Pseudo-T-even Bacteriophage RB49 Encodes CocO, a Cochaperonin for GroEL, Which Can Substitute for Escherichia coli’s GroES and Bacteriophage T4’s Gp31*

Debbie Ang‡‡, Alexandra Richardson‡, Matthias P. Mayer‡‡, France Keppel‡, Henry Krisch, and Costa Georgopoulos‡

From the ²Département de Biochimie Médicale, Centre Médical Universitaire, 1 rue Michel-Servet, CH-1211 Genève 4, Switzerland and ³Laboratoire de Microbiologie et Génétique Moléculaire, CNRS-UMR 5100, 118 route de Narbonne, F-Toulouse Cedex 31062, France

Bacteriophage T4-encoded Gp31 is a functional ortholog of the Escherichia coli GroES cochaperonin protein. Both of these proteins form transient, productive complexes with the GroEL chaperonin, required for protein folding and related functions in the cell. However, Gp31 is specifically required, in conjunction with GroEL, for the correct folding of Gp23, the major capsid protein of T4. To better understand the interaction between GroEL and its cochaperonin cognates, we determined whether the so-called “pseudo-T-even bacteriophages” are dependent on host GroEL function and whether they also encode their own cochaperonin. Here, we report the isolation of an allele-specific mutation of bacteriophage RB49, called ε22, which permits growth on the E. coli groEL44 mutant but not on the isogenic wild type host. RB49 ε22 was used in marker rescue experiments to identify the corresponding wild type gene, which we have named cocO (cochaperonin cognate). CocO has extremely limited identity to GroES but is 34% identical and 55% similar at the protein sequence level to T4 Gp31, sharing all of the structural features of Gp31 that distinguish it from GroES. CocO can substitute for Gp31 in T4 growth and also suppresses the temperature-sensitive phenotype of the E. coli groES24 mutant. CocO’s predicted mobile loop is one residue longer than that of Gp31, with the ε22 mutation resulting in a Q36R substitution in this extra residue. Both the CocO wild type and ε22 proteins have been purified and shown in vitro to assist GroEL in the refolding of denatured citrate synthase.

The Escherichia coli GroEL and GroES proteins form a complex, known as the GroE chaperone machine, that is crucial to the functions related to protein folding (reviewed in Refs. 1–4). The GroEL and GroES proteins are often referred to as chaperonin and cochaperonin, respectively (5). Both GroES and GroEL have been shown to be essential for E. coli viability under all conditions tested (6). Extensive structural analysis has revealed that both proteins are organized into rings with a 7-fold rotational axis (7, 8). GroEL is arranged in two head-to-head stacked rings of seven subunits in whose central cavities various substrate proteins can bind, primarily through hydrophobic interactions (Ref. 9; reviewed in Refs. 2 and 4). GroES is a heptameric ring of 10.5-kDa subunits from which mobile loops extend and interact with GroEL (10, 11).

A typical GroEL substrate is first captured in either of the GroEL central cavities, and the subsequent binding of ATP and GroES to the same ring (referred to as the cis ring) results in the formation of a dome-like structure over the substrate, which is now released into the GroEL cis cavity. The subsequent binding of ATP and substrate (and potentially GroES as well) in the trans GroEL ring results in the release of GroES from the cis ring (4, 12). If the substrate in the cis ring is properly folded, or in a form that no longer binds GroEL, it is released into the medium, and this GroEL ring will become the new trans ring in a new GroEL folding cycle. Houry et al. (13) have shown, by using coimmunoprecipitation experiments, that ~10% of E. coli’s polypeptides bind transiently to GroEL. Some of these proteins are encoded by essential genes, in agreement with the genetic observation that the groES and groEL genes are both essential for E. coli growth (6).

The bacteriophage T4, which is dependent on the host GroEL function for folding of its major capsid protein, Gp23, encodes its own GroES ortholog, termed Gp31. There is very little sequence similarity between GroES and Gp31 (14, 15), although there is a limited conservation in certain regions, most notably the mobile loop that becomes immobilized upon binding to GroEL (10, 16, 17). The functional importance of the mobile loop was first shown by sequencing mutations in groES that blocked bacteriophage λ growth (λ requires both GroEL and GroES host functions for growth (10, 18)). Similarly, T4 gene 31 mutations, which allow T4 to propagate on certain E. coli groEL mutants, also affect the Gp31 mobile loop (14, 19, 20).

Comparative studies of GroES and Gp31 have shown that while Gp31 can substitute for GroES, the reverse is not true; i.e. Gp31 can suppress the temperature-sensitive phenotype of the E. coli groES24 mutant and can participate in the in vitro folding of various substrates, but GroES is unable to substitute for Gp31 in T4 bacteriophage assembly, even when overproduced (21). This may not be totally surprising, since it is likely that Gp31 evolved as a specialized ortholog of GroES that specifically folds the Gp23 substrate. Gp31 is 56 kDa, near the maximum limit allowed by the “Anfinsen cage” formed under the dome of the GroEL-GroES complex (reviewed in Refs. 3 and...
Bacterial strains (Table I). The various strains, bacteriophages, and plasmids used in this study are listed in Table I.

