doi:10.1016/j.jmb.2005.12.005 J. Mol. Biol. (2006) 357, 411–426

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

GroEL Walks the Fine Line: The Subtle Balance of Substrate and Co-chaperonin Binding by GroEL. A Combinatorial Investigation by Design, Selection and Screening

Martin Kawe and Andreas Plückthun*

Biochemisches Institut
Universität Zürich
Winterthurerstrasse 190
CH-8057 Zürich, Switzerland

While support in protein folding by molecular chaperones is extremely efficient for endogenous polypeptides, it often fails for recombinant proteins in a bacterial host, thus constituting a major hurdle for protein research and biotechnology. To understand the reasons for this difference and to answer the question of whether it is feasible to design tailor-made chaperones, we investigated one of the most prominent bacterial chaperones, the GroEL/ES ring complex. On the basis of structural data, we designed and constructed a combinatorial GroEL library, where the substrate-binding site was randomized. Screening and selection experiments with this library demonstrated that substrate binding and release is supported by many variants, but the majority of the library members failed to assist in chaperonin-mediated protein folding under conditions where spontaneous folding is suppressed. These findings revealed a conflict between binding of substrate and binding of the co-chaperonin GroES. As a consequence, the window of mutational freedom in that region of GroEL is very small. In screening experiments, we could identify GroEL variants slightly improved for a given substrate, which were still promiscuous. As the substrate-binding site of the GroEL molecule overlaps strongly with the site of cofactor binding, the outcome of our experiments suggests that maintenance of cofactor binding affinity is more critical for chaperonin-mediated protein folding than energetically optimized substrate recognition.

*Corresponding author

Keywords: combinatorial library; chaperonins; protein folding; GroEL; GroES

Introduction

Protein folding is a key process in biology: it is the three-dimensional structure that ultimately endows proteins with all their functionalities. Whereas this process is very efficient *in vitro* under optimized (also denoted permissive) folding conditions for many proteins, the situation in a living cell is more challenging (denoted non-permissive folding conditions). In particular, the high intracellular concentration of protein and high growth temperature promote aggregation as an undesired side-reaction, competing with productive folding.[1] It is for this reason that all living organisms have evolved diverse classes of helper molecules that serve to prevent misfolding and aggregation of proteins in the crowded environment of the cell.[2] Nevertheless, a large fraction of recombinant proteins still fails to reach the native state and end up in inclusion bodies. The production of natively folded proteins has thus remained a serious bottleneck in basic research, notably structural biology, and biotechnology alike.

Many of these "helper" proteins, generally referred to as molecular chaperones, are heat-shock or stress proteins, but they fulfill essential functions also under normal cellular conditions. One prominent example of these molecular chaperones comprises the family members of chaperonins,[3] which are found in virtually all known organism.[4–6] Type I chaperonins are found in eubacteria,

Abbreviations used: GFP, green fluorescent protein; wt, wild-type; Gdn·HCl, guanidine-hydrochloride; mDHFR, murine dihydrofolate reductase; GVI, gap volume index.

E-mail address of the corresponding author: plueckthun@bioc.unizh.ch

mitochondria and chloroplasts, while type II chaperonins are found in eukaryotes and archaea.[7–9]

All chaperonins share a similar overall structure, i.e. they are large oligomers composed of 14 to 18 subunits built up in the shape of a toroid, usually composed of two rings placed back-to-back.[10] Each oligomer subunit is built up from three domains:[11–13] the equatorial domain that holds the nucleotide-binding site and contributes most of the intra- and intersubunit interactions; the intermediate domain, which serves to relay conformational changes between the equatorial and the apical domain; and the apical domain itself, where both substrate binding and co-chaperonin binding is located in the case of type I chaperonins. In type II chaperonins, the co-chaperonin is included or built into the respective apical domain. The ring architecture of chaperonins allows for a large central cavity, where non-native proteins can undergo productive folding after being captured by the apical domains. The bacterial chaperonin from *Escherichia coli*, GroEL, requires a cofactor, called GroES, for many if not most of its substrates, even though the chaperonin-assisted folding of such a substrate might not always take place in the protected environment of the central cavity.[14] GroES is a heptameric dome-shaped ring composed of 10 kDa subunits, which can cap the GroEL cylinder on one or both ends. Productive protein folding further requires the hydrolysis of ATP.[15]

To summarize the chaperonin-assisted protein folding reaction cycle in a few words: first non-native substrate proteins are captured by hydrophobic patches residing on the apical domains of GroEL. Subsequent binding of ATP and co-chaperonin GroES provokes a large conformational change in GroEL. The hydrophobic polypeptide recognition regions in the apical domain of GroEL are turned upwards and outside by roughly 90°, and subsequently become obscured by co-chaperonin binding. As a consequence, the substrate protein is released and, with unknown success frequency, encapsulated in the inner cavity of the chaperonin *cis* complex or else released to the cytoplasm. The conformational change leads to almost a doubling of the volume of the central cavity and to a change of its character from hydrophobic to hydrophilic, thus providing a relatively polar environment to the released polypeptide, diminishing intermolecular interactions.[16–18] The ATPase cycle that is regulated both by bound substrate and cofactor GroES acts as a timer, providing the substrate protein with ~15 s to complete folding before GroEL is primed to release both GroES and substrate.[15,19]

The relevance of this cooperative action of both chaperonins for efficient cellular protein folding has been underscored by various studies.[20–23] These findings demonstrate that the entire folding machinery is necessary for productive chaperonin-assisted protein folding under non-permissive folding conditions for many of the GroEL-dependent proteins, and show that this machinery has evolved to serve the folding requirements of a broad range of substrates.[24–26] Very recently, it has been hypothesized that the GroEL machinery might have evolved mainly to assist the folding of TIM-barrel proteins or other proteins with complex α/β or $\alpha+\beta$ domains,[27] and it has been speculated that these proteins might share folding intermediates with distinct characteristics that lead to a chaperonin-dependence. By contrast, type II chaperonins seem to have evolved to serve the folding needs of a much smaller subset of proteins.[28] Notably, their essential role in protein folding cannot be complemented by type I chaperonins.[29] Although, for example, GroEL can bind and release the type II substrates actin and tubulin in an ATP-dependent manner, it cannot promote their folding,[29] despite the fact that substrate size is not the limiting determinant in this case.[30] *Vice versa*, type II chaperonins also fail to support folding of typical type I model substrates.[31,32] Thus, both type I and type II chaperonins must have evolved a mechanism to discriminate and tailor their substrate range.

The co-overexpression of GroEL *in vivo* with a substrate protein of interest,[33–35] as well as the provision of fragments encompassing the substrate-binding site of GroEL *in vitro*,[36,37] so-called mini-chaperones, have been investigated to exploit GroEL for better production of proteins. Unfortunately, the application of neither strategy to increase the functional yield of recalcitrant proteins in heterologous host organisms can be generalized. Indeed, for a great number of proteins that form inclusion bodies in *E. coli*, an increase of the cellular GroEL/ES level has no beneficial effect of preventing aggregation. Additionally, it is generally believed that GroEL/ES alone cannot reverse aggregation once it has occurred.[1,38,39]

It is interesting from a biotechnological point of view to examine the adaptability of GroEL to the folding needs of an otherwise aggregation-prone substrate protein. While this work was in progress, Wang and co-workers showed that by altering the ATPase cycle of GroEL, it was possible to generate a chaperonin with improved properties.[40] Interestingly, GroEL variants identified in this study that enhanced the folding of their model substrate protein, green fluorescent protein (GFP), also showed an increased polarity of the folding cavity, but no change in the actual substrate-binding site.

An alternative approach to tailor the GroEL molecule to the needs of a particular substrate of biotechnological interest, suggested by the heterogeneity of the substrate-binding sites found in eukaryotic chaperonins,[41] would be to directly alter its substrate-binding site residing in the apical domain. We considered it conceivable that the interaction energy should be "tuned" to a given substrate, to provide the optimized residence time compatible with the GroEL cycle, where different substrate proteins may show huge differences in hydrophobicity and charge, as well as in individual interactions leading to very different local dissociation rates. It would be irrelevant whether such an engineered GroEL is specific, as $GroEL_{wt}$

would always be present to take care of the requirements of other proteins of the cell.

To explore this latter possibility and to generate substrate-optimized chaperonins, we designed and constructed a combinatorial GroEL library, where the substrate-binding site was randomized, and we subjected this library to screening and selection experiments. While a great number of mutants were able to bind unfolded protein and show activity under permissive conditions, we found that only a very small number of library members were still functional under non-permissive folding conditions *in vitro* and *in vivo*. Inactive GroEL variants showed a decreased affinity for the co-chaperonin GroES. We discuss our findings on the background of the dual role of the randomized region, which is responsible for both substrate and cofactor binding, and with respect to the mechanistic contribution of the co-chaperonin GroES in the context of cellular chaperonin-mediated protein folding.

Results

To explore the mutational variability of the substrate-binding site of GroEL and to answer the question of whether it is possible to generate GroEL molecules adapted to the needs of particular substrate proteins, we designed and constructed a combinatorial library of GroEL, where the substrate-binding site (amino acid residues 191–290) was randomized only in those positions that are potentially involved in direct substrate contacts. Our design was guided by structural and biochemical data.[12,42–45] Furthermore, the side-chain accessibility, i.e. the surface exposure of potential randomization sites, was taken into account. On the basis of these considerations, nine residues in the apical domain of GroEL were chosen for randomization (Figure 1; Table 1). Note that by following this strategy, a total of 63 residues per binding site are randomized due to the 7-fold symmetry in GroEL. This library was subjected to *in vivo* screening and selection experiments, and selected library members were further characterized by different *in vitro* assays.

Library assembly and cloning

To generate a library with nine randomized positions across the apical domain of GroEL, we followed an overlapping oligonucleotide assembly strategy with six oligonucleotides. The codons of the randomized positions were encoded by NNK (allowing all 20 amino acids; positions 203, 230, 231, 238, 242, 267, and 268) and NBB (allowing 13 amino acid types; positions 264 and 270; N = A, T, G, C; K = G, T; B = C, G, T; the allowed amino acids and their expected frequencies by NNK and NBB are shown in Table 1 of the Supplementary Data). The NBB mixture increases the relative content of small, hydrophobic amino acids, compared to the NNK mixture, which allows all 20 amino acids. NBB was used at positions 264 and 270. Position 264 is part of the hydrophobic ValValAsn motif, implicated in GroES binding, while position 270 is buried deeply in the GroES-bound conformation. By using NBB, we therefore intended to disallow residues not compatible with these constraints. The fully assembled library cassette was subsequently PCR-amplified with 5′ and 3′ consensus primers containing recognition sequences for BsaI a type II restriction enzyme. Using these type II restriction enzymes, we could replace the wild-type (wt) sequence with the library cassette in the *groEL* gene on pZA21_BsaI (see Materials and Methods) without further changing its sequence, thereby generating pZA21_GroESLib, which contains our library under the control of an anhydrotetracycline-inducible $P_{LtetO-1}$ promoter.

