide may fold to its native state. When it is unable to reach the native state, it can bind to GroEL again for another folding attempt. The previous trans ring forms a new cis ring to begin the chaperone cycle again.

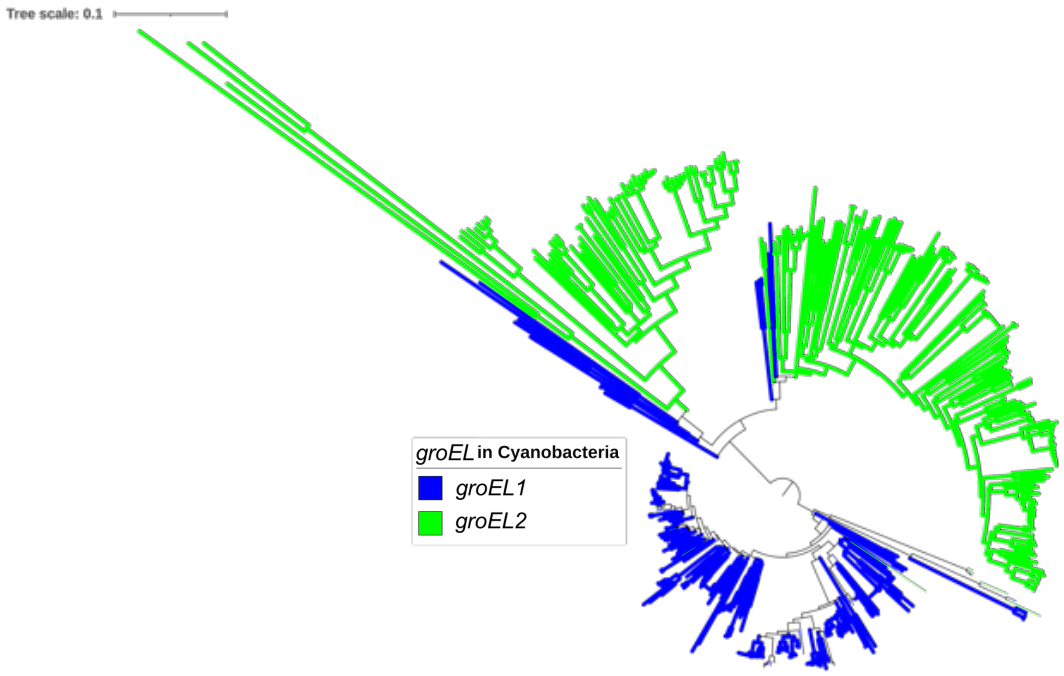

Fig. 3 Phylogenetic tree analysis of the *groEL* genes from 271 cyanobacterial genomes using the Interactive Tree of Life (iTOL) tool. All *groES* (10 kDa chaperonin) and *groEL* (60 kDa chaperonin) genes were retrieved directly from their annotations in GenBank. In this analysis, a *groEL* with *groES* nearby is defined as *groEL1*, while a *groEL* with no *groES* nearby is defined as *groEL2*. In the tree, *groEL1* (in blue) and *groEL2* (in green) originate from the same root, and diverge at the early stage of the evolution.

tree [17].

In this review, we reanalysed the *groESL1* operon and *groEL2* of the updated cyanobacterial genomes as of September 2022. The majority, approximately 85% of the cyanobacterial genomes, represent a single *groESL1* operon and one monocistronic *groEL2*, supporting the analysis by Weissenbech et al [17]. However, many combinations between the number of *groESL1* and *groEL2* were found among the 271 cyanobacterial genomes, as shown in Table 1(a). The number of *groE* genes present in plasmids are not counted in the table although some cyanobacteria including *Acaryochloris* species/strains, *Trichormus variabilis*, and *Nostoc* sp. were found to contain multiple *groE* genes in their plasmids. *Acaryochloris* has been reported to be unique cyanobacteria containing chlorophyll d as the predominant pigment [18]. *Trichormus variabilis*, and *Nostoc* sp. are nitrogen-fixing cyanobacteria with heterocysts. *Calothrix* sp. strain PCC 7716 which has a huge genome was found to contain 3 *groESL1* operons. A few cyanobacteria including *Gloeobacter violaceus* PCC 7421 has two *groESL* operons but no monocistronic *groEL*. Moreover, the phylogenetic tree constructed from the cyanobacterial *groESL1* operon and *groEL2* sequences clearly shows two main patterns of the *groESL1* operon and *groEL2* (Fig. 3).

Control of transcription of *groESL1* and *groEL2* in cyanobacteria

Negative regulation

A distinctive feature of the heat shock response in cyanobacteria is that it is modified not only by heat but also by light. Light accelerates the heat induction of *groESL1* and *groEL2* as well as genes encoding other molecular chaperones in *Synechocystis* sp. PCC 6803. In the dark, the heat shock response of the *groE* genes is not as rapid and intense as in the light. An inhibitor of photosynthetic electron transport, 3-(3,4-dichlorophenyl)-1,1-dimethylurea strongly suppressed heat induction of the *groE* genes in the light [19–21], suggesting that the light acceleration is closely linked to the operation of photosynthetic electron flow. The increase in the cellular level of molecular chaperones in the light is important for the survival of cells because exposure to high temperatures in the light becomes much more lethal to cells than heat-exposure in the dark [20]. In the light, heat may cause more damage/denaturation of cellular proteins than in the dark. What mechanism(s) underlies this unique and physiologically important heat shock response?

It is unlikely that a σ^{32} homologue is involved in the heat- and light-mediated regulation of cyanobacterial *groE* transcription because it is not conserved

in cyanobacterial genomes [21]. Alternative sigma factors such as SigB and SigH in the cyanobacterium *Synechocystis* sp. PCC 6803 are heat-induced, similar to *E. coli* σ^{32} [22–24]. However, it is not known whether these sigma factors transcribe the *groEL* genes in cyanobacteria under heat stress as major regulatory factors.

Transcription of the *groESL1* and *groEL2* genes of *Synechocystis* sp. PCC 6803 is initiated from the same and only one transcriptional start site under normal and heat shock conditions [25], suggesting that the same promoter is used regardless of the conditions. This eliminates the involvement of a stress-response alternative sigma factor acting on special heat shock promoters for *groE* heat induction and suggests that the heat shock response is regulated by the principle sigma factor RpoD1 without switching to an alternative sigma factor upon heat shock.

In *Bacillus subtilis*, *groESL* transcription is heat-induced depending on the principal sigma factor SigA [26]. It is regulated negatively in contrast to the positive regulation by σ^{32} in *E. coli*. Downstream of a SigA-dependent promoter, there is a perfect inverted repeat of 9 bp separated by a 9 bp spacer, designated the CIRCE element (controlling inverted repeat of chaperone expression) [27]. CIRCE acts as an operator to which the HrcA repressor binds to suppress *groE* transcription under normal condition [28]. Upon heat shock, HrcA inactivates or changes its conformation so that it dissociates from the CIRCE element, resulting in the heat induction of the *groESL* operon [29].

The HrcA-CIRCE system is the most widespread among eubacteria [29]. CIRCE is present in upstream regions of *groEL* genes from approximately 80% of cyanobacterial genomes (Table 1(b)). It is shown that both *groESL1* and *groEL2* from some cyanobacterial species including *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120 have CIRCE [25, 30–32]. Our present analysis reveals that CIRCE is present in 179 out of total 290 *groESL1s* from 271 genomes whereas it exists in 125 out of total 260 *groEL2s*. Thus CIRCE is not associated with all *groE* genes. The gene encoding the HrcA repressor protein that binds directly to CIRCE is also conserved in cyanobacterial genomes [33].