**Genotype/phenotype Source or reference**

| Bacterial strains | Genotype/phenotype | Source or reference |
|--------------------|--------------------|---------------------|
| B178               | W3110 galE sup+    | Our collection       |
| CR63               | supD60             | Ref. 26             |
| DA1368             | MC1061 endAΔ::tetR | Refs. 27 and 28     |
| CG3014             | B178 groES+ groEL+ tetR nearby | Ref. 29 |
| CG3010             | B178 groEL4A tetR nearby | Ref. 29 |
| CG3009             | B178 groEL515 tetR nearby | Ref. 29 |
| CG3012             | B178 groEL40 tetR nearby | Ref. 29 |
| CG3013             | B178 groES819 tetR nearby | Ref. 10 |
| DA1415             | B178 groES42 tetR nearby | Ref. 10 |

**Bacteriophages**

| T4Do               | Wild type isolate | Our collection       |
|--------------------|--------------------|---------------------|
| T4 31amNG71        | Amber mutation in gene 31 | Ref. 30 |
| T4 31amB17         | Amber mutation in gene 23 | Ref. 30 |
| T4 31e1            | Grows on E. coli groEL44 | Ref. 31 |
| RB49               | Wild type isolate | Ref. 25 |
| RB49 cocOamE45     | Amber mutation in cocO | This paper |
| (= RB49 amE45)     | Grows on E. coli groEL44 | This paper |
| RB49 cocOe22       | (= RB49 e22)       | This paper |

**Plasmids**

| pMPM401           | ColE1 ori, ApR, SpeI, PBAD | Ref. 32 |
| pMPM41            | pMPM401 with t-SpeI deleted | Ref. 32 |
| pMPM201           | pMPM201 (31'), ApR        | Ref. 32 |
| p31KO             | pMMA4 (cocO'), isolated from library | This paper |
| pFScocO+          | pMMA201 (cocO'), minimal gene | This paper |
| pFScocOe22        | pMMA201 (cocOe22), minimal gene | This paper |

**EXPERIMENTAL PROCEDURES**

4). Structural studies suggest that the GroEL-Gp31 complex may result in the formation of a larger cavity under the dome of the Gp31 heptamer due to a longer mobile loop, the lack of the roof loop present in GroES, and lack of a conserved aromatic residue, which extends into the cavity of GroES (22, 23).

Recently, a large number of bacteriophages with T4-like morphology have been characterized according to their sequence homology with the classical T-even bacteriophages, T2, T4, and T6 (24). The vast majority of these bacteriophages are very closely related to T-evens at the DNA sequence level. However, a few of them have DNA sequences that differ significantly, and these have been termed “pseudo-T-even bacteriophages” (25). We investigated the genomes of the pseudo-T-even bacteriophages to determine whether they possess a gene encoding a homolog of either GroES or Gp31. Here, we report our findings on the bacteriophage RB49 gene 31 homolog, which we have named *cocO* (cochaperonin cognate).

**DNA Library Construction**—One liter of an RB49 lysate (2 x 10⁹ pfu/ml) was concentrated by polyethylene glycol precipitation (34), followed by CsCl banding on a step gradient. The concentrated bacteriophage was dialyzed for 30 min in 500 ml 3 × NaCl, 0.1 × Tris, pH 7.4, followed by two 30-min dialyses in 500 ml of 0.3 × NaCl, 0.1 × Tris, pH 7.4. The dialyzed bacteriophage was treated with proteinase K, followed by phenol/chloroform extraction (34). The DNA was partially digested using Sau3A (New England Biolabs), and fragments of 1.5–5 kb were purified from an agarose gel. The fragments were ligated to the vector pMPM4A (a derivative of pMPM4H) (32) in which digestion by XbaI was used to remove the Ω fragment encoding spectinomycin resistance), which had been digested with BglII. The ligation reaction was transformed into competent DA1368 (MC1061 endAΔ::tetR) bacteria, and transformants were selected on LB-agar-ampicillin plates at 37 °C.

**Marker Rescue**—Cells growing exponentially at a density of A₅₇₀ = 0.1 were infected with the appropriate bacteriophage at a multiplicity of infection of 0.02 bacteriophage/bacterium. The cultures were aerated at 37 °C until lysed, ~2 h. Chloroform was added to ensure lysis and the killing of all remaining cells. To check for wild type recombinants, 0.1 ml of the lysate was mixed with 0.3 ml of a saturated culture of groEL40 cells and 3 ml of molten top agar and then poured onto an LB-agar plate. The plates were incubated at 37 °C overnight to allow plaque formation.