With the nine randomized positions, the theoretical diversity amounts to $20^7 \times 13^2 = 2.1 \times 10^{11}$ at the amino acid level. To assess the quality of our library at the DNA level, 30 individual library members were sequenced before starting selection experiments. About 70% were correct at the DNA level (i.e. no frame-shifts). No obvious sequence bias for certain amino acid types at the randomized positions was observed.

Before starting selection and screening experiments with this library, we characterized unselected library members, i.e. randomly chosen constructs with correct DNA sequences, by expression and

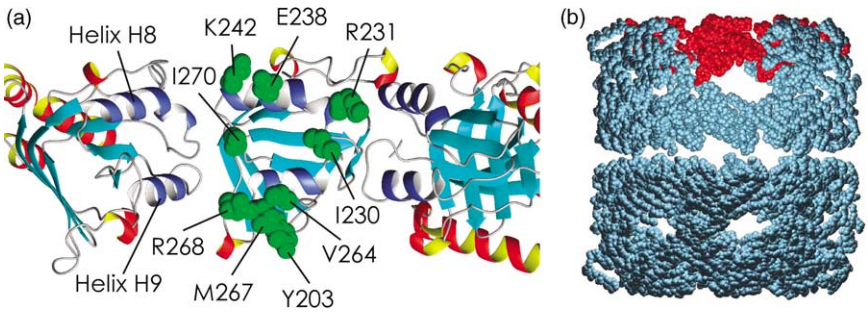

Figure 1. Sites chosen for randomization in the apical domain of GroEL. (a) Ribbon drawing of three adjacent apical domains in the uncomplexed form.[88] The location of these three domains in the whole GroEL structure is depicted in red in the space-filling model of the uncomplexed GroEL (shown in (b)). Note that two domains have been omitted to allow the view into the inner cavity. α-Helices H8 and H9 are depicted in blue; sites chosen for randomization are highlighted in green only for one subunit in a CPK space-filling model, displaying the side-chain of the wild-type amino acid.

Table 1. Properties of residues chosen for randomization

Residue position[a]	Residue location	Conservation[b] (%)	Side-chain solvent accessibility (%)[c]
Y203	Loop between S6 and S7	Y 99 (N 1)	65
I230	Loop between S8 and H8	I 54 (L 25, V 13)	29
R231	H8	Q 43 (R 32, K 22)	38
E238	H8	E 94 (Q 4, R 1)	70
K242	H8	Q 46 (K 37, R 8)	70
V264	H9	V 86 (L 14)	70
M267	H9	L 30 (M 30, I 19)	49
R268	H9	R 99 (K 1)	82
I270	Loop between H9 and S10	I 32 (T 29, G 28)	62

[a] The numbering follows Hemmingsen *et al.*[3] Residues in bold were identified to bind a peptide substrate,[43] underlined residues were identified to be involved in co-chaperonin binding by mutational analysis.[12]

[b] The frequency of occurrence in the sequence alignment of 100 prokaryotic Cpn60 proteins and, in parentheses, the second and third most frequent amino acid.

[c] Side-chain solvent accessibility relative to that in an extended G-X-G tri-peptide,[86] calculated using MOLMOL[87] with a 1.4 Å probe.

solubility tests. Expression screening revealed that all of the above library members that were correct at the DNA level and contained no stop codon could be expressed in soluble form in *E. coli* shake flask cultures (up to 150 mg/l of soluble protein). This is about 100 times the amount of $GroEL_{wt}$ in the cell, and thus the cell makes predominantly the recombinant protein. The corresponding proteins ran at the expected molecular mass in SDS-PAGE (data not shown).

For further usage, the library was introduced into *E. coli* XL-1 Blue F′ by electroporation, yielding an *in vivo* library with a final diversity of 1.2×10^8 individual library members. DNA plasmid preparations of this library were used in the subsequent *in vivo* screening and selection experiments.

In vivo complementation at 37 °C

Although the design and construction of the GroEL apical domain library was intended to identify substrate-optimized chaperonins that may become specialized for their substrate, rather than displaying a broad substrate specificity like $GroEL_{wt}$, we first tested the effects of our GroEL variants on the growth of *E. coli* AI90 (Δ*groEL*::*kan*^{R}) [pBAD-EL][46] at 37 °C. In this strain, the chromosomal *groEL* gene has been deleted, and GroEL is expressed exclusively from a plasmid-borne copy of the gene, which is under the control of the P_{BAD} promoter. As GroEL is essential for cell viability,[22] this strain can be used to test whether GroEL variants are able to complement $GroEL_{wt}$ function, because the expression of GroEL from this plasmid-borne copy of the gene is tightly regulated. P_{BAD} is activated by arabinose but suppressed by glucose, thus AI90 [pBAD-EL] cells cannot grow at 37 °C on medium supplemented with glucose. When we introduced our library into this strain on pZA21_GroESLib, where each cell harbors a $GroEL_{variant}$ under the control of an anhydrotetracycline-inducible $P_{LtetO-1}$ promotor, and cells were plated on LB agar plates containing anhydrotetracycline and glucose to suppress $GroEL_{wt}$ expression from pBAD-EL, only a very small subset of GroEL variants (0.4‰) was able to suppress this *groEL* growth defect (see also Table 2 for identification of the GroEL variants).

Table 2. Naming of GroEL variants

Name	Origin
u_number (e.g. $GroEL_{u25}$)	Unselected library member
s_number (e.g. $GroEL_{s9}$)	Library member obtained from $GroEL_{wt}$ complementation *in vivo* at 37 °C
k_number (e.g. $GroEL_{k4}$)	Library member obtained from *in vivo* screening for tailored GroEL variants

Analyzing these selected mutants of GroEL by expression tests, we found that they could be expressed as soluble proteins in liquid culture comparable to $GroEL_{wt}$ in the same plasmid (data not shown). Sequence analysis of these variants yielded no obvious consensus in the randomized positions, except for some similarity in character in five positions (Figure 2). At position Y203, an enrichment of tryptophan was observed; at position

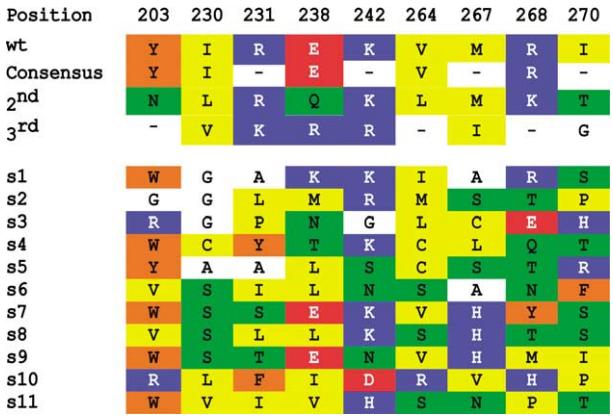

Figure 2. GroEL variants able to complement wt function. Sequence alignment of GroEL variants able to complement $GroEL_{wt}$ function at 37 °C. The names of the clones are given at the left-hand side of the respective sequence. The sequence for $GroEL_{wt}$ and the consensus sequence (residues with 50% conservation are named) calculated from the sequence alignment of 100 prokaryotic Hsp60 proteins are given above the selected sequences, as well as the second and third most frequent amino acid residue. Amino acids are colored according to their character: blue (positively charged), Arg, Lys, His; green (polar), Ser, Thr, Asn, Gln; yellow (aliphatic), Val, Leu, Ile, Pro, Cys, Met; orange (aromatic), Phe, Tyr, Trp; red (negatively charged), Asp, Glu; white (small), Ala, Gly. Note that in this representation only the randomized positions are shown (numbering follows Hemmingsen *et al.*[3]).

I230 mainly small residues were selected (Gly, Ala, Ser) and at position K242 either lysine (as in $GroEL_{wt}$) or another charged or polar residue was selected. At position V264, predominantly hydrophobic amino acids were selected, whereas at position R268 either charged or hydrophilic/neutral amino acids were selected. All other positions, however, showed a broad distribution of amino acids with different side-chain characteristics. It is noteworthy that at position R231, where the prokaryotic chaperonins mainly have a charged residue, the selected GroEL variants predominantly display a hydrophobic amino acid residue.

In a second series of *in vivo* screening experiments, we then tried to find chaperonin variants that showed improved performance for a particular substrate, independent of any need to support the growth of *E. coli*, as the $GroEL_{wt}$ would still be present to support the folding of its natural substrates.

In vivo screening for tailored GroEL variants

To identify GroEL variants that can serve to accomplish the folding needs of a distinct aggregation-prone substrate, we performed *in vivo* screening experiments with GFP from the jellyfish *Aequorea victoria*. GFP is a very attractive candidate for this kind of experiment. (i) Folding of GFP_{wt} is inefficient during recombinant expression in *E. coli* and only moderately enhanced by simultaneous co-overexpression of the chaperonins.[40] GFP is thus a GroEL substrate, but clearly the chaperone action leaves room for improvement. (ii) GFP depends on chaperonin-assisted folding to reach its native state under non-permissive folding conditions *in vitro*[23] and thus benefits directly from improvements of GroEL. (iii) Efficient folding of GFP can be monitored readily in living cells by formation of green fluorescence under UV light. (iv) It is an *in vivo* and *in vitro* substrate of GroEL, even though its three-dimensional fold is not directly related to the architecture of known natural *E. Coli* substrates of GroEL.[27] Importantly, many recombinant proteins causing considerable difficulties in functional expression are those with a high content of β strand, just like GFP. It is thus of particular interest whether GroEL can adapt to such substrates. For these experiments, wt GFP was used, and not one of the mutants that have been adapted to improved bacterial folding by changes in the sequence.

As the screening experiments for tailored GroEL variants were performed in a cell-based assay, we verified before the screening experiments that the chosen level of overexpression of GFP, and GroEL overexpression at any level, had no influence on cell growth (data not shown). The library was introduced into *E. coli* DH5αZ1 cells[47] containing an inducible GFP_{wt} gene on pAT115-wtGFP, and cells were plated on LB agar plates containing glucose and the respective inducers. Note that in this experiment the screening was performed in the presence of an endogenous $GroEL_{wt}$ background, and plasmid-encoded overexpression of both GFP and $GroEL_{variant}$. Thus, cells always contained a certain level of $GroEL_{wt}$. Because of the higher level of expression of the plasmid-encoded variants, the formation of chimeric GroEL molecules consisting of few wt subunits and mostly mutant subunits is likely.