To demonstrate whether the HrcA-CIRCE system regulates cyanobacterial *groEL* gene expression, we disrupted the *hrcA* gene in *Synechocystis* sp. PCC 6803 [25]. In this cyanobacterium, CIRCE is present upstream of both the *groESL1* operon and the *groEL2* gene as already described. Immediately after the -10 promoter sequence of *groESL1*, the CIRCE operator is present. Most of the *groEL2* promoter overlaps with CIRCE. Thus, CIRCE is present around the transcription start sites of the *Synechocystis groESL1* and *groEL2* genes. The transcription of the *groE* genes is derepressed (or greatly enhanced) in the mutant under normal growth conditions at 30 °C. Similarly, the expression of both *groESL1* and *groEL2* is elevated in an *hrcA* null mutant of *Anabaena* sp. PCC 7120 under normal growth compared with the wild type [30]. The electrophoretic mobility shift assay shows that the isolated HrcA repressor protein from *Anabaena* sp. L-31 specifically binds to a DNA fragment containing CIRCE. These results strongly suggest that CIRCE and HrcA constitute a negative regulatory mechanism in the heat induction of the *groE* genes in these cyanobacteria.

Conservation of the CIRCE and HrcA sequences in cyanobacterial genomes does not necessarily mean that it constitutes a 'negative' regulatory mechanism in the *groE* expression. To our surprise, a *hrcA* mutant of *Synechococcus elongatus* PCC 7942 did not show derepression of the *groESL1* transcription under normal growth conditions [33]. Instead, the *groESL1* mRNA levels at 30 °C and a heat shock temperature 45 °C are lower in the mutant than that in the wild type. In *S. elongatus* PCC 7942, a novel heat shock protein, Orf7.5, is involved in positive regulation of the *groESL1* transcription [34]. Orf7.5 is not conserved in some cyanobacteria, including *Synechocystis* sp. PCC 6803. Orf7.5 interacts physically with RpoD1 which is expected to be involved in the transcription of *groESL1* [33]. We suggest that the CIRCE/HrcA system is functionally affected/modified by Orf7.5.

Positive regulation by K-Box

In addition to the negative regulation mediated by the CIRCE/HrcA system, we obtained evidence for the operation of another mechanism for regulation of the cyanobacterial *groEL* gene expression: transcriptions of the *groESL1* and *groEL2* genes in the *hrcA* mutant of *Synechocystis* sp. PCC 6803 were still induced by heat shock at 42 °C [25]. To study the *hrcA*-independent regulatory mechanism(s), we collected and compared conserved sequences present upstream of *groE* genes of various cyanobacterial species. Further upstream from CIRCE there was a conserved sequence that we designated the K-box. Involvement of K-box in the heat- and/or light-induction of *groESL1* transcription in *Synechocystis* sp. PCC 6803 is shown by reporter assays. In the absence of the K-box, *groESL1* promoter activity is completely abolished under low and high light at a normal or a high temperature. Results with the assays indicate that the K-box is essential not only for heat and/or light induction but also for the basal transcription of the *groESL1* operon [21].

Herein, we found that K-box is present in upstream regions of *groESL1* genes from approximately 35% of cyanobacterial genomes (Table 1(b) and Fig. 4(a)). The K-box is also conserved in the upstream regions of *dnaK2* from various cyanobacterial species [21, 35]. There are three homologues of *dnaK* in cyanobacteria including *S. elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803. The *dnaK2* genes from *S. elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803 are heat- and/or

