Molecular Chaperone GroEL – toward a Nano Toolkit in Protein Engineering, Production and Pharmacy

Alexey N. Fedorov and Maria S. Yurkova
A.N. Bach Institute of Biochemistry, Research Center of Biotechnology of the Russian Academy of Sciences, 119071, Russian Federation

Abstract

In this review, we discuss the potential of molecular chaperones belonging to GroEL family as applied to different practical uses. The GroEL, often referred to as chaperonin, plays a fundamental role in protein folding in cells and is essential for the proper folding of a number of proteins. Early on in studying chaperones it has been considered, along with other chaperones, as an attractive tool to assist recombinant protein folding in cells and their refolding in vitro. Several different schemes for using GroEL have been applied for difficult-to-work-with recombinant targets, especially those insoluble and unstable. Fusion constructs of difficult recombinant targets with GroEL as a carrier have been reported to make them better soluble. However, the complex structure of GroEL particle, consisting of two hexamer rings, often makes it difficult to take full advantage of its use as a versatile (re)folding tool. Subsequently, a separately expressed GroEL apical domain, minichaperone, was not only extensively studied as representative of the substrate-binding site of the full-sized GroEL, but also used as an aiding tool for refolding of recombinant proteins in different settings. In separate venues of research, immune-modulating activity of chaperonins has been studied. Mycobacterial GroEL has been used as an immune-stimulating carrier to make therapeutic vaccines. In spite of the progress made, to really meet all the expectations for GroELs, minichaperone and other chaperonin forms, bioengineering and further crafting is required to make them truly useful nano tools for specific applications.

Keywords

GroEL, Protein folding, Minichaperone, Immune-modulating activity, T-cell immune response.

Introduction

Molecular chaperones, initially discovered as heat-shock proteins, are diverse groups of proteins united by one fundamental feature – prevention of protein misfolding and aiding protein folding and stability. Chaperones play important and often indispensable role in the process of biosynthetic protein folding in general and cotranslational folding in particular [1-3]. They form several families with quite conservative structure among species inside a family and various mechanisms of action. Their main function is maintaining solubility and structure of cell proteins not only under stress conditions but also upon regular cell metabolism and functioning [4]. It is this fundamental feature which makes them natural and attractive objects for aiding “difficult” target recombinant proteins, i.e. proteins with limited solubility, impaired ability to fold etc. both in vivo and in vitro. This review is focused on using GroEL family in this regard.
The GroEL, sometimes referred to as chaperonin, is by far the most well studied molecular chaperone. Its mechanisms of functioning in aiding protein folding are studied in great detail [4-6]. The structure of GroEL is extensively studied, and three-dimensional structures are published for GroELs from many sources [6, 7]. GroELs are subdivided into two types [6, 7], recently the third type was described [8]. Mycobacterial GroEL is a stand-alone creature both figural and literally as it exists as a monomer and may also form a dimer.

In most organisms GroEL forms a complex tertiary structure of adjacent two rings consisting of seven subunits each. For GroEL type II the structure is further stabilized with its co-chaperone GroES (Figure 1).

The rings form a cavity inside them with hydrophobic surface, where substrate proteins are bound. Substrate proteins binding and folding occurs in ATP-dependent manner, as is schematically shown in figure 2. Early on in studying GroEL emerged a natural idea of using this chaperone along with other chaperones to aid folding of difficult-to-fold and aggregation-prone recombinant proteins. This has been tried in different approaches, and using not only the whole GroEL/ES particle, but also GroEL as a monomer, and separately its apical domain.

![Figure 1: The structure of GroEL/ES particle. On the right, the scheme of GroEL monomer is given, with its domains shown in corresponding colors.](image)

**GroEL as a folding and stabilizing tool**

The GroEL is able to greatly stabilize proteins and even more than that, maximize the yield of the biologically active form of the protein. Bacterial luciferase, a heterodimer, consists of extremely thermo labile subunits, which alone are inactivated at 30 °C within a minute. The individual subunits bind tightly to GroEL and stay indefinitely long at 37 °C in a folding-competent conformation within the chaperone particle [9]. At the permissive temperature range the chaperone decreases the rate of formation of the folding-incapable species and boosts the final yield of active enzyme by accelerating the rate of the productive folding pathway. These results suggest a mechanism by which the GroEL acts to maximize the yield of the biologically active form of the protein by accelerating the kinetics of folding reactions.