Screening was performed by visual inspection of UV-illuminated plates for bright green colonies. In total, we screened 6×10^6 variants. Unexpectedly, the formation of bright green colonies was a very rare event, and we found only one mutant with a slightly enhanced activity regarding GFP refolding compared to $GroEL_{wt}$ (Figure 3; $GroEL_{variant}$ k8). By contrast, on these plates a large fraction of colonies exhibited a fluorescence that was even lower than the fluorescence observable in colonies expressing only GFP_{wt} in the presence of an endogenous $GroEL_{wt}$ background. These dominant negative variants resulting in a "darker" phenotype did not alter the overall amount of GFP expression. However, cells co-expressing GFP_{wt} and one of these variants showed reduced levels of GFP in the soluble fraction, compared to cells co-expressing GFP and $GroEL_{wt}$ (Figure 4). A typical member of this class of GroEL variants is the clone $GroEL_{k4}$ (Figure 4). It appears as if this and

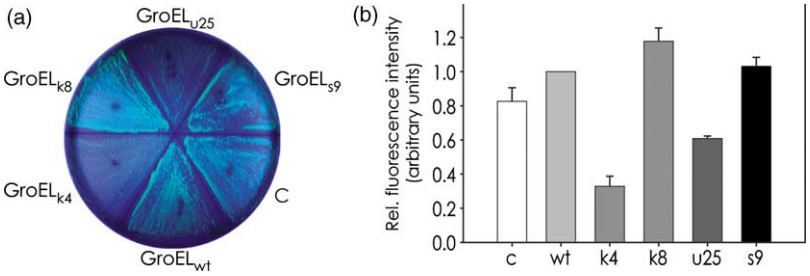

Figure 3. Comparison of chaperonin-assisted folding of GFP_{wt} by different GroEL variants *in vivo*. (a) Cells overexpressing GFP_{wt} and one of the GroEL variants were grown at 37 °C on LB agar plates containing 100 ng/ml of anhydrotetracycline, and 50 μM IPTG for induction of the respective genes and visualized under UV illumination (refer to Materials and Methods for details). (b) Cells co-expressing GFP_{wt} and one of the GroEL variants were normalized by cell density, and the fluorescence of GFP in the soluble fraction was measured at $\lambda_{em}=506$ nm ($\lambda_{ex}=395$ nm). The fluorescence signal of the sample containing $GroEL_{wt}$ was set to 1. Column c, Fluorescence signal due to the endogenous GroEL background; k4/k8, GroEL variants originating from *in vivo* screening for tailored GroEL variants; s9, GroEL variant originating form complementation at 37 °C; u25, unselected library member.

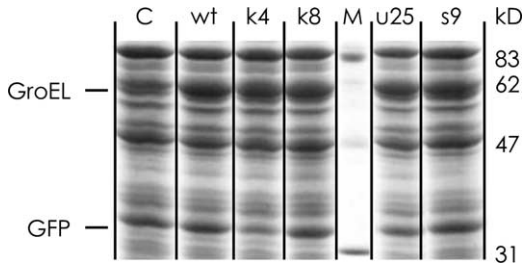

Figure 4. Influence of GroEL variants on the production of soluble GFP. Cells co-expressing GFP_{wt} and one of the GroEL variants, always together with GroES wt, were collected 4 h after induction, normalized by cell density and disrupted by sonication. SDS-PAGE analysis of the soluble fraction was used to determine the levels of soluble GFP in the respective samples. As controls, cells expressing GFP in an endogenous GroEL background (without plasmid-encoded GroEL/GroES, lane C), co-expressing $GroEL_{wt}$ (wt) and an unselected library member (u25) were included.

similar GroEL variants are less active *in vivo*, at least with GFP as substrate. Nevertheless, the "dominant negative" phenotype does not prevent cell growth, and must therefore be benign for the natural substrate proteins of *E. coli*.

Biophysical characterization of GroEL variants

To address the reason for the different behavior of selected GroEL variants observed in the *in vivo* experiments, we performed *in vitro* chaperonin-assisted folding assays with a distinct subset of GroEL mutants under different folding conditions and with different substrates. For this more detailed biophysical characterization, we chose to examine three different GroEL variants exemplifying the outcome of the two *in vivo* experiments. One variant was chosen from the complementation experiments at 37 °C, which was able to suppress the *groEL* growth defect in *E. coli* AI90 ($GroEL_{s9}$). Two GroEL variants, $GroEL_{k8}$ and $GroEL_{k4}$, were chosen from the *in vivo* screening experiments for tailored GroEL variants; $GroEL_{k8}$, because this variant showed a slightly beneficial effect on chaperonin-assisted GFP folding *in vivo*, and $GroEL_{k4}$, because this variant showed a lower efficiency of GFP folding. As internal controls, $GroEL_{wt}$ and an unselected library member, $GroEL_{u25}$, were included in the experiments.

Chaperonin-assisted folding of GFP *in vitro*

We first wanted to compare the results of chaperonin-assisted folding of GFP *in vivo* with equivalent experiments *in vitro*. Guanidine hydrochloride (Gdn·HCl)-denatured GFP is a "non-stringent" substrate protein in chaperonin-assisted folding. This means that denatured GFP can refold spontaneously in dilute solutions, and in the presence of GroEL its folding is independent of the presence of the co-chaperonin GroES. In aqueous buffers containing GroEL, the presence of ATP alone is sufficient to release the GroEL-bound GFP into the bulk solution, where it completes folding on its own.[48] With GroES present, however, GFP is retained to complete folding in the inner cavity of the chaperonins.[17] By contrast, under more stringent folding conditions, Gdn·HCl-denatured GFP can be efficiently refolded only by the entire chaperonin system (i.e. GroEL, GroES and ATP),[23] whereas spontaneous refolding is virtually fully suppressed. We therefore tested our GroEL variants under both folding conditions.

Folding of GFP under permissive conditions resembles closely the results found in the *in vivo* experiments: the "dark" mutant $GroEL_{k4}$ performs significantly worse than $GroEL_{wt}$, while the "brighter" variant $GroEL_{k8}$ is about the same as wt within experimental error. The unselected mutant $GroEL_{u25}$ is inactive, and the variant obtained by *in vivo* complementation, $GroEL_{s9}$, is slightly less active than $GroEL_{wt}$ (Figure 5(a)).

By contrast, folding under non-permissive conditions, which were induced by adding crowding agents to the buffer,[49] and which should actually mimic more closely the *in vivo* situation, is different

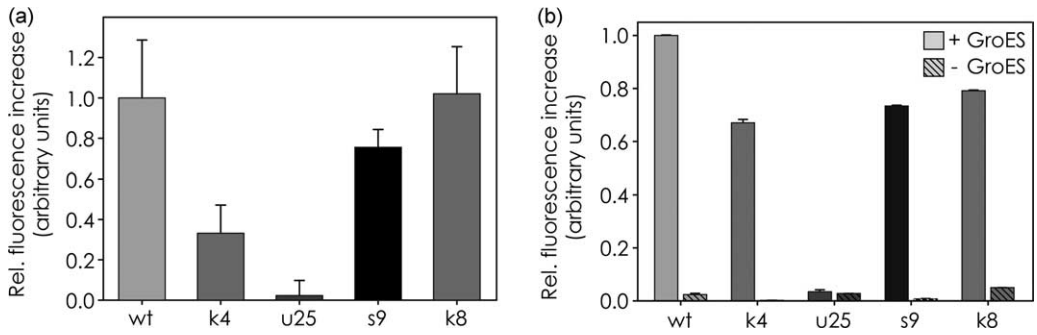

Figure 5. Chaperonin-assisted refolding of GFP_{wt} *in vitro*. The chaperonin-assisted refolding of Gdn·HCl-denatured GFP was followed for different GroEL variants under (a) permissive folding conditions and (b) under non-permissive folding conditions. Under conditions of macromolecular crowding (by including 28% Ficoll in the refolding buffer) the refolding of GFP_{wt} becomes fully dependent on the presence of GroES. Refolding of GroEL-bound GFP was initiated by the addition of ATP alone or by adding ATP and GroES to the refolding buffer. GFP fluorescence was measured before and 15 min after the addition of ATP and/or ATP/GroES (λ_{ex} = 398 nm, λ_{em} = 506 nm). The increase of intrinsic GFP fluorescence due to assisted folding by $GroEL_{wt}$ was set to 1.

with respect to the "deleterious" variant $GroEL_{k4}$ (Figure 5). In the latter case, successful GFP refolding for all variants is dependent on GroES and ATP, as reported previously,[23] and now $GroEL_{k4}$ is capable of assisting significantly in Gdn·HCl-denatured GFP refolding under these non-permissive folding conditions *in vitro* (Figure 5(b)). For the other GroEL variants, no striking difference compared to $GroEL_{wt}$ was observed in these *in vitro* experiments. As expected, the unselected variant $GroEL_{u25}$ was unable to assist GFP folding in both *in vitro* assays (Figure 5), and it failed to promote folding in the *in vivo* situation (Figure 3). Under non-permissive conditions, GroES binding is required, and for the mutants that show GroES binding, at least to some degree (all except $GroEL_{u25}$), the differences between the mutants are small (see Discussion).

Prevention of bovine mitochondiral rhodanese aggregation

The different performance of each of the GroEL variants may result, in part, from their different abilities to recognize GroES, and from differences in the interactions with denatured substrate proteins. To test this latter possibility, we examined the prevention of bovine mitochondrial rhodanese aggregation by the different GroEL variants.

A common feature of the chaperonins is their ability to recognize hydrophobic surfaces of unfolded proteins to which they can bind, and thus, stabilize a broad range of different conformations of unfolded polypeptides,[50] thereby preventing undesired side-reactions. Whereas the nature of chaperonin-assisted folding reactions depends critically on both the folding conditions and the properties of the substrate,[21,24,51] the stabilization of unfolded structures, and thus the inhibition of aggregation in a stoichiometric reaction, is generally achieved by GroEL alone. Thus, aggregation assays probe for a very basic chaperonin activity, requiring only binding of substrate.

Interestingly, all GroEL variants were capable of inhibiting the aggregation of denatured bovine mitochondrial rhodanese upon its dilution into buffer containing one of the GroEL variants each (Figure 6). Thus, even the unselected $GroEL_{u25}$ variant was competent to bind to and to stabilize non-native conformational states of the GroEL substrate proteins tested.

Chaperonin-mediated protein folding of two model substrates

As we did not observe any differences between the GroEL variants in the bovine mitochondrial rhodanese aggregation assay, although they originate from different selection and screening experiments, we were curious to test their behavior in chaperonin-assisted folding of two further GroEL model substrates, namely bovine mitochondrial rhodanese and murine dihydrofolate reductase (mDHFR), to put the results obtained with GFP into perspective.