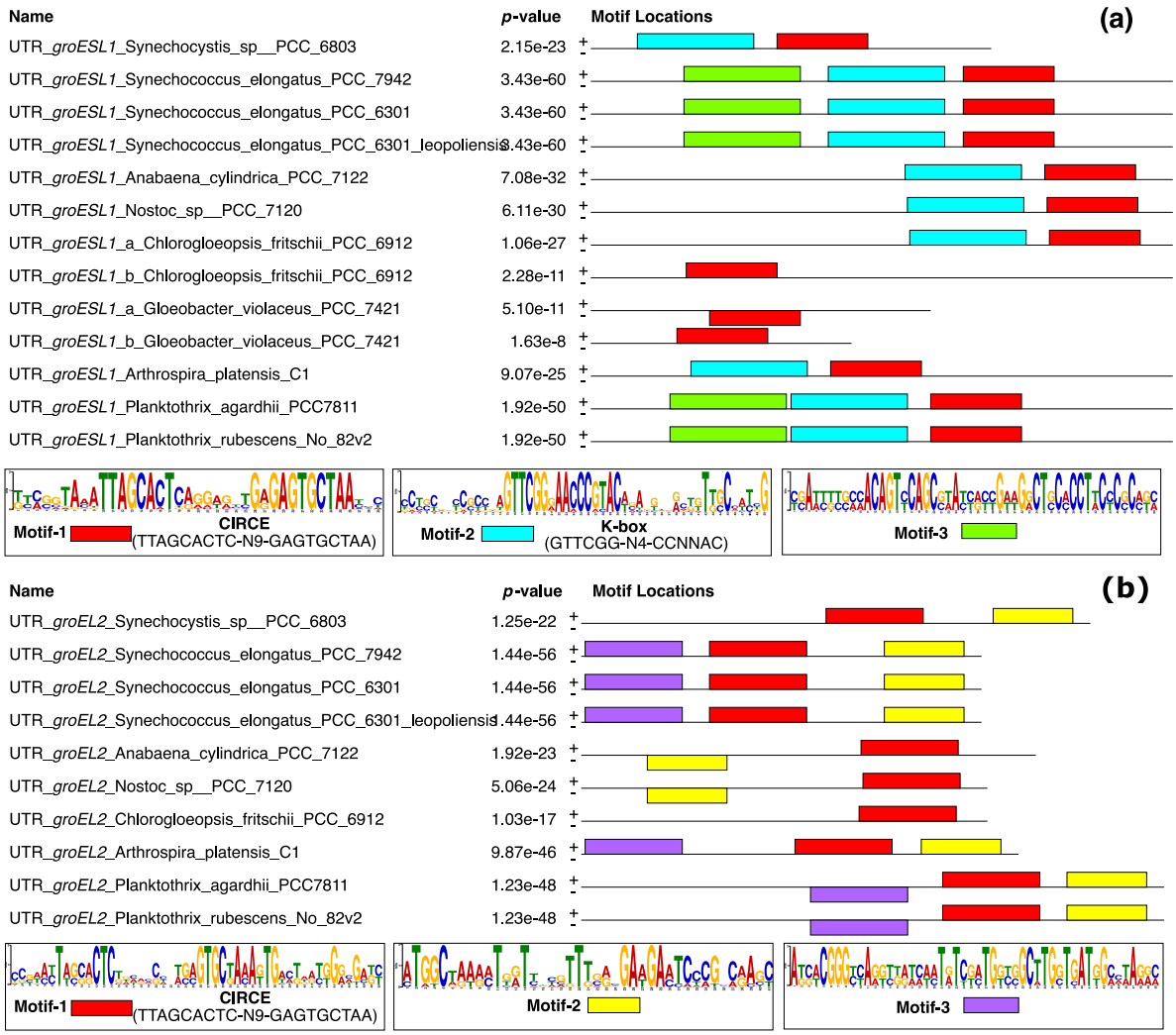

Fig. 4 Alignments of nucleotide sequences of the region from +50 bp to −250 bp from ATG of the *groESL1* operon (a) and the monocistronic *groEL2* (b) from various cyanobacteria. The MEME suite tool for detecting motif logos was used to search for conserved motifs. CIRCE and K-box were searched for with the following sequences. CIRCE, -TTAGCACTC-N9-GAGTGCTAA-; K-box, -GT-N2-GG-N4-CCGAAC-, -GTTCGG-N4-CC-N2-AC-. A motif is statistically significant when p-value is less than 0.05. Some cyanobacteria including *G. violaceus* PCC 7421 contain two *groESL1* operons. Note that the *Gloebacter groESL1* operon has no K-box. Note also that K-box is not conserved in the upstream regions of *groEL2s*. (a) Upstream sequence alignments of some *groESL* operons. Three conserved motifs are shown: motif 1 (CIRCE) in red, motif 2 (K-box) in blue, and motif 3 (a novel motif found in this study) in green. (b) Upstream sequence alignments of some *groEL2* genes. Three conserved motifs are shown: motif 1 (CIRCE) in red, motif 2 (a novel motif) in yellow, and motif 3 (a novel motif) in purple.

light-induced [35, 36]. DnaK2 is the most abundant DnaK protein in the cyanobacteria. Thus, the K-box is an important DNA element that regulates genes encoding highly conserved, major molecular chaperones in cyanobacteria whose expression is modulated by heat and light.

Differential regulation of groESL1 and groEL2 upon heat induction

As described above, heat induction of the *groESL1* and *groEL2* genes is modulated by light. Transient induction is more rapid and intense in the light than in the dark. However, this tendency is much more apparent in *groEL2* than in *groESL1*. Thus, it is suggested that light differentially affects the heat induction of each of the two *groE* genes [19, 20]. An underlying mechanism for this difference is not known. To obtain a clue for this, we analysed the upstream regulatory regions of *groESL1* and *groEL2* from various cyanobacteria. We found that K-box is not well conserved in upstream regions of *groEL2s* [31, 32]. In fact, our present analysis showed that only 2 out of 271 cyanobacterial genomes

Table 1 Distribution of multiple *groEL* genes among 271 cyanobacterial genomes (a) and that of CIRCE and K-box in upstream regions of *groEL* with *groES* (*groEL1*) and *groEL* without *groES* nearby (*groEL2*) (b). In (a), *groES* and *groEL* genes were retrieved directly from their annotations in GenBank. These annotations are either gene names or functional domains, i.e., 10 kDa chaperonin (cpn-10) and 60 kDa chaperonin (cpn-60) domains. In (b), CIRCE and K-box were searched for in the region from +50 bp to −250 bp from ATG of *groE* genes with the following sequences. CIRCE, -TTAGCACTC-N9-GAGTGCTAA-; K-box, -GT-N2-GG-N4-CCGAAC-, -GTTCGG-N4-CC-N2-AC-.

(a)	No. of *groEL1s*	No. of *groEL2s*	No. of organisms
	3	1	1
	2	1	13
	2	–	4
	1	2	7
	1	1	232
	1	–	14

(b)	No. of CIRCE and K-box in upstream regions of *groESL1s*		of *groEL2s*		No. of organisms
	CIRCE	K-box	CIRCE	K-box	
	2	–	1	–	1
	2	–	–	–	3
	1	1	1	–	37
	1	1	–	–	44
	1	–	1	1	1
	1	–	1	–	42
	1	–	–	–	47
	–	1	1	–	5
	–	1	–	–	9
	–	–	1	1	1
	–	–	1	–	38
	–	–	–	–	43

have K-box in their *groEL2s* (Table 1(b)). We also found novel DNA elements (motifs 2 and 3 in Fig. 4(b)) which appear to be unique to *groEL2*. In addition to the conserved regulatory elements/motifs, species-specific ones such as N-box [21] and H-box [30] have been reported to be involved in *Synechocystis* and *Anabaena groESL1* transcription, respectively.

The evolution of regulatory mechanisms in cyanobacterial groEL paralogues

Almost all cyanobacterial genomes encode at least one *groESL* (*groESL1*) operon and a single (two or more in rare cases) monocistronic *groEL* (*groEL2*) gene [17, 35, Table 1(a)]. It is notable that *G. violaceus* has two *groESL* operons but no *groEL2*. *G. violaceus* is an exceptional cyanobacterium because it lacks thylakoid membranes [38] and is suggested to be a member of an early branching lineage [39]. In terms of regulatory elements, both *groESL* operons contain CIRCE. The cyanobacterium also has a gene coding for HrcA. However, there is no K-box upstream of either of the two *groESLs*. Based on the information we have gathered so far, we hypothesize that the cyanobacterial ancestor had only one *groESL* operon. It duplicated to produce two operons. These *groESL* paralogues had the CIRCE-HrcA system to regulate the expression of these genes, similar to *G. violaceus*. CIRCE is widespread/conserved among bacteria [29] including cyanobacteria (Table 1(b)), which suggests that CIRCE and CIRCE-HrcA system are ancient. During evolution, one of the paralogues lost the *groES* gene, resulting in the *groEL2* gene. Furthermore, the other one (*groESL1*) acquired a K-box in addition to CIRCE (Fig. 5). The acquisition of K-box after CIRCE is supported by the fact that K-box is almost always associated with CIRCE (Table 1(b)).

The preservation of the monocistronic *groEL2* gene in cyanobacteria during evolution is suggested to result from neofunctionalization. It is known that duplicated genes will have different fates, including nonfunctionalization, subfunctionalization, and neofunctionalization [40]. In nonfunctionalization, one copy becomes silenced or lost by degenerative mutations, whereas two copies lose functions complementarily in subfunctionalization. In neofunctionalization, one copy acquires a novel, beneficial function, with the other copy retaining the original function. The preservation of multiple functional *groEL* genes in cyanobacteria during evolution suggests that they resulted from either subfunctionalization or neofunctionalization. Differential regulation of *groESL1* and *groEL2* upon heat induction and difference in regulatory elements (e.g. the presence or absence of K-box) between *groESL1* and *groEL2* suggest that the *groEL2* paralogue is an outcome of neofunctionalization. If it is an outcome of subfunctionalization, then the expression of *groEL1* and *groEL2* is expected to be regulated by the same mechanism.

Structure and function of translation products of the *groEL* paralogues

If the *groEL2* paralogue were an outcome of neofunctionalization, *groEL2* would acquire a novel, beneficial function with *groESL1* retaining the original function. In general, the function of a protein is related to or governed by its structure. Thus, a novel function must come from a novel structure. We have been working to show that GroEL1 and GroEL2 have mutually different structures and functions.