**DNA Sequencing**—Chain termination reactions were performed using the ThermoSequenase kit (Amersham Pharmacia Biotech), and the reactions were sequenced on a Li-Cor sequencing apparatus (MWG Biotech). The primers used for sequencing and polymerase chain reaction amplification are as follows: cocO pBAD 5', 5'-CCGACTTCC- TACTGTGTTAATGATGAGCAGTAACTGCAG-3'; cocO 5' RUNd1, 5'-GGGAT- CTTACATGAGCCGTTAAACCTCCTAGAAGAG-3'; cocO 3' XhoI, 5'-GCTC TAGAGGATTATTCCGCGCTTAAA-3'. The resulting plasmid was named pFScocO+.

**Cloning of the Minimal *cocO* Gene for Overexpression**—The cocO 5' RUNd1 and cocO 3' XhoI primers were used to polymerase chain reaction-amplify the minimal *cocO* gene from wild type RB49 bacteriophage. The product was cloned into the EcoRI and XhoI sites of the pMPM201 vector under the regulation of the pBAD promoter as previously described (32). The resulting plasmid was named pFScocO+ (34). The cocOe22 allele was cloned in the same manner, using RB49 cocOe22 DNA as template, resulting in plasmid pFScocOe22.

1 The abbreviations used are: tetR, tetracycline-resistant; HPLC, high pressure liquid chromatography.
Purification of the CocO Protein—Six 1.5-liter cultures inoculated with DA1368 bacteria containing plasmid pSFcocO’ were grown in LB plus 100 μg/ml ampicillin at 37 °C. Synthesis of CocO was induced at A660 ~0.6 with 0.02% arabinose for 4 h. Cells were harvested by centrifugation and washed once in buffer containing 50 mM Tris, pH 7.5, 1 mM EDTA, 5 mM β-mercaptoethanol, and 15% (v/v) glycerol. The pellet was stored at −20 °C and gently thawed before lysis. Cells were resuspended in buffer A (10 mM Tris, pH 7.7, 100 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol) and lysed by two passes through a French press at 1,000 p.s.i. The following steps are adapted from the purification procedure used for Gp31 (21). After removal of the debris by low speed centrifugation (10 min at 10,000 × g), followed by ultracentrifugation for 60 min at 145,000 × g, the supernatant was fractionated by a streptomycin sulfate precipitation. The CocO eluting supernatant was further fractionated by a 35% (w/v) ammonium sulfate precipitation. Again, CocO remained in the supernatant. CocO was dialyzed into buffer B (20 mM Tris, pH 7.7, 1 mM EDTA, 1 mM β-mercaptoethanol) and applied to a Q-Sepharose column, which binds acidic proteins. Because of its neutral pI, CocO eluted in the flow-through. After a 70% (w/v) ammonium sulfate precipitation, CocO was resuspended and dialyzed into buffer C (20 mM sodium phosphate, pH 6.8, 5 mM β-mercaptoethanol, 1 mM EDTA) and loaded onto an hydroxyapatite column to which a 20 to 0.5 M sodium phosphate, pH 6.8, gradient was applied. CocO eluted at ~0.2 M sodium phosphate. After exchanging into buffer D (100 mM Tris, pH 7.5, 1 mM EDTA, 1 mM β-mercaptoethanol), final impurities were removed by gel filtration (HiLoad 16/60 Superdex G200 on a Waters HPLC). CocO eluted at 16 min, suggesting an approximate native molecular mass of 80,000 Da.

Purification of the mutant CocO22 protein was executed following its overproduction from plasmid pFScocO22, using the exact steps described for the wild type protein.

Refolding of Citrate Synthase—The procedure for chaperonin-dependent renaturation of pig heart citrate synthase has been described previously (35). Briefly, citrate synthase at 33 °C was denatured in 6 M guanidine hydrochloride, 3 mM dithiothreitol, and 2 mM EDTA for 30 min at 25 °C. Denatured citrate synthase was diluted to 0.2 μM into a renaturation mix containing 20 mM potassium phosphate, pH 7.4, 10 mM MgCl2, 2 mM ATP, 1 mM oxaloacetic acid and various combinations of the chaperonins at 4.2 μM (concentration given for monomers) as indicated under “Results.” The refolding reactions were performed at 27 °C in a total volume of 400 μl, and citrate synthase activity was measured after 60 min. The enzyme activity was assayed by measuring the decrease in absorption at 232 nm due to the cleavage of acetyl CoA measured after 60 min. The enzyme activity was assayed by measuring its overproduction from plasmid pFScocO22, using the exact steps described for the wild type protein.

TABLE II

| CG3014 wild type | CG3010 groEL44 | CG3009 groEL515 | CG3012 groEL140 | DA1415 groES42 |
|-----------------|---------------|-----------------|-----------------|---------------|
| 37 °C | 42 °C | 37 °C | 42 °C | 37 °C | 42 °C | 37 °C | 42 °C | 37 °C | 42 °C |
| RB49 | + | + | + | + | + | + | + | - | + |
| RB49e22 | + | + | + | + | + | + | + | - | + |
| T4Do | + | + | + | + | + | + | + | - | + |
| T4 31e1 | + | + | + | + | + | + | + | - | + |

We obtained further evidence for a Gp31 homolog in RB49 by showing that the T4 gene 31 product complements the e22 mutation (Table III), allowing bacteriophage RB49 e22 growth on the restricting E. coli wild type host at 37 °C. In contrast, a plasmid overexpressing both the E. coli groEL and groES wild type genes does not complement the e22 mutation (data not shown). Earlier results had shown that overproduction of groES also does not complement the T4 31e1 defect (21).