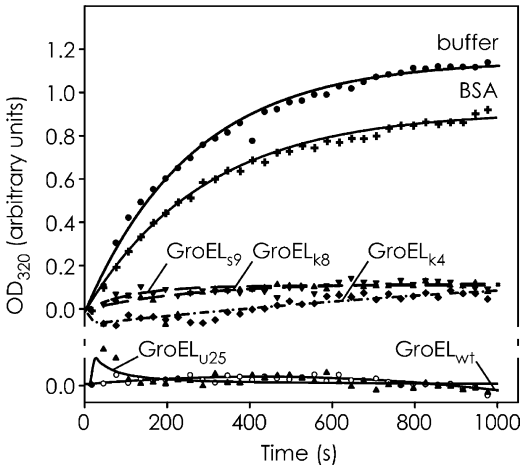

Figure 6. Inhibition of rhodanese aggregation by different GroEL variants. Rhodanese aggregation upon dilution from 25 μM (in Gdn·HCl) to 0.25 μM was monitored by light-scattering at λ = 320 nm and 25 °C for rhodanese in buffer alone, buffer containing 15 μM BSA, or buffer containing 0.25 μM GroEL variants (s9, k8, k4, u25, and wt). For clarity, the curves for $GroEL_{wt}$ and the $GroEL_{u25}$ variant have been moved down.

As the nature of chaperonin-mediated protein folding depends on both the folding conditions and the substrate protein (see above), we chose these two GroEL model proteins for the following reason. Like GFP, mDHFR is a "non-stringent" GroEL substrate. After Gdn·HCl-denatured mDHFR is bound to GroEL, the addition of ATP alone leads to efficient reactivation. Again, GroES is not required for substrate release.[52] Thus, we could employ the refolding of mDHFR to assess if our selected GroEL variants were capable of binding conformational states of denatured proteins tightly enough to prevent aggregation. Furthermore, we can probe with this assay if the proteins were bound in a conformation competent for refolding upon release from GroEL, which was triggered by the addition of ATP, to these preformed binary chaperonin–substrate complexes.

By contrast, the model protein bovine mitochondrial rhodanese has an absolute requirement for the entire chaperonin system, even under permissive folding conditions.[51] Thus, by using this substrate we can assess the different characteristics of our GroEL variants in the context of the concerted action of GroEL, GroES and ATP.

As anticipated from our findings on the prevention of rhodanese aggregation, all GroEL variants were competent to bind denatured mDHFR and, upon addition of ATP, released the substrate into the bulk solution (Figure 7(a)). No significant differences in GroEL-mediated mDHFR refolding could be observed, regardless from which selection

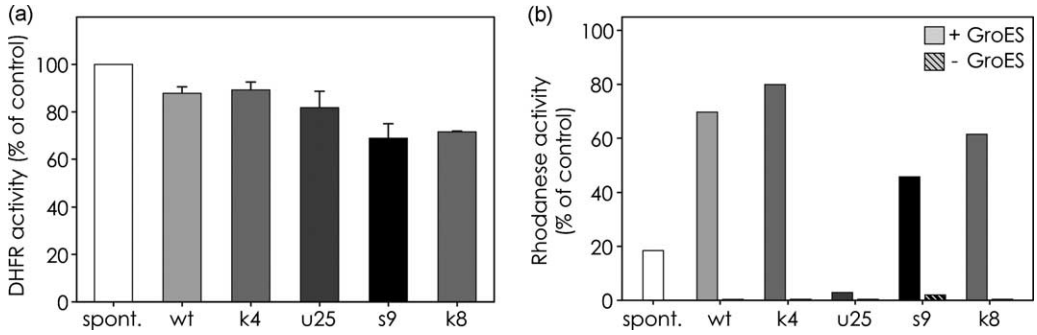

Figure 7. Effect of GroEL variants on mDHFR and rhodanese folding. (a) Gdn·HCl-denatured mDHFR was diluted rapidly into buffer containing the indicated GroEL variants. Release from the GroEL variants was initiated by the addition of ATP and refolding of mDHFR was monitored spectrophotometrically at $\lambda = 340$ nm by following the turnover of NADPH due to refolded mDHFR. The mDHFR activity regained after 5 min by spontaneous folding in buffer containing no chaperonins was set to 100%. (b) Gdn·HCl-denatured rhodanese was allowed to refold in buffer containing ATP and GroES (filled bars) or no GroES (hatched bars) for 30 min. Rhodanese activity regained was measured by adding $Na_2S_2O_3$ and KCN to the buffer for 5 min at 30 °C. The turnover of CN^- to SCN^- was stopped by the addition of formaldehyde and the amount of SCN^- was measured spectrophotometrically by following the formation of an iron-thiocyanide complex at $\lambda = 460$ nm. The activity of an equivalent amount of native rhodanese as used in the refolding reactions was set to 100%.

the variant originated. Especially variants $GroEL_{k4}$ and $GroEL_{k8}$, which were not selected for mDHFR refolding, performed like $GroEL_{wt}$. Thus, it seems that these two GroEL variants are still promiscuous, even though they originated from a selection round for increased performance on a different substrate.

Again, the unselected $GroEL_{u25}$ variant is fully active in this assay. This suggests that a basic functionality of chaperonins, namely the binding to and the protection of aggregation-prone regions displayed by denatured proteins, is still retained in all GroEL library members analyzed.

By contrast, if the entire chaperonin system is needed for efficient protein folding, as exemplified by chaperonin-assisted refolding of bovine mitochondrial rhodanese, only the selected GroEL variants were capable of serving in this reaction, whereas the unselected variant $GroEL_{u25}$ was not (Figure 7(b)). Again, the behavior of the variant $GroEL_{k4}$ was surprising. Although deleterious in the GFP folding experiments *in vivo* (Figure 3), it was active in chaperonin-mediated folding of both bovine mitochondrial rhodanese and GFP under non-permissive folding conditions *in vitro* (Figures 5(b) and 7(b)). Note that chaperonin-mediated folding of bovine mitochondrial rhodanese was indeed GroES-dependent for all variants analyzed, as reported previously,[51] as no activity was found in the absence of GroES (Figure 7(b)).

Interaction of the GroEL variants with the co-chaperonin GroES

To determine the reason for the different behavior of the variant $GroEL_{k4}$ in the different assays, we investigated the interaction of our GroEL variants with the co-chaperonin GroES. The interaction of GroEL with its cofactor GroES can be assayed by proteinase protection. The GroEL double-ring toroid is a compact structure that, in the presence of nucleotides, is resistant to proteolytic cleavage by proteinase K. Only the last 16 C-terminal residues of each subunit are prone to proteolytic processing by this proteinase.[53] Two properties of the GroEL oligomer may permit proteinase K to cleave these C-terminal residues: (i) the C-terminal tail of GroEL is very flexible and may reach out of the central cavity; (ii) proteinase K is small enough to enter the cavity and exert its proteolytic activity there.

Thus, this cleavage reaction is suppressed when GroES binds to the top of the GroEL cylinder, thereby shielding the interior of the central cavity. As a consequence, proteinase K cannot gain access to the tails of the seven GroEL subunits in this ring. The opposite ring, however, which is not sealed by GroES, is still susceptible to proteolysis. Thus, proteolytic cleavage of GroEL by proteinase K in the presence of GroES results in two bands of equal intensity after analysis by SDS-PAGE. Therefore, this assay is appropriate to obtain a qualitative measure of the binding parameters of GroEL/GroES interaction, and to test for functionality of the GroEL and GroES oligomers.

As can be seen in Figure 8, binding of the co-chaperonin GroES was impaired for the unselected variant $GroEL_{u25}$ and the "deleterious" variant $GroEL_{k4}$ under the assay conditions. Also, the co-chaperonin binding of the two "active" variants $GroEL_{k8}$ and $GroEL_{s9}$ seemed to be weakened slightly, compared to $GroEL_{wt}$ (Figure 8), but this observation might well be attributed to the experimental inaccuracy of this assay. The inability to bind GroES was more pronounced in the unselected variant $GroEL_{u25}$ than in the other variants. This explains its failure to assist in chaperonin-mediated folding of substrate proteins under all non-permissive folding conditions tested, where the entire chaperonin system is needed for efficient folding.

To underscore our findings that the main reason for the different behavior of the GroEL variants was due to a change in their ability to

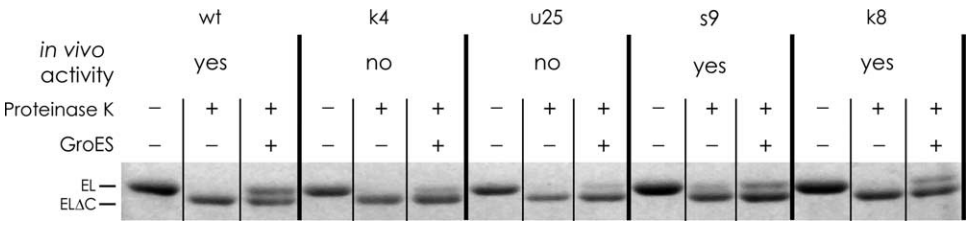

Figure 8. Proteinase protection of GroEL carboxy termini by GroES. The proteinase K protection of GroEL carboxy termini for different GroEL variants by GroES is shown. The band in the first lane of each sample corresponds to the full-length GroEL subunit (EL). The band in the second lane of each sample corresponds to truncated GroEL subunits (ELΔC), generated by treatment of the chaperonin with proteinase K. Binding of GroES on top of GroEL protects the adjacent chaperonin ring from attack by proteinase K, depending on the strength of GroEL–GroES interaction (lane 3).

bind the chaperonin cofactor, we took a closer look at the GroEL–GroES protein–protein interface. First, we investigated the H-bonding pattern of our variants. The native, wild-type H-bonding network is increasingly disrupted in unselected library members compared to $GroEL_{wt}$ (Figure 1 of the Supplementary Data). This may indicate that in unselected or inactive GroEL variants the interaction with GroES is weakened, compared to active variants. These findings underscore the results gained in the proteinase protection experiments.

Second, we compared the gap volume index of GroEL variants competent to suppress the *groEL* growth defect in *E. coli* AI90 with those who were unable to do so (Figure 9). The gap volume index (GVI) is a measure of the surface complementarity of a given protein–protein interface,[54] here, the GroEL–GroES interface. We found that GroEL variants competent to complement $GroEL_{wt}$ in *E. coli* AI90 had a GVI similar to that of $GroEL_{wt}$. Variants unable to complement the *groEL* growth defect, however, had an increased GVI (Figure 9). Thus, the GroEL–GroES interface is perturbed in inactive variants, while it is retained or only slightly altered in active variants. Note that, although we analyzed only the GroEL–GroES interface of a single subunit, the effect becomes multiplied due to the 7-fold symmetry of the GroEL–GroES complex.