Functional difference between GroEL1 and GroEL2

GroEL1 is functionally equivalent to E. coli GroEL, whereas GroEL2 is not. To evaluate whether *groEL1* and *groEL2* are equivalent to *E. coli groEL*, complementation tests with *E. coli groEL* mutants were carried out. In early complementation tests, the mutant strain *groEL44*, which carries the E191G mutation in GroEL, was used. It exhibits a temperature-sensitive phenotype, growing at 30 °C or 37 °C but not at 42 °C. The mutant is transformed with a plasmid harbouring

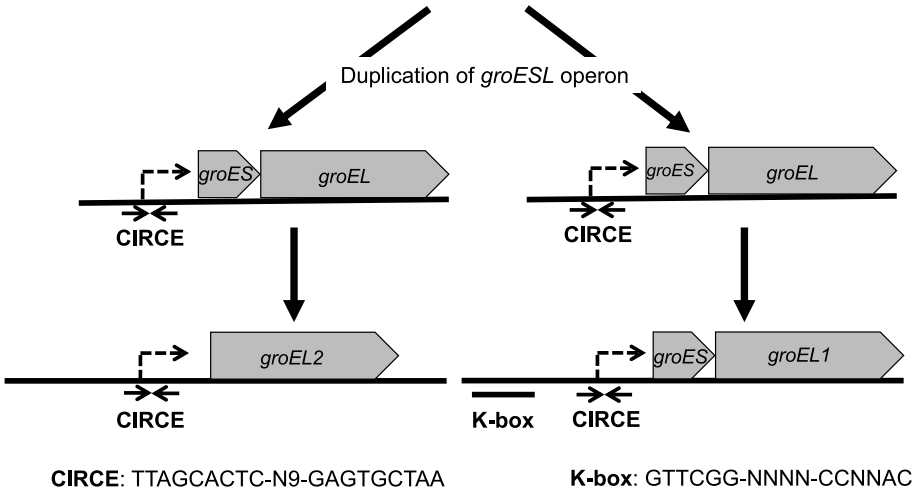

Fig. 5 Hypothetical model for the evolution of *groESL1* and *groEL2* in majority of cyanobacteria. Ancient cyanobacteria have two *groESL* operons with CIRCEs around their promoters, similar to *G. violaceus* PCC 7421 which is thought to have diverged very early during the evolution of cyanobacteria. During evolution, one of the duplicated genes lost *groES* to become the *groEL2* gene. In some cyanobacterial species, *groEL2* has acquired a new regulatory element(s). The other *groESL* operon has a conserved *groES* gene, CIRCE, and further acquired a K-box.

groESL1, *groEL1*, or *groEL2* from the thermophilic cyanobacterium *Synechococcus vulcanus* or *Synechocystis* sp. PCC 6803 [13, 41, 42]. The level of complementation is evaluated by the recovery of thermotolerance. In more recent complementation assays, the mutant strain MGM100 [43] was used. The expression of the native *groESL* operon in MGM100 is controlled by an arabinose-inducible pBAD promoter. The *groES* and *groEL* genes are essential [9], so the strain is kept viable in the presence of arabinose. Growth/survival of MGM100 takes place if *groES*/*groEL* introduced by a plasmid can complement the native *groESL* gene in the MGM100 strain in the absence of arabinose.

The results thus far show that the *groEL1* gene from the thermophilic unicellular cyanobacterium *S. vulcanus* or the mesophilic unicellular cyanobacterium *Synechocystis* sp. PCC 6803 can complement the native *E. coli groEL* gene [41, 42]. Furthermore, the *groEL1* or *groEL1.2* gene expressed with the *groES1* gene from the filamentous heterocystous cyanobacterium *C. fritschii* PCC 6912 can complement the native *E. coli groESL* operon [17]. Thus, both GroEL1 and GroEL1.2 require GroES to play the essential role of native GroEL and GroES in *E. coli* cells. On the other hand, the *groEL2* genes from *S. vulcanus* and *C. fritschii* PCC 6912 are unable to do so under normal or heat-stressed conditions regardless of the presence or absence of a *groES* gene [13, 17]. The *groEL2* gene from *Synechocystis* sp. PCC 6803 can complement the native *E. coli groEL* gene, but only partially [42]. In contrast, the *groEL1* or *groEL2* gene from *Anabaena* sp. strain L-31 can complement the native *groESL* gene in the MGM100 strain without co-expression of the *groES* gene at a high temperature [44]. Although there may be exceptions, we infer that GroEL1 is functionally equivalent to *E. coli* GroEL, whereas GroEL2 is not.

GroEL1 is essential, whereas GroEL2 plays an important role under stress. There is a report that it is impossible to disrupt the *groEL1* gene in all genome copies of *S. elongatus* PCC 7942 [35], suggesting that the gene is essential. This essentiality indicates that GroEL1 plays an essential role for which GroEL2 is unable to substitute. On the other hand, the *groEL2* gene in all genome copies of the thermophilic cyanobacterium *T. elongatus* can be disrupted [45], indicating that *groEL2* is nonessential under normal growth conditions at 50 °C. The *groEL2* mutant strain grows similarly to the wild type at 50 °C. However, it is unable to grow at 62 °C or at 40 °C, indicating that it is high- and low-temperature sensitive. Thus, GroEL2 plays a crucial role under the temperature stresses. In other words, GroEL2 exerts beneficial effects under stress conditions. We must point out that GroEL1 is present at the same or even higher level at low temperature in a *groEL2* mutant cell [45]. This suggests that GroEL1 is unable to substitute for the function of GroEL2. Consistent with GroEL2 function under temperature stresses, the wild type induced the *groEL2* gene at both 40 °C and 63 °C. The *groEL2* gene in all genome copies of the mesophilic cyanobacterium *S. elongatus* PCC 7942 is also disruptable (Huq and Nakamoto, unpublished data). As shown in Fig. 6, growth at a slightly high temperature of 42 °C is greatly impaired due to the mutation, although there is no difference in growth between the mutant and the wild type at 30 °C. The *groEL2* mutant strain also shows higher

sensitivity to high light and high salt than the wild-type strain (Fig. 6). Thus, we conclude that GroEL2 is nonessential in both mesophilic and thermophilic cyanobacteria, in contrast with the essential GroEL1 and *E. coli* GroEL. However, we hypothesize that GroEL2 confers cell tolerance to various stresses, such as heat, cold, high light, and high salt. GroEL2 exerts beneficial effects under various abiotic/environmental stresses. Cyanobacteria must have encountered various environmental stresses, and thus, we suggest that GroEL2 has played an important role during evolution of cyanobacteria.

In vitro chaperone function of GroEL1 and GroEL2. Prevention of protein aggregation is characteristic of evolutionarily conserved molecular chaperones, including GroEL. Non-native/denatured proteins are specifically recognized and captured by molecular chaperones, which usually results in suppression of aggregation of the proteins. Prevention of protein aggregation forms a first line of cellular defence under stress. This anti-aggregation activity can be quantified by measuring the increase in apparent absorbance, turbidity or light scattering in a solution where denatured proteins aggregate.

Both *Synechococcus* GroEL1 and GroEL2 suppress the aggregation of heat (45 °C)-denatured malate dehydrogenase (MDH), similar to *E. coli* GroEL at pH 8.0 [46]. *Anabaena* GroEL1 also suppresses the aggregation of heat (55 °C)-denatured MDH at pH 7.4 [44]. These results do not show any difference in the anti-aggregation activity between GroEL1, GroEL2, and *E. coli* GroEL. However, we have observed differential effects of pH on the anti-aggregation activity of each GroEL [47]. When the pH decreases from 8.5 to 7.0, GroEL1 loses its ability to suppress the aggregation of heat-denatured MDH most sharply, whereas the activity of *E. coli* GroEL is most resistant to pH changes. Compared with GroEL1, GroEL2 shows a modest response to the change. Thus, cyanobacterial GroELs are more sensitive to pH changes than *E. coli* GroEL. It is known that the pH in the cyanobacterial cytosol and chloroplast stroma increases one pH unit upon illumination [48, 49]. Our *in vitro* observations suggest that GroEL1 exerts more chaperone activity in the light when ATP-producing photosynthesis takes place. GroEL1 has much higher ATPase activity than GroEL2 [46]. In other words, GroEL1 consumes more ATP than GroEL2. The differential effect of pH on the anti-aggregation activity of GroEL1 and GroEL2 may indicate that cells distinguish chaperone usage depending on the cellular energy status.