Cloning and Genetic Properties of cocO—To identify the RB49 gene where the e22 mutation resides, we constructed an RB49 DNA library in the multicopy plasmid pMPMA4. Although the construction of the RB49 DNA library was initiated with ~1.5–5.0-kb-long DNA fragments, the resulting library was enriched in small DNA insert fragments, probably because of the toxicity of many of the RB49 DNA sequences in E. coli.

We transformed the entire library into the E. coli groEL44 mutant, which is permissive for RB49 e22 and used this library in a marker rescue scheme. Since it was unclear whether the wild type clone corresponding to e22 would be underrepresented in our RB49 DNA library, we started with an initial pool of ~10^5 independent groEL44 transformants obtained with plasmid DNA purified from the original amplified library (itself consisting of ~10^6 independent clones). This pool was divided into 45 groups each containing ~2,000 transformants and allowed to amplify. RB49 e22 was grown on these 45 groups of transformants, and 0.1 ml of each lysate was plated on the nonpermissive groEL140 mutant strain at 37 °C (Table II) to detect the presence of RB49 wild type recombinants or revertants. Whereas most of the transformant pools gave fewer than 20 plaques per plate, five of them gave significantly higher numbers, ranging from 50 to 200. The marker rescue scheme was repeated (five independent lysates each) for two of these pools, confirming that the elevated number of plaque formers (putative wild type recombinants) was not due to statistical fluctuation. One of these pools was chosen for further screening and, by repeating the marker rescue experiment with progressively smaller pools of transformants (250, 10, and finally individual transformants), we progressively obtained higher frequencies of RB49 wild type recombinants. Thus, we were able to identify the in vivo biological effects of this mutation.

Isolation of RB49 Mutations in the Putative Gene 31 Homolog—In analogy with the isolation of the bacteriophage T4 31e1 mutant (23, 31), spontaneous mutants of RB49 can be isolated at a frequency of ~5 × 10^-6 that grow on groEL44 mutant bacteria. One such isolate, RB49 e22, forms wild type-size plaques on the E. coli groEL44 mutant host at 37 °C (Table II). It was found that RB49 e22 cannot form plaques on either wild type, groEL140, or groES42 hosts at 37 °C; however, it does form plaques on these same hosts at 42 °C. For comparison, we include the T4 mutant T4 31e1 in Table II, which was originally isolated as a spontaneous plaque former on groEL44 (31). Previous analysis has shown that the T4 31e1 mutation results in an L35I substitution in the middle of the mobile loop of Gp31 (14, 19, 20).

Plating properties of various bacteriophages on various bacterial hosts at 37 and 42 °C

| Plaques | 37 °C | 42 °C |
|---------|------|------|
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |
| 0.6 | BD44 | BD44 |

Each of the indicated bacterial strains was seeded in molten top agar on LB-agar plates. Stocks of the four indicated bacteriophage strains were serially diluted 10-fold, and 5 μl of each dilution was spot-tested on each strain. The plates were incubated overnight at 37 or 42 °C, and the plaques were counted. The approximate efficiency of plating of each bacteriophage strain on a particular bacterial host is shown by either a plus sign, indicating an efficiency of plating of approximately 1.0, or a minus sign, indicating an efficiency of plating of less than 10^-6.

RESULTS

Dependence of RB49 on groEL—Both polymerase chain reaction amplification and DNA hybridization experiments with the pseudo-T-even bacteriophage RB49 suggested that this bacteriophage has no gene 31 homolog with significant DNA sequence similarity to the T4 gene (data not shown). Thus, we determined if RB49 growth is dependent on the host GroEL and GroES functions. We found that, like bacteriophage T4, bacteriophage RB49 is totally dependent on the E. coli GroEL function for successful completion of its life cycle but is not dependent on GroES function. Table II shows that wild type RB49, which plates with an efficiency of 1.0 on the wild type host, does not form plaques on either the E. coli groEL44 or groEL515 mutant hosts but does propagate on the groEL140 and groES42 hosts. Since bacteriophage T4 plaques normally on groEL515, it appears that RB49 is more sensitive to the integration of the groEL515 gene than T4.
TABLE III

| T4D0 | CG3014 | CG3014 (p31<sup>+</sup>) | CG3014 (pocoO<sup>34-2</sup>) | CR63 |
|------|--------|--------------------------|-----------------------------|-----|
|      | +      | +                        | +                           | +   |
| T4 31amNG71 | -      | -                        | -                           | -   |
| T4 23amB17 | +      | +                        | +                           | +   |
| T4 31e1 | -      | +                        | +                           | +   |
| RB49 | +      | +                        | +                           | +   |
| RB49 cocoE22 | +      | +                        | +                           | +   |
| RB49 cocoOamE45 | -      | -                        | -                           | -   |