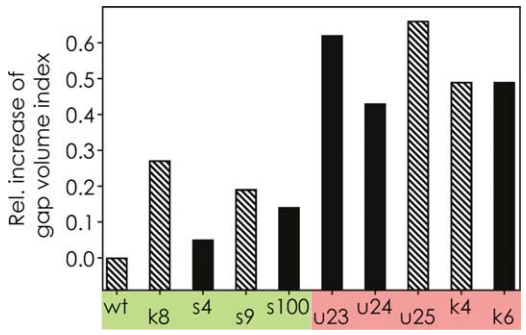

Figure 9. Gap volume index of selected GroEL variants. The relative increase of the gap volume index for different GroEL variants is given to the gap volume index calculated for the native GroEL–GroES interface of one GroEL–GroES subunit. Green bar, variants able to complement $GroEL_{wt}$ in *E. coli* AI90; red bar, variants not able to complement $GroEL_{wt}$ in *E. coli* AI90. Hatched bars, GroEL variants analyzed in more detail in the present study. The gap volume index was calculated using the program SURFNET;[89] the mutant PDB files used as input files to this program were calculated using the Swiss-PDB viewer v3.7.[90]

Discussion

Chaperonin function

Although the information required for a protein to attain its native, functional conformation is stored in its amino acid sequence,[55] cellular protein folding is a much more problematic task, compared to folding under optimized *in vitro* conditions, since the crowded environment of the cell favors a multitude of unwanted side-reactions.[2] Nature has tried to respond to this cellular protein folding problem by creating a large number of proteins, the molecular chaperones, that assist the folding of other proteins. The chaperonins, GroEL/ES in particular, have evolved to recognize and promote the folding of a broad range of proteins. Chaperonins assist protein folding in a dual manner.

First, chaperonins temporarily bind to and protect unfolded and aggregation-prone regions of misfolded proteins. Thereby, chaperonins prevent undesired side-reactions that would lead to further misfolding and aggregation. Upon binding, the protein is thought to be partially unfolded,[56,57] and it can thus escape a local energy minimum or folding trap. It is especially kept isolated from intermolecular interactions with other molecules of the same type, which would lead to domain-swapped aggregates. Finally, upon release from GroEL, if it is not yet native, the polypeptide can rebind and is thus given another opportunity to reach its native state (iterative annealing model).[58–60]

Second, the chaperonins provide a protected environment in which the protein can fold by itself (Anfinsen cage). Thus, the folding chain is isolated from the highly crowded cytoplasm, but at the same time confined within the chaperonin folding cage.[61]

Whereas the former chaperonin function depends solely on providing a suitable binding surface to bind the folding intermediate, thereby preventing its aggregation, the latter folding

assistance depends on the GroEL–GroES complex formation and the encapsulation of the substrate within its cavity. The successful folding of model proteins *in vitro*, just by providing a suitable binding surface, has been reported.[36,37] However, this chaperonin functionality is insufficient to complement the entire chaperonin system *in vivo*.[62] The fully assembled GroEL 14-mer has a timed release and a conformational change built in, allowing "annealing" of the substrate and multiple binding cycles, independent of GroES. Only the full system with GroES allows substrate confinement, however.

Despite this sophisticated machinery, a multitude of cloned foreign proteins do not reach the native state in bacteria, a huge challenge for biotechnology and protein research in general. For this reason, we investigated in this study whether it would be feasible to optimize the binding region of the intact GroEL molecule for a given target protein in the cellular context. For example, the binding site might interact too strongly with some proteins, not releasing them, or too weakly by e.g. unfavorable electrostatic effects, thus releasing them too fast. Hence, optimizing this chaperonin–substrate interaction for a given protein might conceivably result in an increased folding yield in chaperonin-assisted refolding of this protein and proteins with similar binding characteristics.

Our endeavor was further motivated by the fact that the eukaryotic counterparts of bacterial chaperonins have obviously evolved a mechanism to adapt to their substrates. This substrate discrimination is partially reflected by their molecular architecture, which is built up from different subunits displaying different binding regions.[10] Hence, although the number of proteins assisted by eukaryotic chaperonins is greater than previously thought, it is still evident that eukaryotic chaperonins are not as promiscuous as GroEL, but seem to have evolved for a smaller subset of essential proteins,[28,41] with different chaperonin subunits for different substrates. It seemed thus an interesting question whether a prokaryotic chaperone can also be adapted to a given protein, e.g. one of biotechnological importance.

There is a wide range of GroEL variants able to support binding to unfolded polypeptides. A smaller subset is able to allow release, and indeed, this is substrate-dependent (cf. Figures 5(a) and 7(a)). We were unable to select any greatly improved chaperonins only by engineering the substrate-binding site of GroEL. The mutational freedom in the apical domain of GroEL appears to be very narrow and seems to be related to GroES binding (Figure 8), as GroES is needed for some substrates under non-permissive conditions only (Figure 5(b)), while for others it is needed under all conditions (Figure 7(b)).

Some variants (e.g. $GroEL_{k4}$) showed deleterious effects on the folding of GFP_{wt} in the *in vivo* screening experiments, which could be visualized by reduced levels of GFP in the soluble fraction of cells co-expressing GFP and $GroEL_{k4}$, compared to cells not overexpressing any GroEL at all (Figure 4, cf. lanes C and k4). As these *in vivo* screening experiments were performed in the presence of an endogenous $GroEL_{wt}$ background, the deleterious effect of the variant $GroEL_{k4}$ may most likely result from a dominant negative effect of these mutants and/or reducing the active proportion of endogenous $GroEL_{wt}$ molecules by the formation of mixed chimeric GroEL molecules. Presumably, these mutant or mixed mutant GroEL molecules will also show a decreased affinity for the co-chaperonin GroES, comparable to the purified $GroEL_{k4}$ variant, which will ultimately result in an overall decreased efficiency for chaperonin-mediated GFP folding in this case. These mutants are thus not simply inert; they appear to interfere with the normal operation of the wt system. Nevertheless, the effects are benign, as the cells grow normally, and GroEL/ES is needed for normal growth.

In the *in vitro* folding assays, however, this decreased affinity for the co-chaperonin is compensated by the fact that the ratio of GroES to GroEL in the various folding assays was at least five to ten times higher than in the natural *in vivo* situation and above the K_D, thereby shifting the equilibrium to the complexed state,[63] thus compensating for a loss in binding affinity for GroES. This may explain why the variant $GroEL_{k4}$ can assist in GroES-dependent chaperonin-assisted folding reactions *in vitro* but not *in vivo* (Figures 5(a) and 7(a)).

The unselected mutant $GroEL_{u25}$ has decreased activity in GFP folding *in vivo* and performs poorly in GFP folding under permissive and non-permissive conditions (with and without GroES), poorly in rhodanese refolding (GroES dependent), but well in DHFR refolding (GroES-independent) and in preventing rhodanese aggregation by binding alone (GroES-independent). This can be explained by having a very poor interaction with GroES, even poorer than $GroEL_{k4}$, and an inappropriate interaction with GFP (too tight or too weak). On the other hand, the interaction with rhodanese must be tight enough, as this substrate binds (Figure 6).

In contrast to promoting folding by substrate binding, the second mechanism of chaperonin-mediated protein folding relies on the formation of the GroEL–GroES complex.[17] Encapsulation and confinement of substrate protein within the hydrophilic space of the chaperonin cavity is thought to "smooth" the energy landscape of folding for the encapsulated protein, removing folding traps and allowing the protein to reach the native state.[61,64] This requires a relative adaptation of the chaperonin interaction with the substrate protein. As the modification of the binding site is constrained by the necessity of interaction with GroES, the second possibility is to change the "timer". Therefore, by adaptation of the ATPase cycle, which acts as a timer of encapsulation to the folding needs of a given protein, it should be possible to obtain substrate-optimized chaperonins. Different times of encapsulation would allow different proteins to

reach its native state during a single round of encapsulation and release. Although this has been proposed theoretically for some time,[65] and shown experimentally some time ago,[66] only recently Wang *et al.* successfully implemented this strategy with the purpose to obtain substrate-optimized chaperonins.[40] It should be noted that in this case substrate adaptation comes as a side-effect of the slower ATPase cycle, rather than by specific substrate recognition. Wang *et al.* found that rhodanese was equally well folded by the substrate-optimized chaperonins as the original substrate GFP_{wt}.[40] This can be explained by the fact that rhodanese needs several cycles in chaperonin-mediated folding to reach its native state.[49] Thus, several GroEL–GroES reaction cycles were substituted by a slower one, due to a reduced ATPase activity. The broad applicability of this approach for biotechnological purposes has yet to be proven, as the timing of encapsulation and release is of critical importance for the efficiency in chaperonin-mediated protein folding.[67] It is noteworthy that Wang *et al.* did not observe any mutations in the substrate-binding region of GroEL in their substrate-optimized chaperonins, although this region was included in their shuffling experiments.[40] This finding is consistent with our finding of a narrow mutational window of the apical domain of GroEL.

Co-chaperonin *versus* substrate binding

The importance of the cofactor in chaperonin-assisted protein folding has long been known and GroEL has evolved together with its co-chaperonin GroES to meet the folding requirements of its substrates in the cellular context.[20–22] In addition, it is likely that also under non-permissive folding conditions *in vitro* the entire chaperonin system is advantageous for efficient folding.[20,21,23,49]

GroEL recognizes and binds its substrates mainly *via* exposed hydrophobic patches,[68] a general feature of partially denatured and unfolded proteins. As hydrophobic interactions are generally unspecific, GroEL is able to interact with its broad substrate spectrum. Furthermore, this interaction with its substrates is very tight, in the low nanomolar to high picomolar range.[69] Thus, further improvement of the binding affinity might hamper the release of substrate into the inner cavity upon cofactor binding and could clog the chaperonin. In contrast, it is conceivable that some substrates are binding too tightly, such that they are not released, thus clogging the normal cycle. In this case, a lower binding affinity might be advantageous by engineering the GroEL binding site. Hence, in a dual system with both $GroEL_{wt}$ and $GroEL_{variant}$ present, $GroEL_{wt}$ still takes care of the normal *E. coli* proteins, and the variant might then bind a given recombinant protein with a lifetime commensurate with the cycle time; albeit in the presence of binding of that protein also to $GroEL_{wt}$ and the presence of mixed oligomers with unknown effects on *E. coli* proteins.

Indeed, tight binding is not seen in the eukaryotic counterparts. Group II chaperonins typically bind their substrates in the micromolar range.[8] Thus, in these cases there might be room for substrate discrimination and adaptation due to the overall lower substrate affinity.