Structural difference between GroEL1 and GroEL2

Oligomer structure of GroEL: *E. coli* forms a 14-mer or double ring of the 7-mer (Fig. 1). This oligomer structure is essential for GroEL to build a folding chamber. Therefore, the functional difference between GroEL1 and GroEL2, if any, may be reflected in the oligomer state. There is a technical hurdle in the analysis of the oligomer states of cyanobacterial GroELs. To the best of our knowledge, GroEL1 and GroEL2 from *Synechocystis* sp. PCC 6803 and *S. elongatus* PCC 7942 are much more unstable than *E. coli* GroEL, which hampers reliable determination of their oligomer states and chaperone functions. Perhaps due to this instability, oligomer states of cyanobacterial GroEL1 and GroEL2 are reported to vary depending on cyanobacterial species, analytical methods, and experimental conditions. GroEL1, GroEL2, or both from *Synechocystis* sp. PCC 6803 is detected in native PAGE as a tetradecamer only in the presence of glycerol which is known to stabilize proteins [42]. In *S. elongatus*, analysis using native PAGE showed that GroEL1 and GroEL2 do not form a 14-mer, whereas the *E. coli* GroEL gives a clear band of the 14-mer under the same conditions [46]. Oligomers ranging from pentamer to dimer of GroEL1 are detected. On the other hand, GroEL2 always forms a dimer. Glycerol and MgATP, which stabilize a 14-mer of *Synechocystis* GroELs, do not affect the oligomeric state of GroEL1 and GroEL2. High salt and/or high protein concentrations often stabilize some proteins. We also analysed the oligomeric state of GroELs in the presence of 300 mM NaCl and varying concentrations of GroEL1 and GroEL2 using gel filtration chromatography. The results showed that GroEL1 can form a 14-mer at a high protein concentration, but the largest oligomer of GroEL2 is a heptamer under the same conditions. The nitrogen-fixing cyanobacterium *Anabaena* L-31 also has two kinds of GroELs. *Anabaena* GroEL1 shows two peaks in gel filtration chromatography [44]. The calculated sizes of the oligomers are >700 kDa and 61.7 kDa, corresponding to a large oligomer (>12-mer) and a monomer, respectively. Based on these results, we conclude that only GroEL1 and not GroEL2, is able to form a 14-mer, although it is less stable than the *E. coli* GroEL.

Recently, bacterial two-hybrid analysis was employed to study interactions between GroEL and GroES [17]. In this study, GroELs and GroESs were from *C. fritschii*, whose genome contains two *groESL1* operons (*groESL1* and *groESL1.2*) and one *groEL2* gene. This study indicated that GroEL2 does not interact with either itself or with GroEL1/GroEL1.2. GroEL1 can self-interact, but it does not interact with GroEL1.2. These results suggest that GroEL1 forms homo-oligomer(s), whereas GroEL2 is a monomer. Such finding is noteworthy, although further biochemical analysis with isolated GroELs and GroESs is necessary to verify these physical interactions.

Interaction of GroEL1 and GroEL2 with GroES: *E. coli* GroEL cooperates with GroES to assist in the folding of non-native proteins as described. GroES acts like a lid to close the folding chamber of the GroEL

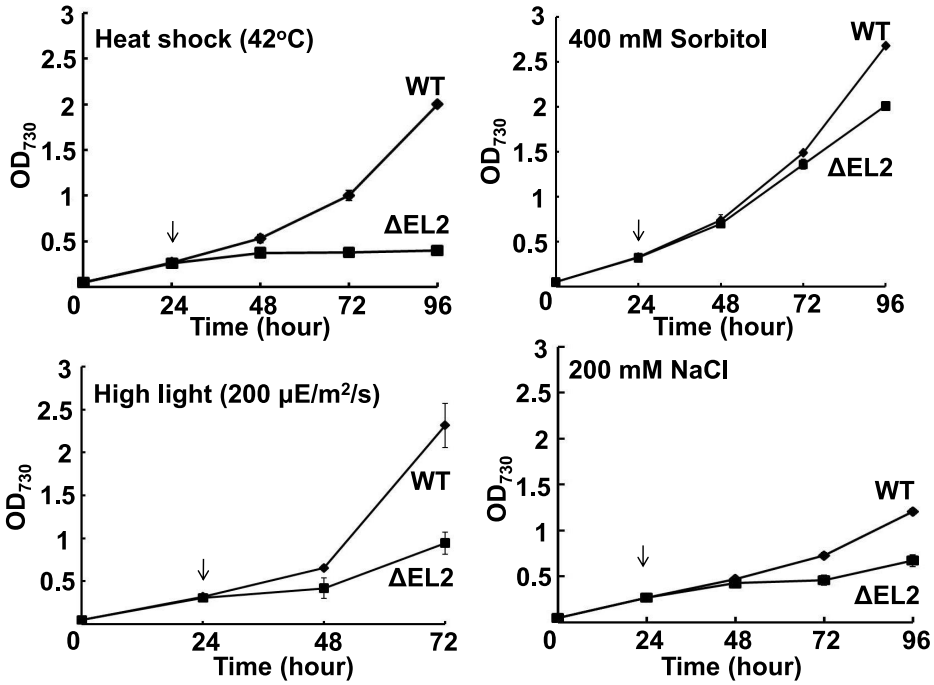

Fig. 6 Growth of the wild-type *S. elongatus* PCC 7942 and its *groEL2* deletion mutant under various conditions.

heptamer (Fig. 1). *Synechocystis* GroES physically interacts with GroEL1 and/or GroEL2 when analysed using sucrose gradient centrifugation and gel filtration column chromatography [42]. *Chlorogloeopsis* GroEL1 and GroEL1.2 interact with GroES1/GroES1.2 and GroES1, respectively, as shown by bacterial two-hybrid analysis [17]. In contrast, GroEL2 interacts with none of the GroESs. The 'functional' interaction of *Chlorogloeopsis* GroEL1 and GroEL1.2 with GroES1 is confirmed using complementation analysis in which *groEL1* or *groEL1.2* can complement the native *groEL* in *E. coli* MGM100 only when the *groES1* gene is co-expressed [17].

Structural difference in relevance to functional difference: These structural and protein-interaction studies indicate that GroEL1 and GroEL2 have mutually different oligomer structures. GroEL1 appears to conserve 'the original structure', similar to *E. coli* GroEL, whereas GroEL2 acquires a novel structure that is very different from *E. coli* GroEL or GroEL1. It is yet to be determined whether the lower oligomer state of GroEL2 than GroEL1 is related to a beneficial function, such as stress tolerance, in cells. When the GroEL oligomer is dissociated, it loses the chaperone activity to facilitate protein folding as well as ATPase activity. In fact, the ATPase activity of GroEL2 is much lower than that of GroEL1 [46]. However, as described above, GroEL2 retains its anti-aggregation activity. Under stress, chaperone-assisted refolding of a denatured protein may be useless because the refolded protein may be denatured again. Chaperone-assisted refolding is costly or energetically unfavourable because it consumes a large amount of ATP. In contrast, ATP is not required for GroEL to prevent protein aggregation at least *in vitro*. One of the benefits for a cell in maintaining GroEL2 may be to protect and prevent aggregation of a denatured protein with less energy under stress. The denatured protein may be kept in a folding-competent state by forming a complex with GroEL2 under stress. When stress is gone, the denatured protein is transferred from GroEL2 to the other chaperone system, such as DnaK, and refolds back to its native state with the assistance of the chaperone system [31].