Ability of plasmid-encoded Gp31 and CocO proteins to specifically complement the growth of bacteriophage 31 or cocO mutants

The wild type *E. coli* strain CG3014 was transformed with either plasmid p31<sup>+</sup>, encoding the T4 31 gene, or pocoO<sup>34-2</sup>, encoding the RB49 cocoO gene. See the legend to Table II for a description of the plating conditions and explanation of symbols. The plating experiments were carried out at 37 °C. The CR63 strain carries the supD suppressor, enabling bacteriophages that carry amber (am; nonsense) mutations to plaque.

| Q36R |   |   |   |   |
|------|---|---|---|---|
|      | 10 | 20 | 30 | 40 | 50 |
| Coco9 | MTW | KF | LGA | DRY | Y11 | E45 | 31amGp31 | 34–2 | 11 1 1 | 1 34–2 plasmid enable normal phenotype of this mutation clearly indicates that the last eight amino acid residues (residues 79–90 in Coco9) conserved between the two bacteriophage-encoded cochaperonins but absent in GroES and the rest of its cochaperonin homologs. In the three-dimensional structure of Gp31, this region constitutes an extra loop extending down from the external face of the Gp31 domain (22).

Alignment of the GroES, Gp31, and CocO amino acid sequences reveals a stretch of 12 amino acid residues (residues 79–90 in Coco9) conserved between the two bacteriophage-encoded cochaperonins but absent in GroES and the rest of its cochaperonin homologs. In the three-dimensional structure of Gp31, this region constitutes an extra loop extending down from the external face of the Gp31 domain (22).

**Fig. 1.** Suppression of the *groES42* temperature-sensitive growth phenotype by the RB49 cocoO gene. Strain DA1415 (*groES42*) was transformed with either plasmid pMPMA4 (control), p31<sup>+</sup> (expressing wild type Gp31), or pocoO<sup>34-2</sup> (expressing wild type CocO). See "Experimental Procedures" for details.

**FIG. 2.** Comparison between Gp31, CocO, and GroES predicted amino acid sequences. The alignment between Gp31 and GroES is taken from the article by Hunt *et al.* (22). Alignment between CocO and GroES1 was performed using BLAST (36).

| Q36R |   |   |   |   |
|------|---|---|---|---|
|      | 10 | 20 | 30 | 40 | 50 |
| Coco9 | MTW | KF | LGA | DRY | Y11 | E45 | 31amGp31 | 34–2 | 11 1 1 | 1 34–2 plasmid enable normal phenotype of this mutation clearly indicates that the last eight amino acid residues (residues 79–90 in Coco9) conserved between the two bacteriophage-encoded cochaperonins but absent in GroES and the rest of its cochaperonin homologs. In the three-dimensional structure of Gp31, this region constitutes an extra loop extending down from the external face of the Gp31 domain (22).

Alignment of the GroES, Gp31, and CocO amino acid sequences reveals a stretch of 12 amino acid residues (residues 79–90 in Coco9) conserved between the two bacteriophage-encoded cochaperonins but absent in GroES and the rest of its cochaperonin homologs. In the three-dimensional structure of Gp31, this region constitutes an extra loop extending down from the external face of the Gp31 domain (22).

**Fig. 1.** Suppression of the *groES42* temperature-sensitive growth phenotype by the RB49 cocoO gene. Strain DA1415 (*groES42*) was transformed with either plasmid pMPMA4 (control), p31<sup>+</sup> (expressing wild type Gp31), or pocoO<sup>34-2</sup> (expressing wild type CocO). See "Experimental Procedures" for details.

**Fig. 2.** Comparison between Gp31, CocO, and GroES predicted amino acid sequences. The alignment between Gp31 and GroES is taken from the article by Hunt *et al.* (22). Alignment between CocO and GroES1 was performed using BLAST (36).

**Purification of CocO Protein**—The CocO wild type protein was overproduced and purified from bacteria carrying the pFScocoO<sup>+</sup> plasmid (see “Experimental Procedures” for details). Similarly, the CocO<sub>22</sub> mutant protein was overproduced and purified from bacteria carrying the pFScocoO<sub>22</sub> plasmid. Both proteins were at least 90% pure, as judged by staining of the purified protein preparations following separation by SDS-polyacrylamide gel electrophoresis (Fig. 3A). The elution profiles of both proteins during HPLC gel filtration suggest an
CocO, a Gp31 Homolog