The substrate-binding site of GroEL overlaps with the site of cofactor binding,[12,43,70] and even residues that do not contact GroES directly alter its binding affinity to GroEL.[71] This indicates also that the binding of GroES, which is essential for chaperonin function, is highly regulated *via* an allosteric network that is still not fully understood.[72] These reasons render mutational alterations in this region of GroEL a highly difficult task. Co-chaperonin binding is achieved by a complex network of interactions that is sensitive to alterations. Sequence analysis of active and deleterious $GroEL_{variants}$ shows that active $GroEL_{variants}$ do not differ in their overall amino acid side-chain characteristics compared to $GroEL_{wt}$ (Figure 10), i.e. they never comprise a hydrophobic amino acid where $GroEL_{wt}$ displays a charged or polar one and *vice versa*. The only exception in that respect is the position of Y203, which is structurally located in a loop region beneath the two substrate-binding helices H8 and H9 (Figure 1(a)). Here, tryptophan was predominantly found in all active $GroEL_{variants}$ (Figure 10). An explanation for this observation might be that the residue found in $GroEL_{wt}$, Y203, can contribute to protein–protein interactions with the substrate both by hydrophobic contacts and by contributing to an H-bonding network. This residue becomes buried in the GroEL/GroEL interdomain interface upon the conformational changes induced by ATP and co-chaperonin binding. Whether it is the interaction with the neighboring domain in the ATP/GroES-bound state or the interaction with

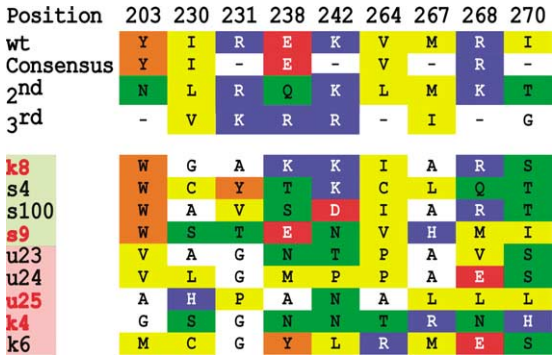

Figure 10. GroEL variants. Sequence alignment of GroEL variants analyzed in more detail in this study (green bar on the left, active variants; red bar, inactive variants). The names of the clones are given at the left-hand side. The sequence for $GroEL_{wt}$ and the 50% consensus sequence (residues occurring in more than 50% of the sequences are given as letters) calculated from the sequence alignment of 100 prokaryotic Hsp60 proteins are given above the selected sequences as well as the second and third most frequent amino acid residue. Amino acids are colored according to their character as in Figure 2. Numbering follows Hemmingsen *et al.*[3]

the substrate that drives the Trp selection is currently unknown. Smaller hydrophobic residues as found for the inactive variants appear to be inactive (Figure 10). Generally speaking we found that active $GroEL_{variants}$ contain amino acid residues with side-chain characteristics similar to those of $GroEL_{wt}$. This explains also why active GroEL variants found in our selection experiments are still promiscuous and are capable of fully complementing wild-type function: GroEL has to assure cofactor binding in the first place. This binding task seems to be more inflexible than the accommodation of various conformations and surface features of non-native proteins, and these GroES binding constraints do not allow the alteration of this region of GroEL except within a rather narrow window.

These findings are corroborated by reports on specialized co-chaperonins from bacteriophages.[73] Bacteriophage T4 major capsid protein Gp23 depends on chaperonin-assisted folding to reach its native conformation, which is vital for phage propagation. However, it is too large to be encapsulated by the endogenous bacterial chaperonin system. Thus, T4 encodes its own co-chaperonin, Gp31, which can substitute GroES. This substitution leads to an enlargement of the central cavity that can now accommodate the capsid protein Gp23.[73] In this case, the critical GroEL-co-chaperonin interface is not affected. It should be noted that, although Gp31 shows only 43% sequence similarity to GroES within the GroEL binding region (the mobile loop), the H-bonding pattern corresponds to that of GroEL/S. In addition, the GroEL–Gp31 interface is very complementary as judged by the gap volume index (0.87; GroEL/ES=1).

Conclusions

Our findings suggest that it may be very difficult to improve chaperonin-mediated protein folding of a given recalcitrant recombinant proteins by engineering the substrate-binding site of GroEL. The chaperonin system is allosterically regulated, and mutations altering this allostery can have substrate-specific effects on its folding function.[74] Interaction with the co-chaperonin GroES is an important element of this allosteric regulation.[75] Thus, alteration of the GroEL–GroES interaction directly affects efficient protein folding under non-permissive conditions, and this interaction is, most likely, required for the substrates that are most aggregation-prone. As pairs of homologous isozymes were observed to bind with different affinities to GroEL and, furthermore, had different requirements for both GroES and ATP to become released efficiently, specific residues may serve as determinants of polypeptide binding by GroEL.[76,77] Nevertheless, these effects could equally well be attributed to mutations causing differences in branch-points or rates of putative steps in the overall folding pathway of the substrates investigated, thereby leading to crucial intermediates being populated to a very different extent, and these intermediates interact then with GroEL,[78] rather than the mutations being in direct contact with GroEL.

Our findings suggest that co-chaperonin binding is less variable than substrate binding, and that this interaction has to be maintained to assure the GroEL–GroES reaction cycle. If the binding site is altered, another copy of GroES may be needed as well, which can interact with the mutated $GroEL_{variant}$. Too tight as well as too weak an interaction will result in a loss of chaperonin function, as the strength of interaction is crucial for the rate of the chaperonin cycling and hence for chaperonin functionality.[75]

As an alternative, group II chaperonins might offer an easier possibility for substrate optimization: First, they have already evolved different substrate-binding sites and second, they have a built-in lid, and thus upon alteration of the substrate-binding site, cofactor binding is not affected. Future experiments might elucidate this possibility.

Materials and Methods

Molecular biology

Unless stated otherwise, all experiments were performed according to standard protocols.[79] Enzymes and buffers were from New England Biolabs (NEB; Beverly, MA) or Fermentas (Vilnius, Lithuania). All PCR reactions were performed using the proofreading Pfu^{Turbo}-polymerase (Stratagene).

Plasmids

Plasmids used in this study are listed in Table 3. The sequences of all inserts in plasmids that were generated by PCR were confirmed by sequencing.

The vector for the expression of GFP_{wt} in all experiments, pAT115-wtGFP, was constructed by inserting the PCR-amplified *gfp* gene into pAT115,[80] thereby replacing its NcoI/HindIII fragment. The GFP gene sequence was amplified from pBADGFPAC2[81] with oligonucleotides MAK_#04F (5′-GTGAATTCGGTAC CATGGGTAAGGGA GAAG-3′) and MAK_#09R (5′-GCGGAAGCTTCATTAA GATCTCAGATCCTCTTCTGAGATGAGTTTTTGTTCG-GATCCAGATTTGTAC-3′), the product was digested with the restriction enzymes NcoI and HindIII and ligated into pAT115 digested with the same restriction enzymes. The back-mutations of this engineered "enhanced GFP"[81] to the wt sequence (AS2G, S100F, T154M, A164V, I168T) were introduced by site-directed mutagenesis.

pZA21_GroESL is a derivative of pZA21[47] and expresses the *groES* and the *groEL* genes bicistronically under control of the promoter $P_{LtetO-1}$. It was constructed by inserting the PCR-amplified *groEL/S* genes into pZA21 cut by KpnI and HindIII. The *groEL/S* genes were amplified from pOF39[82] with oligonucleotides MAK_#16F (5′-CGGAATTCCGGT-TACCATATGAATATTCGTCC-3′) and MAK_#17R (5′-CGGAATTCAAGCTTCATTACATCATGCC-3′). The product was digested with the restriction enzymes KpnI and HindIII and ligated into pZA21 digested with the same restriction enzymes.

pZA21_BsaI is a derivative of pZA21_GroESL and contains two BsaI recognition sites at base-pairs 1122

Table 3. Plasmids

Plasmid	Relevant details	Source/reference
pZA21_ GroESL	Kan^R, p15A, encodes GroESL under control of $P_{LtetO-1}$	This work
pZA21_BsaI	Kan^R, p15A, encodes GroESL under control of $P_{LtetO-1}$; carries two BsaI recognition sites at positions 1122 and 1199 for uptake of the GroEL apical domain library	This work
pZA21_ GroESLib	Kan^R, p15A, encodes *groES* gene and *groEL* apical domain library under control of $P_{LtetO-1}$	This work
pAT115-wtGFP	Ap^R, ColE1, encodes RGS-His_6-tag-wtGFP under control of $P_{T5/lac}$	This work
pQE16	Ap^R, ColE1, encodes mDHFR-His_6-tag under control of $P_{T5/lac}$	QIAgen, Hilden, Germany
pAT115	Ap^R, ColE1, RGS-His_6-tag-gpD under control of $P_{T5/lac}$	80
pOF39	Ap^R, ColE1, encodes GroESL	82
pZA21	Kan^R, p15A, $P_{LtetO-1}$	47
pBADGFPAC2	Ap^R, ColE, encodes GFP under control of P_{BAD}	81

Ap^R, ampicillin resistant; Kan^R, kanamycin resistant.

and 1199 within the *groEL* gene for uptake of the GroEL apical domain library. It was constructed by PCR amplification of pZA21_GroESL with oligonucleotides MAK_#26F (5′-GCACGAATTCATTAAAGAG-3′), MAK_#27R (5′-TCCCCCCGGGTACCGTAGCGGCTTA-TAGCTACGGTTTGGTCTCGGAGACAGGTAGCCA-3′), MAK_#26R (5′-CTCTTTAATGAATTCGGTC-3′) and MAK_#27F (5′-TCCCCCCGGGTACCGTAGCGGCTTA-TAGCTACGGTTGGTCTCTGAAAGTCGCTGCGG-3′). The two products were digested with the restriction enzymes EcoRI and XmaI and ligated to yield pZA21_BsaI.

pZA21_GroESLib is a derivative of pZA21_BsaI and contains the GroEL apical domain library between basepairs 1123 and 1334. It was constructed by digesting the GroEL apical domain library with the restriction enzyme BsaI, which was then ligated into pZA21_BsaI cut with the same restriction enzyme.