PROTEOMIC ANALYSIS OF THE WILD TYPE AND A *groEL2* MUTANT OF *S. elongatus* PCC 7942

As described, GroEL1 appears to be essential and equivalent to *E. coli* GroEL. It has been shown that there are ∼80 proteins in *E. coli*, some of which are essential proteins and are dependent on GroEL/GroES [50, 51]. GroEL substrates range from 21 to 68 kD in size [51]. The size of the GroEL/GroES folding chamber is thought to limit the size of the substrates. GroEL2 does not form a folding chamber, thus, the sizes of substrates for GroEL2 may not be restricted. GroEL2 is not indispensable but is required for (better or competitive) growth under stresses, including heat and cold. To determine how GroEL2 plays a role under stress, proteins in *S. elongatus* PCC 7942 wild type and a *groEL2*-deletion mutant grown under low- and high-temperature stresses were examined using

Table 2 Up- or down-regulated proteins identified by comparative proteome analysis of wild type (WT) and *groEL2* deletion mutant (MT). n/a: not available.

Growth temp.	Log flow change ([MT] vs [WT]) :normalized	Regulation ([MT] vs [WT]) :normalized	Locus_tag	Protein name	NCBI category	KEGG Pathway name
Up-regulation						
16 °C	2.6264749	up	Synpcc7942_0304	conserved hypothetical protein	Coenzyme transport and metabolism	n/a
	2.4622715	up	Synpcc7942_0443	conserved hypothetical protein	n/a	n/a
	2.4622715	up	Synpcc7942_0443	conserved hypothetical protein	n/a	n/a
	3.9303222	up	Synpcc7942_0656	photosystem II 44 kDa subunit reaction center protein	Energy production and conversion	Photosynthesis
	3.9303222	up	Synpcc7942_0656	photosystem II 44 kDa subunit reaction center protein	Energy production and conversion	Metabolic pathways
35 °C	3.5074956	up	Synpcc7942_0039	6-phosphogluconate dehydrogenase (decarboxylating)	Carbohydrate transport and metabolism	Pentose phosphate pathway
	2.0916889	up	Synpcc7942_0443	conserved hypothetical protein	n/a	n/a
	2.2557704	up	Synpcc7942_1123	SSU ribosomal protein S18P	Translation, ribosomal structure and biogenesis	n/a
	4.9901752	up	Synpcc7942_1298	diguanylate cyclase (GGDEF domain)	Signal transduction mechanisms	n/a
	3.681321	up	Synpcc7942_1541	Flavodoxin, long chain	Energy production and conversion	n/a
	2.346918	up	Synpcc7942_1656	catalase/peroxidase HPI	Inorganic ion transport and metabolism	Phenylalanine metabolism
45 °C	3.865192	up	Synpcc7942_0039	6-phosphogluconate dehydrogenase (decarboxylating)	6-phosphogluconate dehydrogenase	Carbohydrate transport and metabolism
	2.8405306	up	Synpcc7942_0240	hypothetical protein	Allophycocyanin alpha chain	Energy production and conversion
	5.192276	up	Synpcc7942_0304	conserved hypothetical protein	Ubiquinone biosynthesis protein Coq4	Coenzyme transport and metabolism
	5.1311707	up	Synpcc7942_0443	conserved hypothetical protein	n/a	n/a
	1.5691743	up	Synpcc7942_0505	D-fructose 1,6-bisphosphatase/sedoheptulose 1,7-bisphosphatase	Fructose-1,6-bisphosphatase/sedoheptulose 1,7-bisphosphatase or related protein	Carbohydrate transport and metabolism
	5.1854677	up	Synpcc7942_0682	conserved hypothetical protein	ADP-heptose:LPS heptosyltransferase	Cell wall/membrane/envelope biogenesis
	2.5925174	up	Synpcc7942_0700	conserved hypothetical protein	Nucleotide-binding universal stress protein, UspA family	Signal transduction mechanisms
	3.8947449	up	Synpcc7942_0788	glutathione S-transferase	Glutathione S-transferase	Posttranslational modification, protein turnover, chaperones
	4.034714	up	Synpcc7942_1298	diguanylate cyclase (GGDEF domain)	GGDEF domain, diguanylate cyclase (c-di-GMP synthetase) or its enzymatically inactive variants	Signal transduction mechanisms
	7.7302375	up	Synpcc7942_1314	FtsH-2 peptidase. Metallo peptidase. MEROPS family M41	ATP-dependent Zn proteases	Posttranslational modification, protein turnover, chaperones
	2.6406696	up	Synpcc7942_1541	Flavodoxin, long chain	Flavodoxin	Energy production and conversion
	2.5284424	up	Synpcc7942_1678	conserved hypothetical protein	Photosystem II oxygen-evolving enhancer PsbQ/CyanoQ	Energy production and conversion
	1.5520245	up	Synpcc7942_1781	conserved hypothetical protein	Molybdopterin synthase sulfur carrier subunit MoaD	Coenzyme transport and metabolism
	2.717964	up	Synpcc7942_2095	conserved hypothetical protein	n/a	n/a
	3.7349293	up	Synpcc7942_2297	transaldolase	Transaldolase/fructose-6-phosphate aldolase	Carbohydrate transport and metabolism
Down regulation						
16 °C	−3.6590142	down	Synpcc7942_0697	photosystem II core light harvesting protein	Energy production and conversion	Photosynthesis
	−1.643912	down	Synpcc7942_2010	cytochrome c550	Energy production and conversion	Photosynthesis
	−2.5985527	down	Synpcc7942_2315	ATP synthase F1, beta subunit	Energy production and conversion	Oxidative phosphorylation
	−2.5997577	down	Synpcc7942_2342	hypothetical protein	Energy production and conversion	Photosynthesis
	−2.6526523	down	Synpcc7942_2445	Phosphate binding protein	Inorganic ion transport and metabolism	ABC transporters
35 °C	−2.9589658	down	Synpcc7942_0294	photosystem II manganese-stabilizing polypeptide	Energy production and conversion	Photosynthesis
	−2.8014147	down	Synpcc7942_0700	conserved hypothetical protein	Signal transduction mechanisms	n/a
	−3.120836	down	Synpcc7942_0788	glutathione S-transferase	Posttranslational modification, protein turnover, chaperones	Glutathione metabolism
	−1.5390956	down	Synpcc7942_2535	diguanylate cyclase (GGDEF domain) with PAS/PAC and GAF sensors	Signal transduction mechanisms	n/a
45 °C	−3.0720017	down	Synpcc7942_0321	nitrogen regulatory protein P-II	Amino acid transport and metabolism	Signal transduction
	−2.330923	down	Synpcc7942_0325	Lc 7.8 apoprotein (core components of the phycobilisomes)	Energy production and conversion	Energy metabolism
	−2.3214746	down	Synpcc7942_0834	conserved hypothetical protein	n/a	n/a
	−3.1545506	down	Synpcc7942_1427	ribulose 1,5-bisphosphate carboxylase small subunit	Carbohydrate transport and metabolism	Carbohydrate metabolism
	−3.409163	down	Synpcc7942_2342	hypothetical protein	Energy production and conversion	Energy metabolism
	−3.6714973	down	Synpcc7942_2535	diguanylate cyclase (GGDEF domain) with PAS/PAC and GAF sensors	Signal transduction mechanisms	n/a

quantitative proteome and phosphoproteome analysis, representing protein regulation at the translational and posttranslational levels, respectively. The proteins regulated at both levels were identified and designated regulatory hub proteins [52]. It was found that 64 proteins, including 6 molecular chaperones (GroEL1, GroEL2, ClpB2, DnaK3, DnaJ and HtpG), are involved in the hub. In addition to chaperones, proteins involved in signal transduction, photosynthesis, energy production, CO_2 concentration, nitrogen metabolism, the Calvin cycle, protein biosynthesis and protein transport were regulated at both translational and posttranslational levels.