**Functional Analysis of the CocO Wild Type and CocOe22 Proteins**—To test whether the purified CocO wild type and CocOe22 proteins are active and capable of interacting with their GroEL partner, we carried out *in vitro* folding assays with denatured citrate synthase. Previously, it was established that the correct *in vitro* folding of pig heart citrate synthase requires both the GroEL and GroES proteins (37, 38). As shown in Fig. 3B, the CocO cochaperonin is indeed capable of assisting wild type GroEL in the folding of denatured citrate synthase. Even more convincing are the observations that (a) the wild type CocO protein cannot assist the GroEL44 mutant protein in citrate synthase folding, thus reproducing the *in vivo* inability of bacteriophage RB49 to propagate on the *groEL44* mutant host (Table II), and (b) the CocOe22 mutant protein restores the ability of the GroEL44 mutant protein to fold citrate synthase, again in agreement with the *in vivo* result that the bacteriophage RB49 cocOe22 mutant grows perfectly well on the *groEL44* mutant host (Table II). These experiments have been performed using various chaperonin/cochaperonin ratios, with essentially the same results (data not shown). In contrast to the *in vivo* bacteriophage plating phenotype (Table II), the CocOe22 mutant protein was capable of assisting the wild type GroEL protein in citrate synthase folding. This result is not entirely unexpected, since bacteriophage RB49 *cocOe22* can form plaques on CG3014 wild type bacteria at high temperatures, thus demonstrating significant interaction between GroEL wild type and CocOe22 proteins, at least at high temperature (Table II). Most likely, citrate synthase, although a relatively large substrate (monomer is 51,629 Da), is not as problematic to fold as the RB49 major capsid protein.

**DISCUSSION**

Although the x-ray crystallographic structures of the GroEL chaperonin and its GroES cochaperonin have been solved, both as a complex and as individual proteins, many details of their interaction with each other and their substrates are not understood (2, 7, 8, 11). The primary contact that GroES makes with GroEL has been localized to a 16-amino acid residue segment of GroES. This so-called mobile loop, although relatively unstructured in free GroES, adopts a β-hairpin structure upon binding to GroEL (10). In the crystal structure, the only direct contact that GroES makes with GroEL is with a conserved, hydrophobic tripeptide located in the middle of its mobile loop (10, 11). The contributions of other elements to both the initial binding of GroES to GroEL and to the subsequent stabilization of the GroEL-GroES complex are unknown. Another detail that remains to be clarified is whether GroES binding to the *trans* ring of GroEL can take place while GroES is still bound to the *cis* ring (“football”) and whether such binding may be required for the efficient release of some substrates from the *cis* ring of GroEL.

Genetic and biochemical studies have identified the bacteriophage T4-encoded Gp31 protein (product of gene 31) as essential for the correct folding of Gp23 (product of gene 23), the major bacteriophage capsid protein (39–42). In the absence of a functional gene 31, the Gp23 capsid protein aggregates in lumps on the *E. coli* inner membrane (39). The realization that some *groEL* mutations (but none of the *groES* mutations) block bacteriophage T4 morphogenesis by interfering with the action of Gp31 led to the proposal that Gp31 may serve as a more specialized GroES-like homolog, capable of assisting the correct folding of Gp23 by GroEL (14, 43) despite the fact that the GroES and Gp31 proteins share little sequence identity (~14%; Refs. 14–16). Gp31 is indeed a cochaperonin for GroEL; it substitutes for GroES in protein folding, binds GroEL, and modulates its ATPase activity as GroES does and even suppresses the temperature-sensitive phenotype of *groES* mutations (21). But only Gp31, and not GroES, can assist in correctly folding Gp23 *in vivo* (21, 44).

What then is the feature(s) of Gp31 that distinguishes it from GroES? Hunt et al. (22) emphasized four structural features that distinguish Gp31 from GroES and all of the other GroES-like cochaperonins. First, since Gp31’s mobile loop is substantially longer (22 versus 16 amino acids) compared with that of GroES (10, 19), this may result in the formation of a “taller” dome structure over GroEL. Such a larger space in the GroEL-Gp31 cavity could better accommodate Gp23, a relatively large
substrate at 56 kDa. A large mobile loop is also a feature of the CocO cochaperonin (Fig. 2), which is actually one amino acid longer than that of Gp31. The e22 mutation (Q36R) that results in an efficient interaction with the GroEL44 mutant chaperonin actually alters this extra residue in CocO. Previous biochemical analyses have shown that Gp31 and GroEL44 do not efficiently form a complex in vitro (20). Thus, CocO and GroEL44 could suffer from a similar defect, which is overcome by the Q36R substitution of CocO. In this respect, it is interesting that the only mutation of bacteriophage T4 gene 31 that is capable of restoring substantial interaction with GroEL44 is the T4 31e1 mutation, an L35I substitution in Gp31’s mobile loop (23). This substitution results in a more hydrophobic amino acid (isoleucine 35 for leucine) in the first position of the hydrophobic tripeptide of Gp31. In contrast, the Q36R substitution in CocO’s mobile loop enables it to interact even better with GroEL44, as judged by the fact that T4 31e1 (L35I) grows on groEL44 hosts at 37 °C but not at 42 °C, while RB49 cocOe22 (Q36R) grows on groEL44 bacteria at both 37 and 42 °C (Table II). The Q36R substitution is located two amino acid residues C-terminal to the conserved hydrophobic tripeptide and may bypass the GroEL44 block by a different mechanism than the L35I substitution of Gp31.