One of the approaches to express “difficult” recombinant proteins in bacterial cells is co-expressing GroEL and other chaperones. Using the natural potential of chaperones in their co-expression in host cells with target proteins which tend to misfold and aggregate in order to obtain their soluble forms has been demonstrated in Kim et al. [10], where the production of active cycloexextrin glycosyltransferase (CGTase) of Bacillus macerans, otherwise mainly expressed in inclusion bodies, was greatly enhanced by co-expression of folding accessory proteins, such as human peptidyl-prolyl cis-trans isomerase followed by co-expression of DnaK-DnaJ-GrpE together with GroEL-GroES. Similar results were reported for the expression in E. coli system of codon-optimized sarcosine oxidase from Thermomicrobium roseum, whose soluble expression was significantly enhanced via the co-expression of chaperones [11], for the expression of humanized single-chain antibody in E. coli [12], and others. Same strategy applies for the expression in yeast system, as, for instance, for the expression of xylose isomerase in Saccharomyces cerevisiae in functional state [13]. There is a quite recent interesting work describing that co-expressed GroEL-GroES assisted folding of simultaneously over-expressed proteins maltodextrin glucosidase and yeast mitochondrial aconitase, both of which are relatively large and prone to aggregation [14]. It should be noted that despite some successful stories in most instances co-expression of chaperones didn’t help to produce soluble recombinant proteins. Very brief explanation is that GroEL may help when productive folding is limited without participation of chaperones and GroEL in particular. In this scenario, increasing GroEL concentration may shift the balance between aggregation and productive folding toward the latter. However, if the final target protein structure is poorly soluble itself and forms aggregates, the only help it can get is to stay soluble while bound to and shielded within the GroEL particle. This, however, requires one to one stoichiometry between a target polypeptide and the entire GroEL particle, huge, nearly 1 MDa complex, which substantially limits practicality of this approach.

Yet another interesting spin is studying structures of GroEL/GroELS by a variety of techniques [4-6, 15] and applying these approaches to studying protein and peptide structures in a bound state with GroEL by X-ray, NMR and cryo-electron microscopy [16-20].

**GroEL as a carrier in fusion constructs**

Another approach is based on using GroEL in cis with target proteins, i.e. making fusion protein constructs of GroEL and designated target polypeptide. The problem with using native unperturbed chaperonin is that, in its native
conformation, GroEL type I has its N- and C-ends well buried inside the equatorial domain of the protein globule, and connecting a fusion partner to any of its ends destabilizes the structure of the oligomer as well as GroEL monomer, which makes it functionally unfit. This steric obstruction greatly reduces the possibilities of using fusions with GroEL for aiding target protein folding and stability. Indeed, while fusion system with GroEL as a carrier has been reported [21], it has never been used by researchers due to the limitations described above.

A very attractive option would be to insert a target polypeptide inside the GroEL cavity thus making the target isolated from cell cytoplasm, which is one of positive considerations if planning to use chaperonin as a leader for a fusion construct. The idea seems not to work for GroEL type I due to its tertiary structure. The problem due to positioning of its N- and C-ends has been described above. As for inserts along the GroEL polypeptide chain, there was an attempt made to subject groEL gene to insertion mutagenesis using transposon ISlacZ/in [22]. Four GroEL-LacZ fusions and the corresponding insertion mutants were obtained, but none of them retained the ability to fold properly, and three out of four were totally or partially degraded. The conclusion has been made that it’s impossible to insert polypeptide sequences within the GroEL polypeptide chain.

Still, there is an example of successful expression of a model protein fused to GroEL type II in an ingenious work by M. Furutani et al. [23]. In this work, Thermococcus sp. KS-1 chaperonin α-subunit (TCP) which forms an eightfold symmetric double-ring structure was used. Expression plasmids were constructed which carried two or four TCP genes ligated head to tail in phase and a target protein gene at the 3’-end of the linked TCP genes. It was shown that a fusion protein of tetra-TCP and green fluorescent protein was able to form the double-ring complexes with green fluorescence. Also, using this tetra-TCP fusion strategy, two virus structural proteins (21–25 kDa) toxic to host cells or two antibody fragments (25–36 kDa) prone to aggregate were expressed in the soluble fraction of Escherichia coli.