Library construction

The sites within the apical domain targeted for randomization span a region of ~211 bp on the DNA level and therefore cannot be assembled in a one-step gene synthesis. Therefore, this part of the *groEL* gene was assembled by overlapping-extension-ligation-PCR of six oligonucleotides harboring degenerated codons of the NNK and NBB-type (N = A, T, G, C; K = G, T; B = C, G, T; see also Table 1 in Supplementary Data) at the respective positions: oli_01 (5′-CGGGATCCGCGGTCTCGTCTC-CTNNKTTCATCAACAAGCCGGAAACTGG-3′), oli_02 (5′-GTTGTTCGGCCTTTGACCGCGTCATCTTGACCT-TTCGGGCAAGTAGGACGACCGACTGTTCTTTTAG-AGG-3′), oli_03 (5′-GCTGACAAGAAAATCTCCAACN-NKNNKGAAATGCTGCCGGTTCTG-3′), oli_04 (5′-GCGATGATCAGCAGCGGTTTGCCTGCMNNGGCAA-CAGCMNNCAGAACCGGCAGCATTTC-3′), oli_05 (5′-CCGCTGCTGATCATCGCTGAAGATGTAGAAGGCG-AAGCGCTGGCAACTCTGG-3′) and oli_06 (5′-CCCA-AGCTTCGCGGTCTCCTTTCACVVNGCCMNN-MNN-GGTGTTVVNAACCAGAGTTGCCAGCGCTTCG-3′). The assembled and randomized part of the *groEL* gene was further amplified by "outer" oligonucleotides MAK_#25F (5′-CGGGATCCGCGGTCTCGTC-3′) and MAK_#25R (5′-CCCAAGCTTCGCGGTCTCCTTTC-3′), annealing to the conserved 5′ and 3′ ends of this DNA stretch. This PCR product was then digested with the restriction enzyme BsaI and ligated into pZA21_BsaI, cut with the same restriction enzyme.

Protein production and purification

$GroEL_{wt}$ and $GroEL_{variants}$ were expressed from plasmids $pZA21\_GroESLib_{wt}$ and $pZA21\_GroESLib_{variant}$, respectively, in *E. coli* XL-1 Blue F′ and purified as described.[83] His-tagged GFP_{wt} and mDHFR were expressed from plasmids pAT115-wtGFP or pQE16 in *E. coli* XL-1 Blue F′ and purified according to the protocols of the manufacturer (Qiagen GmbH, Hilden, Germany). GroES and bovine liver rhodanese were obtained from Sigma. Protein concentrations were determined spectrophotometrically as described.[84]

Chaperonin assays in aqueous solutions and in the presence of crowding agents

Protein refolding of GFP in aqueous solutions and in the presence of crowding agents was performed as described.[23] Briefly, GFP denatured in 6 M Gdn·HCl was bound to GroEL by 150-fold dilution into buffer or buffer containing 28% (v/v) Ficoll. Samples always contained at least a fivefold molar excess of GroEL. Refolding was initiated by adding 3 mM ATP and GroES (fivefold to tenfold molar excess over total GroEL) to the respective buffer solutions. The fluorescence of native GFP was measured as described.[48]

mDHFR refolding was done following the protocol described by Viitanen *et al*.[52] Briefly, mDHFR was denatured in 5 M Gdn·HCl and bound to GroEL by 150-fold dilution into buffer containing at least a fivefold molar excess of GroEL. Refolding was initiated by adding 3 mM ATP and the recovery of enzyme activity was followed spectrophotometrically by the oxidation of NADPH in the presence of dihydrofolate at 340 nm.

Rhodanese refolding and measurement of rhodanese enzyme activity was done following published procedures.[51]

Suppression of rhodanese aggregation was done as described elsewhere.[85] Briefly, rhodanese was denatured in 6 M Gdn·HCl and diluted 100-fold into buffer containing a stoichiometric amount of the chaperonins. The formation of protein aggregates was followed spectrophotometrically by light-scattering at 320 nm.

Assaying the GroEL/GroES interaction by protease protection followed published protocols.[53] In brief, GroEL was incubated with or without GroES in buffer containing proteinase K or no proteinase. Proteolytic digestion of the C-terminal residues of GroEL was stopped by the addition of 1 M PMSF to all samples and samples were subsequently analyzed by SDS-PAGE.

All assays were performed at 25 °C.

In vivo screening and selection experiments

In vivo screening for substrate-optimized GroEL variants was done by introducing the GroEL apical domain library on pZA21_GroESLib into *E. coli* DH5αZ1 [pAT115-wtGFP][47] and plated on LB agar plates containing 1% (w/v) glucose, 50 μg/ml of ampicillin, 30 μg/ml of kanamycin, 100 ng/ml of anhydrotetracycline, and 50 μM IPTG. Cells were grown at 37 °C overnight, left at room temperature for a maximum of 10 h, followed by visual inspection of UV-illuminated plates for bright green or pale/colorless colonies.

In vivo complementation of $GroEL_{wt}$ at 37 °C in *E. coli* AI90 [pBAD50] was performed as described.[46]

DNA of positive clones found in these assays was isolated and re-introduced into the respective *E. coli* strain, and the screening/selection step was repeated to confirm the phenotype and to eliminate false positives. The DNA of a subset of clones that were twice confirmed as positive was sequenced using standard DNA sequencing.

In vivo GFP folding assay

E. coli DH5αZ1 cells containing pAT115-wtGFP and one of the respective GroEL variants on pZA21_GroESLib were grown under screening conditions (see above). Cells were harvested, normalized to A_{600} and disrupted by sonification. The fluorescence of native GFP in the soluble fraction was measured as described.[48] The fluorescence signal of cells containing $GroEL_{wt}$ was set to 1.

Acknowledgements

The authors sincerely thank Dr Patrik Forrer for his continuous technical advice throughout the project and for the critical reading of the manuscript, Dr Jörg Martin for valuable technical advice and information concerning refolding experiments, Drs Hermann Bujard and Peter Lund for kindly providing *E. coli* strains and plasmids, and members of the Plückthun laboratory M66 for valuable discussions. This work was supported by the Schweizerische Nationalfonds.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.12.005

References

1. Jaenicke, R. & Seckler, R. (1997). Protein misassembly *in vitro*. *Advan. Protein Chem.* **50**, 1–59.
2. Hartl, F. U. & Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **295**, 1852–1858.
3. Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P. *et al*. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature*, **333**, 330–334.
4. Ellis, R. J. (1996). *The Chaperonins*, Academic Press, San Diego, CA.
5. Glass, J. I., Lefkowitz, E. J., Glass, J. S., Heiner, C. R., Chen, E. Y. & Cassell, G. H. (2000). The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature*, **407**, 757–762.
6. Qamra, R., Srinivas, V. & Mande, S. C. (2004). *Mycobacterium tuberculosis* GroEL homologues unusually exist as lower oligomers and retain the ability to suppress aggregation of substrate proteins. *J. Mol. Biol.* **342**, 605–617.
7. Fenton, W. A. & Horwich, A. L. (2003). Chaperonin-mediated protein folding: fate of substrate polypeptide. *Quart. Rev. Biophys.* **36**, 229–256.
8. Gutsche, I., Essen, L. O. & Baumeister, W. (1999). Group II chaperonins: new TRiC(k)s and turns of a protein folding machine. *J. Mol. Biol.* **293**, 295–312.
9. Bukau, B. & Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell*, **92**, 351–366.
10. Carrascosa, J. L., Llorca, O. & Valpuesta, J. M. (2001). Structural comparison of prokaryotic and eukaryotic chaperonins. *Micron*, **32**, 43–50.
11. Braig, K., Adams, P. D. & Brunger, A. T. (1995). Conformational variability in the refined structure of the chaperonin GroEL at 2.8 Å resolution. *Nature Struct. Biol.* **2**, 1083–1094.
12. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. *Nature*, **371**, 614–619.
13. Ditzel, L., Lowe, J., Stock, D., Stetter, K. O., Huber, H., Huber, R. & Steinbacher, S. (1998). Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell*, **93**, 125–138.
14. Chaudhuri, T. K., Farr, G. W., Fenton, W. A., Rospert, S. & Horwich, A. L. (2001). GroEL/GroES-mediated folding of a protein too large to be encapsulated. *Cell*, **107**, 235–246.
15. Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R. & Horwich, A. L. (1999). GroEL–GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. *Cell*, **97**, 325–338.
16. Xu, Z. & Sigler, P. B. (1998). GroEL/GroES: structure and function of a two-stroke folding machine. *J. Struct. Biol.* **124**, 129–141.
17. Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M. & Horwich, A. L. (1996). Characterization of the active intermediate of a GroEL–GroES-mediated protein folding reaction. *Cell*, **84**, 481–490.
18. Mayhew, M., da Silva, A. C., Martin, J., Erdjument-Bromage, H., Tempst, P. & Hartl, F. U. (1996). Protein folding in the central cavity of the GroEL–GroES chaperonin complex. *Nature*, **379**, 420–426.
19. Ranson, N. A., Burston, S. G. & Clarke, A. R. (1997). Binding, encapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. *J. Mol. Biol.* **266**, 656–664.
20. Wang, J. D., Michelitsch, M. D. & Weissman, J. S. (1998). GroEL–GroES-mediated protein folding requires an intact central cavity. *Proc. Natl Acad. Sci. USA*, **95**, 12163–12168.
21. Schmidt, M., Buchner, J., Todd, M. J., Lorimer, G. H. & Viitanen, P. V. (1994). On the role of groES in the chaperonin-assisted folding reaction. Three case studies. *J. Biol. Chem.* **269**, 10304–10311.
22. 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.
23. Martin, J. (2002). Requirement for GroEL/GroES-dependent protein folding under nonpermissive