Hunt et al. (22) also noted the presence of a universally conserved aromatic amino acid residue at position 71 in GroES and its homologs. This residue juts into the central cavity of GroES and limits the volume available within the GroEL-GroES complex for substrate folding. This aromatic amino acid is absent from the equivalent position in both Gp31 and the CocO cochaperonin (Fig. 2). Another feature that distinguishes both Gp31 and CocO from GroES and the other cochaperonins is the lack of a “roof loop” structure at the top of the GroES dome (Fig. 2). Again, this difference could allow Gp31 and CocO to fold a larger substrate, such as Gp23. Finally, both Gp31 and CocO possess an extra loop at their C-terminal region that is absent in GroES and all other GroES-like cochaperonins. In the Gp31 crystal structure, this extra loop domain is located outside of the dome. Although it is not clear what role it plays in Gp31 and CocO cochaperonin function, deletion of this loop domain abolishes Gp31 specificity for Gp23 but not for its E. coli protein substrates (45).

Despite the various structural differences between the bacteriophage-encoded Gp31(CocO) cochaperonin class and the GroES-type cochaperonins, the molecular mechanism(s) that underlies the unique capacity of the Gp31(CocO) cochaperonin to fold the Gp23 capsid protein remains enigmatic. The T-even bacteriophages assemble an icosahedral head structure composed of 960 Gp23 subunits (46). If one assumes a burst size of 200 bacteriophage particles per infected bacterium, then ~200,000 Gp23 subunits must mature during the short time (10–20 min at 37 °C) available for virion morphogenesis. In addition to being relatively large, the Gp23 capsid protein also has a tendency to aggregate, as judged by the formation of lumps in the absence of GroEL (39). Thus, Gp23 may necessitate many cycles of GroEL binding and release for its proper maturation. It could be that the Gp31(CocO)-GroEL folding cycle is substantially shorter than that of GroES-GroEL, thus favoring the fast and efficient maturation of Gp23. Another possibility is that the Gp31(CocO) cochaperonin can bind to the trans ring of GroEL more efficiently than GroES and, in doing so, accelerate the release of Gp23 from the cis ring. Finally, the size and biochemical properties of the Gp31(CocO)-GroEL cavity could be specifically customized for Gp23 folding. Of course, none of these possibilities are mutually exclusive, and they could all contribute to the correct and rapid maturation of Gp23. Whatever this difference between Gp31 (CocO) and GroES may be, it is clear that the Gp31(CocO) cochaperonin class is capable of folding all of E. coli’s essential GroEL substrates, since either Gp31 or CocO can totally substitute for GroES in bacterial growth.2 Thus, Gp31 (CocO) must maintain all features necessary for the correct folding of many polypeptide substrates. Clearly, more experiments are necessary, such as domain swap experiments, to pinpoint the Gp31 (CocO) features that enable preferential Gp23 folding.

One aspect of this work deserves further comment. Here, we have studied the structure and function of the Gp31 protein of an evolutionarily distant T4-type bacteriophage and compared its properties with those of the T4 homolog and the E. coli ortholog. There is extremely little sequence homology between bacteriophage T4 Gp31 and E. coli GroES, although their cochaperonin functions are very similar. More sequence homology is apparent between GroES and RB49 CocO sequences, particularly in the N-terminal portion of these proteins. As Gp31 homologs are analyzed from additional distant T4-type bacteriophages, we anticipate that their evolutionary relationship to the GroES protein will become increasingly obvious. Such an expanded phylogenetic analysis of Gp31 could also provide valuable information about the motifs and domains that are critical for function and thus conserved.

Acknowledgments—We thank Eric Morency and “La science appelle les jeunes” for their catalytic role in isolating the RB49 cocOe45 mutant; Françoise Schwager for excellent technical assistance; and Luli Billecchi-Mestre for cheerful editing of the manuscript.