According to the proteome and phosphoproteome data, GroEL2-dependent proteins are thought to be those whose expression level and/or phosphorylation state are altered in the absence of GroEL2. Proteins involved in DNA repair (such as GyrA), chlorophyll biosynthesis (such as ChlN), and the cytochrome b_6f complex which mediates electron transfer from PSII to PSI, were found in the absence of GroEL2 [52].

In addition to the regulatory hub proteins reported earlier [52], the protein expression level was considered and the significantly up- and down-regulated proteins compared between the wild-type and mutant strains can be classified as shown in Table 2. In the absence of GroEL2, the metabolic pathway affected under low-temperature stress was energy production via photosynthesis. However, proteins involved in nitrogen assimilation, carbohydrate and coenzyme transport, signal transduction and posttranslational modification were regulated at the protein expression level under high temperature. It should be noted that the increased or decreased levels of phycobiliproteins might be due to its denaturation under heat stress [53].

Taken together with the data concerning the regulatory hub proteins, in the absence of GroEL2, the cells responded to cold stress by controlling their energy storage, carbon metabolism and circadian clock, probably via the two-component transcriptional regulator Synpcc7942_1453. However, the response to high-temperature stress in the absence of GroEL2 involved carbohydrate/coenzyme transport, cell wall biogenesis and protein turnover, probably controlled by universal stress protein and diguanylate cyclase (GGDEF domain).

Moreover, it is noteworthy that glutamate synthase (GltB), which synthesizes glutamate from 2-oxoglutarate (2-OG), the compound at the interconnection of C and N metabolism, was identified in the regulatory hub of both the wild type and *groEL2* mutant under temperature stress. Its level is well known as a signal for nitrogen regulatory protein PII (GlnB) phosphorylation in response to N limitation [54], which was found to be downregulated in the *groEL2* deletion mutant after exposure to high temperature. Thus, despite the absence of GroEL2, the quantitative and phosphoproteome analysis indicate that the regulation of the C/N ratio under temperature stress is similar to that under N starvation.

E.coli GroEL Cyanobacterial GroEL1		Cyanobacterial GroEL2
Yes	Essential?	No
Yes	Stress inducible?	Yes
Yes	Compliments an *E. coli groEL* mutant?	No
Yes	Suppresses aggregation of denatured proteins?	Yes
Yes	Forms an oligomer of 14-subunits?	No

Fig. 7 Comparison of *E. coli* GroEL, cyanobacterial GroEL1, and cyanobacterial GroEL2.

CONCLUDING REMARKS

Fig. 7 compares the characteristics of cyanobacteria GroEL1 and GroEL2 with those of *E. coli* GroEL. Clearly, GroEL2 is different from GroEL1 and *E. coli* GroEL. We conclude that the *groEL2* gene is the outcome of neofunctionalization. It has acquired a novel, beneficial structure and function and has been preserved by natural selection, with the *groEL1* gene retaining its original function. GroEL2 may have played an essential function under environmental stress conditions, as we proved in GroEL2 from the thermophilic cyanobacterium *T. elongatus*, and thus has become preserved by natural selection.

Acknowledgements: This work was supported in part by Grants-in-Aid for Scientific Research (C) (No. 18K05407, to H.N.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1. Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332.
2. Richter K, Haslbeck M, Buchner J (2010) The heat shock response: life on the verge of death. *Mol Cell* **40**, 253–266.
3. Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* **181**, 223–230.
4. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J Mol Biol* **222**, 599–620.
5. Minton AP (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J Biol Chem* **276**, 10577–10580.

6. Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* **26**, 597–604.
7. Horwich AL, Fenton WA, Chapman E, Farr GW (2007) Two families of chaperonin: physiology and mechanism. *Annu Rev Cell Dev Biol* **23**, 115–145.
8. Mayer MP (2010) Gymnastics of molecular chaperones. *Mol Cell* **39**, 321–331.
9. Fayet O, Ziegelhoffer T, Georgopoulos C (1989) The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J Bacteriol* **171**, 1379–1385.
10. Rosen R, Ron EZ (2002) Proteome analysis in the study of the bacterial heat-shock response. *Mass Spectrom Rev* **21**, 244–265.
11. Guisbert E, Yura T, Rhodius VA, Gross CA (2008) Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev* **72**, 545–554.
12. Chitnis PR, Nelson N (1991) Molecular cloning of the genes encoding two chaperone proteins of the cyanobacterium *Synechocystis* sp. PCC 6803. *J Biol Chem* **266**, 58–65.
13. Furuki M, Tanaka N, Hiyama T, Nakamoto H (1996) Cloning, characterization and functional analysis of *groEL*-like gene from thermophilic cyanobacterium *Synechococcus vulcanus*, which does not form an operon with *groES*. *Biochim Biophys Acta* **1294**, 106–110.
14. Schroda M (2004) The *Chlamydomonas* genome reveals its secrets: chaperone genes and the potential roles of their gene products in the chloroplast. *Photosynth Res* **82**, 221–240.
15. Wongsamart R, Yokthongwattana C, Yokthongwattana K (2022) *Chlamydomonas* plastid chaperonin subunits expressed in *E. coli* can interact with one another inside the bacterial cell and putatively confer enhanced tolerance toward singlet oxygen. *ScienceAsia* **48**, 496–505.
16. Hill JE, Hemmingsen SM (2001) *Arabidopsis thaliana* type I and II chaperonins. *Cell Stress Chaperones* **6**, 190–200.
17. Weissenbach J, Ilhan J, Bogumil D, Hülter N, Stucken K, Dagan T (2017) Evolution of chaperonin gene duplication in stigonematalean cyanobacteria (subsection V). *Genome Biol Evol* **9**, 241–252.
18. Ohkubo S, Miyashita H (2012) Selective detection and phylogenetic diversity of *Acaryochloris* spp. that exist in association with didemnid ascidians and sponge. *Microbes Environ* **27**, 217–225.
19. Glatz A, Horváth I, Varvasovszki V, Kovács E, Török Z, Vigh L (1997) Chaperonin genes of the *Synechocystis* PCC 6803 are differentially regulated under light-dark transition during heat stress. *Biochem Biophys Res Commun* **239**, 291–297.
20. Asadulghani, Suzuki Y, Nakamoto H (2003) Light plays a key role in the modulation of heat shock response in the cyanobacterium *Synechocystis* sp PCC 6803. *Biochem Biophys Res Commun* **306**, 872–879.
21. Kojima K, Nakamoto H (2007) A novel light- and heat-responsive regulation of the *groE* transcription in the absence of HrcA or CIRCE in cyanobacteria. *FEBS Lett* **581**, 1871–1880.
22. Huckauf J, Nomura C, Forchhammer K, Hagemann M (2000) Stress responses of *Synechocystis* sp. strain PCC 6803 mutants impaired in genes encoding putative alternative sigma factors. *Microbiology (Reading)* **146**, 2877–2889.
23. Imamura S, Yoshihara S, Nakano S, Shiozaki N, Yamada A, Tanaka K, Takahashi H, Asayama M, et al (2003) Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium. *J Mol Biol* **325**, 857–872.
24. Nikkinen HL, Hakkila K, Gunnelius L, Huokko T, Pollari M, Tyystjärvi T (2012) The SigB σ factor regulates multiple salt acclimation responses of the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* **158**, 514–523.
25. Nakamoto H, Suzuki M, Kojima K (2003) Targeted inactivation of the *hrcA* repressor gene in cyanobacteria. *FEBS Lett* **549**, 57–62.
26. Yuan G, Wong SL (1995) Regulation of *groE* expression in *Bacillus subtilis*: the involvement of the sigma A-like promoter and the roles of the inverted repeat sequence (CIRCE). *J Bacteriol* **177**, 5427–5433.
27. Zuber U, Schumann W (1994) CIRCE, a novel heat shock element involved in regulation of heat shock operon dnaK of *Bacillus subtilis*. *J Bacteriol* **176**, 1359–1363.
28. Yuan G, Wong SL (1995) Isolation and characterization of *Bacillus subtilis* groE regulatory mutants: evidence for *orf39* in the *dnaK* operon as a repressor gene in regulating the expression of both *groE* and *dnaK*. *J Bacteriol* **177**, 6462–6468.
29. Schumann W (2003) The *Bacillus subtilis* heat shock stimulon. *Cell Stress Chaperones* **8**, 207–217.
30. Rajaram H, Apte SK (2010) Differential regulation of *groESL* operon expression in response to heat and light in *Anabaena*. *Arch Microbiol* **192**, 729–738.
31. Nakamoto H, Kojima K (2017) Non-housekeeping, non-essential GroEL (chaperonin) has acquired novel structure and function beneficial under stress in cyanobacteria. *Physiol Plant* **161**, 296–310.
32. Nakamoto H (2021) Evolution and diversification of the *groEL*/chaperonin paralogs in cyanobacteria. In: Rastogi RP (ed) *Ecophysiology and Biochemistry of Cyanobacteria*, Springer Nature Singapore, Singapore, pp 181–207.
33. Saito M, Watanabe S, Nimura-Matsune K, Yoshikawa H, Nakamoto H (2020) Regulation of the *groESL1* transcription by the HrcA repressor and a novel transcription factor Orf7.5 in the cyanobacterium *Synechococcus elongatus* PCC 7942. *J Gen Appl Microbiol* **66**, 85–92.
34. Nakamoto H, Tanaka N, Ishikawa N (2001) A novel heat shock protein plays an important role in thermal stress management in cyanobacteria. *J Biol Chem* **276**, 25088–25095.
35. Sato M, Nimura-Matsune K, Watanabe S, Chibazakura T, Yoshikawa H (2007) Expression analysis of multiple *dnaK* genes in the cyanobacterium *Synechococcus elongatus* PCC 7942. *J Bacteriol* **189**, 3751–3758.
36. Rupprecht E, Düppre E, Schneider D (2010) Similarities and singularities of three DnaK proteins from the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol* **51**, 1210–1218.
37. Lund PA (2009) Multiple chaperonins in bacteria – why so many? *FEMS Microbiol Rev* **33**, 785–800.
38. Rippka R, Waterbury J, Cohen-Bazire G (1974) A cyanobacterium which lacks thylakoids. *Arch Microbiol* **100**, 419–436.
39. Honda D, Yokota A, Sugiyama J (1999) Detection of seven major evolutionary lineages in cyanobacteria

based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *J Mol Evol* **48**, 723–739.

40. Rensing SA (2014) Gene duplication as a driver of plant morphogenetic evolution. *Curr Opin Plant Biol* **17**, 43–48.
41. Tanaka N, Hiyama T, Nakamoto H (1997) Cloning, characterization and functional analysis of *groESL* operon from thermophilic cyanobacterium *Synechococcus vulcanus*. *Biochim Biophys Acta* **1343**, 335–348.
42. Kovács E, van der Vies SM, Glatz A, Török Z, Varvasovszki V, Horváth I, Vígh L (2001) The chaperonins of *Synechocystis* PCC 6803 differ in heat inducibility and chaperone activity. *Biochem Biophys Res Commun* **289**, 908–915.
43. McLennan N, Masters M (1998) GroE is vital for cell-wall synthesis. *Nature* **392**, 139.
44. Potnis AA, Rajaram H, Apte SK (2016) GroEL of the nitrogen-fixing cyanobacterium *Anabaena* sp. strain L-31 exhibits GroES and ATP-independent refolding activity. *J Biochem* **159**, 295–304.
45. Sato S, Ikeuchi M, Nakamoto H (2008) Expression and function of a *groEL* paralog in the thermophilic cyanobacterium *Thermosynechococcus elongatus* under heat and cold stress. *FEBS Lett* **582**, 3389–3395.
46. Huq S, Sueoka K, Narumi S, Arisaka F, Nakamoto H (2010) Comparative biochemical characterization of two GroEL homologs from the cyanobacterium *Synechococcus elongatus* PCC 7942. *Biosci Biotechnol Biochem* **74**, 2273–2280.
47. Akter T, Nakamoto H (2021) pH-mediated control of anti-aggregation activities of cyanobacterial and *E. coli* chaperonin GroELs. *J Biochem* **169**, 351–361.
48. Mangan NM, Flamholz A, Hood RD, Milo R, Savage DF (2016) pH determines the energetic efficiency of the cyanobacterial CO_2 concentrating mechanism. *Proc Natl Acad Sci USA* **113**, 5354–5362.
49. Werdan K, Heldt HW, Milovancev M (1975) The role of pH in the regulation of carbon fixation in the chloroplast stroma. Studies on CO_2 fixation in the light and dark. *Biochim Biophys Acta* **396**, 276–292.
50. Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, Stines AP, Georgopoulos C, Frishman D,et al (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* **122**, 209–220.
51. Fujiwara K, Ishihama Y, Nakahigashi K, Soga T, Taguchi H (2010) A systematic survey of *in vivo* obligate chaperonin-dependent substrates. *EMBO J* **29**, 1552–1564.
52. Hongsthong A, Senachak J, Nakamoto H (2017) Genome-and proteome-wide analyses for targeted manipulation and enhancement of bioproducts in cyanobacteria. In: Rastogi R, Madamwar D, Pandey A (eds) *Algal Green Chemistry*, Elsevier, Chennai, India, pp 39–64.
53. Nakamoto H, Honma D (2006) Interaction of a small heat shock protein with light-harvesting cyanobacterial phycocyanins under stress conditions. *FEBS Lett* **580**, 3029–3034.
54. Forchhammer K (2010) The network of P(II) signalling protein interactions in unicellular cyanobacteria. *Adv Exp Med Biol* **675**, 71–90.
55. Chaudhry C, Horwich AL, Brunger AT, Adams PD (2004) Exploring the structural dynamics of the *E. coli* chaperonin GroEL using translation-libration-screw crystallographic refinement of intermediate states. *J Mol Biol* **342**, 229–245.