Refolding device

Another direction for using GroEL is making it a "refolding device" for renaturation of difficult recombinant proteins. There are many works studying the interactions of chaperones with substrate proteins in vitro in order to facilitate their correct refolding after expression. In Pack et al. [24] the importance of electrostatic interactions in binding of substrate to GroEL is shown, as well as the preference of GroEL to a positively charged substrates. The work of Teshima et al. [25], demonstrates that GroEL is able to assist denatured proteins in their refolding being itself immobilized onto carboxylated poly(styrene/acrylamide) microspheres, as schematically shown on figure 3. Thermus thermophilus holo-chaperonin was covalently immobilized onto these microspheres and showed sufficiently high ability to facilitate refolding of two guanidine hydrochloride-denatured enzymes, of Leuconostoc mesenteroides glucose-6-phosphate dehydrogenase and of pig heart lactate dehydrogenase at 30 °C, and of Bacillus stearothermophilus lactate dehydrogenase at 60 °C. Microspheres with covalently bound GroEL could be used repeatedly, and the conclusion is made that poly (styrene/ acrylamide) microspheres on which approximately 10 mg/m² of T. thermophilus holo-chaperonin is immobilized are very effective for refolding of various guanidine hydrochloride-denatured enzymes over a wide temperature range. Also, the ability of GroEL to prevent aggregation of client proteins was studied as compared with human small heat shock protein HspB5 (α-crystallin) and synthetic polyanion, poly (styrene sulfonate) [26]. Assistance of GroEL in the productive (re) folding of proteins would be diminished if it is blocked by misfolded proteins or polypeptides forms that in principle cannot fold correctly [27].

Functionality in minimal size: minichaperone

The practical usefulness of GroEL particle is limited by its overall size and a complex quaternary structure. The GroEL monomer consists of three well distinguished domains – equatorial, intermediate and apical, with different functions. It is the apical domain’s function to bind substrate proteins; and apical domain is so structurally isolated that it can be expressed as a separate protein that folds into a correct structure, same as it has in the whole native monomer GroEL, and retains its ability to bind substrate proteins and assist in their folding. The substrate binding surface is formed by α-helices 8 and 9, see figure 4. The ability of the apical domain to independent functioning was first discovered in the lab of A. Fersht and described in Zahn et al. [28]. The polypeptide corresponding to residues 191 to 345, called in this work “minichaperone”, had the activity both in facilitating the refolding of rhodanese and cyclophilin A in the absence of ATP and in catalyzing the unfolding of native barnase. This minichaperone has been
thoroughly investigated, including resolution of its crystal structure (residues 191-376) [29], its minimal functional size (residues 193-335) [30], reversibility of its folding [31, 32], the extension of substrate-binding surface and amino acid residues involved [33], mutations stabilizing the minichaperone structure [34] and salt bridges stabilizing minichaperone analog from thermophilic eubacterium Thermus thermophilus [35]. It was also demonstrated that both GroEL and its isolated apical domain form amyloid-like fibrils under physiological conditions, and that the fibrillation of the apical domain is accelerated under acidic conditions, but, despite its fibrillation propensity, the apical domain exhibits a pronounced inhibitory effect on the fibril growth of β(2)-microglobulin [36].

Of course, the elegant idea of making “minimal” chaperones still possessing fundamental chaperone capacities but more useful for practical considerations, easier to deal with and lacking complexities of native full length chaperones was explored by many investigators. Refolding of target proteins in vitro in the presence of minichaperone was performed in many instances, first for rhodanese and cyclophilin A [28], later for interferon gamma [37, 38]. Rhodanese was also used as a reporter enzyme in the development of in vitro methods of protein refolding using minichaperone. In Ramón-Luing et al. [39] the apical domain of GroEL, fused to the carbohydrate binding module type II of Cellulomonas fimi, was expressed in Escherichia coli. This recombinant protein was purified and immobilized in microcrystalline cellulose particles or cellulose fabric membranes. Assisted refolding of rhodanese by the immobilized mini-chaperone showed a two-fold improvement as compared to a control. In Antonio-Pérez et al. [40] also, the GroEL apical domain fused to a carbohydrate-binding module and immobilized on microcrystalline cellulose particles significantly improved the oxidative chromatographic refolding of lysozyme. Besides, AD, DsbA and DsbC immobilized on cellulose exhibited significant operational stability under the extreme denaturing conditions. Figures 5 and 6 illustrate the functioning of minichaperone in solution and in immobilized state, correspondingly.