conditions of macromolecular crowding. *Biochemistry*, **41**, 5050–5055.
24. Ewalt, K. L., Hendrick, J. P., Houry, W. A. & Hartl, F. U. (1997). *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell*, **90**, 491–500.
25. Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F. & Hartl, F. U. (1999). Identification of *in vivo* substrates of the chaperonin GroEL. *Nature*, **402**, 147–154.
26. Viitanen, P. V., Gatenby, A. A. & Lorimer, G. H. (1992). Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of *Escherichia coli* proteins. *Protein Sci.* **1**, 363–369.
27. Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H.-C., Stines, A. P. *et al*. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell*, **122**, 209–220.
28. Spiess, C., Meyer, A. S., Reissmann, S. & Frydman, J. (2004). Mechanism of the eukaryotic chaperonin: protein folding in the chamber of secrets. *Trends Cell Biol.* **14**, 598–604.
29. Tian, G., Vainberg, I. E., Tap, W. D., Lewis, S. A. & Cowan, N. J. (1995). Specificity in chaperonin-mediated protein folding. *Nature*, **375**, 250–253.
30. Sakikawa, C., Taguchi, H., Makino, Y. & Yoshida, M. (1999). On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. *J. Biol. Chem.* **274**, 21251–21256.
31. Klunker, D., Haas, B., Hirtreiter, A., Figueiredo, L., Naylor, D. J., Pfeifer, G. *et al*. (2003). Coexistence of group I and group II chaperonins in the archaeon *Methanosarcina mazei*. *J. Biol. Chem.* **278**, 33256–33267.
32. Kim, S., Schilke, B., Craig, E. A. & Horwich, A. L. (1998). Folding *in vivo* of a newly translated yeast cytosolic enzyme is mediated by the SSA class of cytosolic yeast Hsp70 proteins. *Proc. Natl Acad. Sci. USA*, **95**, 12860–12865.
33. Xu, H.-M., Zhang, G.-Y., Ji, X.-D., Cao, L., Shu, L. & Hua, Z.-C. (2005). Expression of soluble, biologically active recombinant human endostatin in *Escherichia coli*. *Protein Expr. Purif.* **41**, 252–258.
34. Dale, G. E., Schönfeld, H. J., Langen, H. & Stieger, M. (1994). Increased solubility of trimethoprim-resistant type S1 DHFR from *Staphylococcus aureus* in *Escherichia coli* cells overproducing the chaperonins GroEL and GroES. *Protein Eng.* **7**, 925–931.
35. Carrillo, N., Ceccarelli, E., Krapp, A., Boggio, S., Ferreyra, R. & Viale, A. (1992). Assembly of plant ferredoxin-$NADP^+$ oxidoreductase in *Escherichia coli* requires GroE molecular chaperones. *J. Biol. Chem.* **267**, 15537–15541.
36. Altamirano, M. M., Golbik, R., Zahn, R., Buckle, A. M. & Fersht, A. R. (1997). Refolding chromatography with immobilized mini-chaperones. *Proc. Natl Acad. Sci. USA*, **94**, 3576–3578.
37. Zahn, R., Buckle, A. M., Perrett, S., Johnson, C. M., Corrales, F. J., Golbik, R. & Fersht, A. R. (1996). Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. *Proc. Natl Acad. Sci. USA*, **93**, 15024–15029.
38. Ben-Zvi, A. P. & Goloubinoff, P. (2001). Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. *J. Struct. Biol.* **135**, 84–93.
39. Horwich, A. L. (2004). Chaperoned protein disaggregation—the ClpB ring uses its central channel. *Cell*, **119**, 579–581.
40. Wang, J. D., Herman, C., Tipton, K. A., Gross, C. A. & Weissman, J. S. (2002). Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell*, **111**, 1027–1039.
41. Dunn, A. Y., Melville, M. W. & Frydman, J. (2001). Review: cellular substrates of the eukaryotic chaperonin TRiC/CCT. *J. Struct. Biol.* **135**, 176–184.
42. Buckle, A. M., Zahn, R. & Fersht, A. R. (1997). A structural model for GroEL–polypeptide recognition. *Proc. Natl Acad. Sci. USA*, **94**, 3571–3575.
43. Chen, L. & Sigler, P. B. (1999). The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. *Cell*, **99**, 757–768.
44. Kobayashi, N., Freund, S. M., Chatellier, J., Zahn, R. & Fersht, A. R. (1999). NMR analysis of the binding of a rhodanese peptide to a minichaperone in solution. *J. Mol. Biol.* **292**, 181–190.
45. Wang, Q., Buckle, A. M. & Fersht, A. R. (2000). From minichaperone to GroEL 1: information on GroEL–polypeptide interactions from crystal packing of minichaperones. *J. Mol. Biol.* **304**, 873–881.
46. Ivic, A., Olden, D., Wallington, E. J. & Lund, P. A. (1997). Deletion of *Escherichia coli* GroEL is complemented by a *Rhizobium leguminosarum* GroEL homologue at 37 °C but not at 43 °C. *Gene*, **194**, 1–8.
47. Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli via* the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucl. Acids Res.* **25**, 1203–1210.
48. Makino, Y., Amada, K., Taguchi, H. & Yoshida, M. (1997). Chaperonin-mediated folding of green fluorescent protein. *J. Biol. Chem.* **272**, 12468–12474.
49. Martin, J. & Hartl, F. U. (1997). The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proc. Natl Acad. Sci. USA*, **94**, 1107–1112.
50. Gomez-Puertas, P., Martin-Benito, J., Carrascosa, J. L., Willison, K. R. & Valpuesta, J. M. (2004). The substrate recognition mechanisms in chaperonins. *J. Mol. Recogn.* **17**, 85–94.
51. Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L. & Hartl, F. U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a "molten globule"-like intermediate. *Nature*, **352**, 36–42.
52. Viitanen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. & Gatenby, A. A. (1991). Complex interactions between the chaperonin 60 molecular chaperone and dihydrofolate reductase. *Biochemistry*, **30**, 9716–9723.
53. Langer, T., Pfeifer, G., Martin, J., Baumeister, W. & Hartl, F. U. (1992). Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *EMBO J.* **11**, 4757–4765.
54. Jones, S. & Thornton, J. M. (1995). Protein–protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.* **63**, 31–65.
55. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–230.
56. Zahn, R., Spitzfaden, C., Ottiger, M., Wüthrich, K. & Plückthun, A. (1994). Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature*, **368**, 261–265.
57. Nieba-Axmann, S. E., Ottiger, M., Wüthrich, K. & Plückthun, A. (1997). Multiple cycles of global unfolding of GroEL-bound cyclophilin A evidenced by NMR. *J. Mol. Biol.* **271**, 803–818.
58. Todd, M. J., Lorimer, G. H. & Thirumalai, D. (1996). Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl Acad. Sci. USA*, **93**, 4030–4035.

59. Zahn, R., Perrett, S., Stenberg, G. & Fersht, A. R. (1996). Catalysis of amide proton exchange by the molecular chaperones GroEL and SecB. *Science*, **271**, 642–645.
60. Jackson, G. S., Staniforth, R. A., Halsall, D. J., Atkinson, T., Holbrook, J. J., Clarke, A. R. & Burston, S. G. (1993). Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: implications for the mechanism of assisted protein folding. *Biochemistry*, **32**, 2554–2563.
61. Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U. & Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. *Cell*, **107**, 223–233.
62. Chatellier, J., Hill, F., Lund, P. A. & Fersht, A. R. (1998). *In vivo* activities of GroEL minichaperones. *Proc. Natl Acad. Sci. USA*, **95**, 9861–9866.
63. Lorimer, G. H. (1996). A quantitative assessment of the role of the chaperonin proteins in protein folding *in vivo*. *FASEB J.* **10**, 5–9.
64. Baumketner, A., Jewett, A. & Shea, J. E. (2003). Effects of confinement in chaperonin assisted protein folding: rate enhancement by decreasing the roughness of the folding energy landscape. *J. Mol. Biol.* **332**, 701–713.
65. Betancourt, M. R. & Thirumalai, D. (1999). Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity. *J. Mol. Biol.* **287**, 627–644.
66. Yifrach, O. & Horovitz, A. (2000). Coupling between protein folding and allostery in the GroE chaperonin system. *Proc. Natl Acad. Sci. USA*, **97**, 1521–1524.
67. Grallert, H., Rutkat, K. & Buchner, J. (2000). Limits of protein folding inside GroE complexes. *J. Biol. Chem.* **275**, 20424–20430.
68. Lin, Z., Schwartz, F. P. & Eisenstein, E. (1995). The hydrophobic nature of GroEL–substrate binding. *J. Biol. Chem.* **270**, 1011–1014.
69. Clarke, A. R. & Lund, P. A. (1996). Kinetic and energetic aspects of chaperonin function. In *The Chaperonins*, pp. 167–212, Academic Press, London.
70. Xu, Z., Horwich, A. L. & Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex. *Nature*, **388**, 741–750.
71. Kawata, Y., Kawagoe, M., Hongo, K., Miyazaki, T., Higurashi, T., Mizobata, T. & Nagai, J. (1999). Functional communications between the apical and equatorial domains of GroEL through the intermediate domain. *Biochemistry*, **38**, 15731–15740.
72. Horovitz, A., Fridmann, Y., Kafri, G. & Yifrach, O. (2001). Review: allostery in chaperonins. *J. Struct. Biol.* **135**, 104–114.
73. Hunt, J. F., van der Vies, S. M., Henry, L. & Deisenhofer, J. (1997). Structural adaptations in the specialized bacteriophage T4 co-chaperonin Gp31 expand the size of the Anfinsen cage. *Cell*, **90**, 361–371.
74. Fridmann, Y., Ulitzur, S. & Horovitz, A. (2000). *In vivo* and *in vitro* function of GroEL mutants with impaired allosteric properties. *J. Biol. Chem.* **275**, 37951–37956.
75. Shewmaker, F., Kerner, M. J., Hayer-Hartl, M., Klein, G., Georgopoulos, C. & Landry, S. J. (2004). A mobile loop order–disorder transition modulates the speed of chaperonin cycling. *Protein Sci.* **13**, 2139–2148.
76. Staniforth, R. A., Cortes, A., Burston, S. G., Atkinson, T., Holbrook, J. J. & Clarke, A. R. (1994). The stability and hydrophobicity of cytosolic and mitochondrial malate dehydrogenases and their relation to chaperonin-assisted folding. *FEBS Letters*, **344**, 129–135.
77. Zahn, R., Lindner, P., Axmann, S. E. & Plückthun, A. (1996). Effect of single point mutations in citrate synthase on binding to GroEL. *FEBS Letters*, **380**, 152–156.
78. Clark, A. C. & Frieden, C. (1997). GroEL-mediated folding of structurally homologous dihydrofolate reductases. *J. Mol. Biol.* **268**, 512–525.
79. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
80. Stumpp, M. T. (2004). From leucine-rich repeat proteins to ankyrin repeat proteins. PhD dissertation, University of Zürich.
81. Crameri, A., Whitehorn, E. A., Tate, E. & Stemmer, W. P. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotechnol.* **14**, 315–319.
82. Fayet, O., Louarn, J. M. & Georgopoulos, C. (1986). Suppression of the *Escherichia coli* dnaA46 mutation by amplification of the *groES* and *groEL* genes. *Mol. Gen. Genet.* **202**, 435–445.
83. Zahn, R., Axmann, S. E., Rücknagel, K. P., Jaeger, E., Laminet, A. A. & Plückthun, A. (1994). Thermodynamic partitioning model for hydrophobic binding of polypeptides by GroEL. I. GroEL recognizes the signal sequences of beta-lactamase precursor. *J. Mol. Biol.* **242**, 150–164.
84. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.
85. Weber, F. & Hayer-Hartl, M. (2000). Prevention of rhodanese aggregation by the chaperonin GroEL. In *Chaperonin Protocols* (Schneider, C., ed.), vol. 140, pp. 111–115, Humana Press, Totowa, NJ.
86. Miller, S., Janin, J., Lesk, A. M. & Chothia, C. (1987). Interior and surface of monomeric proteins. *J. Mol. Biol.* **196**, 641–656.
87. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
88. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature*, **371**, 578–586.
89. Laskowski, R. A. (1995). SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.* **13**, 323–330.
90. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.

Edited by C. R. Matthews

(Received 5 August 2005; received in revised form 23 November 2005; accepted 1 December 2005)
Available online 20 December 2005