REFERENCES

1. Georgopoulos, C. & Welch, W. J. (1993) Annu. Rev. Cell Biol. 9, 601–635
2. Lorimer, G. (1997) Nature 388, 720–723
3. Bukau, B. & Horwich, A. L. (1996) Cell 92, 351–366
4. Sigler, P. B., Xu, Z., Rye, H. S., Burton, B. G., Fenton, W. A. & Horwich, A. L. (1996) Annu. Rev. Biochem. 67, 581–608
5. Ellis, R. J. & van der Vies, S. M. (1996) Annu. Rev. Biochem. 65, 321–347
6. Fayet, O., Zieglerhoff, T. & Georgopoulos, C. (1989) J. Bacteriol. 171, 1297–1305
7. Braig, K., Otwonowski, Z., Hegde, R., Boivest, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994) Nature 371, 578–586
8. Hunt, J. F., Weaver, A. J., Landry, S. J., Giersch, L. & Deisenhofer, J. (1996) Nature 379, 37–45
9. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994) Nature 371, 611–619
10. Landry, S. J., Zelista-Rylls, J., Fayet, O., Georgopoulos, C. & Giersch, L. M. (1993) Nature 364, 255–258
11. Xu, Z., Horwich, A. L. & Sigler, P. B. (1997) Nature 385, 741–750
12. Rye, H. S., Rosenman, A. M., Chen, S., Furtak, R., Fenton, W. A., Saibil, H. R. & Horwich, A. L. (1999) Cell 97, 325–328
13. Houry, W. A., Frishman, D., Eckerskorn, C., Lottespeich, F. & Hartl, F. U. (1999) Nature 402, 147–154
14. Keppel, F., Lipinska, B., Ang, D. & Georgopoulos, C. (1996) Gene (Amst.) 90, 19–25
15. Nivinskas, R. & Black, L. W. (1988) Gene (Amst.) 73, 251–257
16. Koomin, R. V. & van der Vies (1993) Trends Biochem. Sci. 20, 14–15
17. Richardson, A., Landry, S. J. & Georgopoulos, C. (1998) Trends Biochem. Sci. 23, 138–143
18. Tiley, K., Muriadko, H. & Georgopoulos, C. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 1629–1633
19. Landry, S. J., Taher, A., Georgopoulos, C. & van der Vies, S. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 11622–11627
20. Richardson, A., van der Vies, S., Keppel, F., Taher, A., Landry, S. J. & Georgopoulos, C. (1999) J. Biol. Chem. 274, 52–58
21. van der Vies, S. M., Gatenby, A. A. & Georgopoulos, C. (1994) Nature 368, 654–656
22. Hunt, J. F., van der Vies, S. M., Henry, L. & Deisenhofer, J. (1997) Cell 90, 361–371
23. Richardson, A. & Georgopoulos, C. (1999) Genetics 14, 1449–1457
24. Repolla, F., Téteard, F., Bouet, J.-Y. & Krisch, H. M. (1994) EMBO J. 13, 4181–4192
25. Morel, C., Repolla, F., Kutateladze, M., Téteard, F. & Krisch, H. M. (1997) J. Mol. Biol. 267, 257–259
26. Brenner, S. & Beckwith, J. R. (1985) J. Mol. Biol. 13, 629–637

2 F. Keppel and C. Georgopoulos, unpublished results.
CocO, a Gp31 Homolog

27. Meissner, P. S., Sisk, W. S. & Berman, M. L. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4171–4175
28. Cherepanov, P. P. & Wackernagel, W. (1995) Gene (Amst.) 158, 9–14
29. Zeilstra-Ryalls, J., Fayet, O., Baird, L. & Georgopoulos, C. (1993) J. Bacteriol. 175, 1134–1141
30. Epstein, R. H., Bolle, A., Steinberg, C. M., Kellenberger, E., Bey de la Tour, E. & Chevalley, R. Edgar, R.S., Susman, M., Denhardt, G. H., and Lielausis, A. (1963) Cold Spring Harbor Symposia on Quantitative Biology, vol XXVIII, Synthesis and Structure of Macromolecules, pp. 375–394, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
31. Georgopoulos, C., Hendrix, R. W., Kaiser, A. D. & Wood, W. B. (1972) Nat. New Biol. 239, 38–41
32. Mayer, M. P. (1995) Gene (Amst.) 163, 41–46
33. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995) J. Bacteriol. 177, 4121–4130
34. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, p. 85, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
35. Zhi, W., Srere, P. A. & Evans, C. T. (1991) Biochemistry 30, 9281–9286
36. Tatusova, T. A. & Madden, T. L. (1999) FEMS Microbiol. Lett. 174, 247–250
37. Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R., Rudolph, R., Schmid, F. X. & Kiefhaber, T. (1991) Biochemistry 30, 1586–1591
38. Zhi, W., Landry, S. J., Gierasch, L. M. & Srere, P. A. (1992) Protein Sci. 1, 522–529
39. Laemmli, U. K., Begaun, F. & Gujer-Kellenberger, G. (1970) J. Mol. Biol. 47, 69–85
40. Castillo, C. J. & Black, L. W. (1978) J. Biol. Chem. 253, 2132–2139
41. Simon, L. D. & Randolph, B. (1984) J. Virol. 51, 321–328
42. Doermann, A. H. & Simon, L. D. (1984) J. Virol. 51, 315–320
43. Georgopoulos, C., Ang, D., Liberek, K. & Zylicz, M. (1990) in Stress Proteins in Biology and Medicine, (Morimoto, R. I., Tissieres, A., and Georgopoulos, C., eds), pp. 191–221, Cold Spring Harbor Press, Cold Spring Harbor, NY
44. Andreadis, J. D. & Black, L. (1998) J. Biol. Chem. 273, 34075–34086
45. Richardson, A. (2000) Function/Structure Studies of the Bacteriophage T4-Encoded Co-chaperonin, Gp31, with Emphasis on Co-chaperonin Mobile Loop Interactions with GroEl, Ph.D. Thesis, University of Geneva
46. Baschong, W., Baschong-Presianotto, C., Engel, A., Kellenberger, E., Lustig, A., Reichelt, R., Zulauf, M. & Aebi, U. (1991) J. Struct. Biol. 106, 93–101