The use of minichaperone as a partner for an insoluble protein in a fusion construct was also explored. For two originally insoluble proteins, E6 of human papilloma virus type 16 and N-terminal fragment of E2 of hepatitis C virus, fusion constructs were made with apical domain of Thermus thermophilus GroEL. E6 was fused to C-terminal end of GroEL apical domain, and E2 N-terminal fragment either to
its C- or N-terminal end. All three constructs were expressed in inclusion bodies, but afterwards were easily renatured and remained soluble in native buffers throughout standard biochemical procedures, including concentration, freezing-thawing and lyophilization. HPV16 E6 even retained its functional activity in fusion construct, HCV E2 does not possess any activities and, consequently, was not tested [41].

Immunological tool: adjuvant activities of mycobacterial GroEL

Besides their recognized function of assisting in folding of cell proteins, chaperones were shown to possess immune-modulating activities. They were shown to induce autoimmune response in mice not only as upregulated targets of adaptive immunity during inflammatory stress, but also as triggering factors for innate immunity through activation via Toll-like receptors [42-44]. In mice immunized with chaperones isolated from mouse tumor cells (donor cells) was observed activation of the production of CD8+ cytotoxic T lymphocytes that recognize donor cell peptides in association with the major histocompatibility complex class I proteins of the responding mouse [45]. In the aspect of immune-modulating activity, the attention of most of the investigators has been drawn to *Mycobacterium tuberculosis* GroEL. *Mycobacterium tuberculosis* is one of a small proportion of bacteria to encode two chaperonin 60 proteins. In contrast to most other bacterial species, *Mycobacterium tuberculosis* genome carries a duplicate set of cpn60 genes, one of which occurs on the groESL operon (cpn60.1), while the other is separately arranged on the chromosome (cpn60.2) [46]. Unlike chaperonins from most other species, *Mycobacterium tuberculosis* chaperonins do not form oligomers, but were shown to exist as dimers. For *M. tuberculosis* chaperonin 60.1, the oligomeric state of the active proteins has been shown to be dimers using native gel electrophoresis and protein cross-linking. In the same work, it has been demonstrated that chaperonin 60.1 from this bacterium is a very active stimulator of human monocytes, and that its monocyte-stimulating activity resides in the monomeric subunit and within this subunit the biological activity is due to the equatorial domain [47]. In another work, the crystal structure of *Mycobacterium tuberculosis* chaperonin 60.2 was published [48]. This protein also crystallized in its dimeric state. The unusual dimerization of chaperonin 60.2 leads to the exposure of certain hydrophobic patches on the surface of the protein, which might have relevance in binding to substrate peptides. It was shown that *Mycobacterium tuberculosis* chaperonins, existing in low oligomeric state, do not require ATP or the presence of co-chaperonin GroES for their refolding activity [48]. The authors hypothesize that the ATP-independent chaperones might offer benefit to the pathogen by promoting its persistence in the latent phase of its life cycle. Another feature of *Mycobacterium tuberculosis* chaperonins, and of much more practical use, is their ability to up-regulate immune response [49], which suggests their use as vaccine adjuvants targeted to infections and cancers. *Mycobacterium sp.* chaperonins are potent inducers of innate and antigen-specific immunity. They prime multiple host defense pathways, namely activate dendritic cells partly through toll-like receptors, activate natural killer cells, increase presentation of antigens to effector cells and augment T-cell and humoral immune responses against their associated antigens. It has been demonstrated that up-regulation of cellular immune response requires the presence of the protein itself, rather than DNA-vaccine, while humoral immune response can be elicited by both [50]. Immunogenic properties of mycobacterial chaperonins have been used in the development of therapeutic vaccines against cancer, autoimmune and other diseases, such as melanoma [51], diabetes [52], hepatitis B [53], and others, but there are not many examples of their practical use. One instance is that BCG vaccine, whose most potent immune-stimulating components are mycobacterial chaperonins, is successfully used in intravesical treatment of superficial bladder cancer [54]. Also, there was a study that reached as far as the second stage of clinical trials, in which was used a fusion protein comprising *Mycobacterium bovis* bacille Calmette-Guérin (BCG) hsp65 and E7 protein of human papilloma virus type 16 (HPV16). Hypothetic structure of such a fusion is shown in figure 7. Infection with HPV16 is strongly associated with a number of disease states, including cervical and anal cancers, and the persistent infection is characterized by the synthesis of the viral E6 and E7 oncoproteins in cervical epithelial cells. The expression of E7 in dysplastic and transformed cells and its recognition by the immune system as a foreign antigen make it an ideal target for immunotherapy. Studies have shown that fusion proteins with mycobacterial chaperonin as a carrier were capable of inducing potent antigen-specific activity of cytotoxic T lymphocytes in experimental animal models, which is important in eradication of HPV-induced lesions. Utilizing the E6 and E7-expressing murine tumor cell line, TC-1, as a model of cervical carcinoma, an immunotherapy based on the administration of an adjuvant-free fusion protein comprising BCG hsp65 and HPV16 E7 was developed [55, 56]. It was shown that prophylactic immunization with hspE7 protected mice against challenge with TC-1 cells and that these tumor-free animals were also protected against re-challenge.
with TC-1 cells. In addition, therapeutic immunization with hspE7 induced regression of palpable tumors, conferred protection against tumor re-challenge and was associated with long-term survival. In vivo studies using mice with targeted mutations in CD8⁺ or MHC class II or depleted of CD8⁺ or CD4⁺ lymphocyte subsets demonstrated that tumor regression following therapeutic HspE7 immunization was CD8⁺ dependent and CD4⁺ independent. Later studies demonstrated that HspE7 primed potent E7-specific CD8⁺ T cells with cytolytic and cytokine secretion activities. These CD8⁺ T-cells could differentiate into memory T cells with effector functions in the absence of CD4⁺ T-cell help [57]. Other authors added HBc antigen into the construct, demonstrating that it could induce a significant increase in tumor regression following therapeutic HspE7 immunization [58]. It was shown that it could induce a significant increase in E7-specific CD8⁺ T-cell responses, which makes it a promising therapeutic vaccine for treatment of cervical cancer with possible therapeutic potential in clinical settings.

**Protein engineering of GroEL to make it a better tool for different downstream applications**

For the full-size GroEL, the supposed directed changes in the chaperonin must be based on complete knowledge of the details of its functioning. The reaction cycle involved in GroEL-mediated protein folding for the substrates that fit into the cavity and its allosteric transitions are thoroughly described [6, 7, 59], there are some considerations concerning the folding of large substrates that do not fit into the cavity [60]. Besides hydrophobic interactions, the participation of electrostatic interactions in binding of substrate has been studied [24], and the question about requirements to the substrate itself is being discussed [61]. As for protein engineering applied to GroEL, circular permutations were used to study the structural and functional characteristics of the GroEL subunit [62]. In this work, three soluble, partially active mutants with polypeptide ends relocated into various positions of the apical domain of GroEL were isolated and studied. In all three mutants, the ability to facilitate the refolding of rhodanese was roughly equal, while the basal ATPase activity and ATPase inhibition ability to facilitate the refolding of rhodanese was roughly equal, while the basal ATPase activity and ATPase inhibition of substrate may come at the cost of diminished activity of GroEL system even for this particular protein, because proper binding of GroES is required for the chaperonin cycling. A clear consequence of these results is that the options to changes in protein-binding region of GroEL are small and require extensive effort.

**Conclusions and Future Directions**

The GroEL is being used in several different approaches to assist folding and stability of recombinant proteins in vivo as well as refolding in vitro. Its developed miniature derivative, minichaperone, carries fundamental features of the chaperonin with regard to binding and assisting folding of different recombinant proteins and has been used with recombinant proteins similar to its parental chaperonin particle. Truly exciting opportunities open up in development of GroEL-based immune biology drugs.

However, it’s clearly seen that still there is a long road for GroEL particle, its reduced form – minichaperone, mycobacterial chaperonin monomer and may be other GroEL-based constructs to become useful tools in a variety of applications, both the ones that have been tried already with mixed successes and many other new applications where they may become a clue for success. Structures and functions of chaperonins may be changed and crafted for further use for specific purposes thus making them more useful for difficult target proteins of different kinds. Making GroEL a truly efficient tool for specific purposes by protein engineering is the way that suggests itself. Making minichaperone and exploring its potential is one step in that direction.

The field of directly changing GroELs for specific substrates by means of protein engineering is in its nascent stage, but it makes a great promise.

**References**


chaperonin from a thermophilic bacterium. *Nat Commun* 8(1): 827. https://doi.org/10.1038/s41467-017-00980-z


Production and Pharmacy

Molecular Chaperone GroEL – toward a Nano Toolkit in Protein Engineering, Production and Pharmacy

Fedorov and Yurkova.

Based fusion system for producing insoluble proteins in soluble stable forms. *Protein Eng Des Sel* 29(2): 57-64. https://doi.org/10.1093/protein/gxo060


