

Dispatch
R661

Chaperomics: *In Vivo* GroEL Function Defined

A recent proteome analysis of protein folding inside cells of *Escherichia coli* predicts that only 84 of the approximately 2400 cytosolic proteins expressed in minimal media depend absolutely on the GroEL/GroES chaperone system to avoid aggregation. These proteins are enriched in α/β domains and 13 are essential for growth.

R. John Ellis

The late biochemist David Green pointed out that a clever engineer can make a vacuum cleaner from the wreck of an automobile, but this does not show that cars contain vacuum cleaners. The chaperonin GroEL will bind to virtually any partly folded protein it is presented with in the test tube, but which of the ~2400 different cytosolic protein molecules expressed in *Escherichia coli* does it actually bind inside the cell, and which of these proteins must bind to GroEL to fold correctly? The Hartl laboratory [1] has now provided comprehensive answers to these questions by a combined quantitative proteomics and functional analysis approach. Only 84 proteins are predicted to depend absolutely upon GroEL to fold correctly, but these include at least 13 essential proteins, explaining why GroEL is the only chaperone in *E. coli* required under all growth conditions tested. Almost all these 84 proteins contain either α/β or $\alpha+\beta$ domains that expose substantial hydrophobic regions during folding. The new paper [1] is the first publication in the new discipline of chaperomics, which aims to identify all the essential substrates of a given molecular chaperone *in vivo*.

A universal cellular problem is posed by the tendency of many partly folded polypeptide chains to aggregate with one another via exposed hydrophobic regions rather than form functional structures. Such partly folded chains may appear during their initial synthesis and when the mature proteins unfold either as a result of their innate instability or

after an environmental stress such as high temperature. It is the job of molecular chaperones to combat this aggregation problem [2]. GroEL binds to its cochaperone GroES to create a large complex that prevents aggregation by encapsulating individual polypeptide chains one at a time inside a molecular cage (Figure 1). Inside this cage, the chain continues to fold in the absence of other folding chains until the hydrophobic residues that cause aggregation are buried within the final folded structure. This complex is termed an Anfinsen cage [3] to indicate that inside this cage the chain folds spontaneously in a manner determined by its amino acid sequence, as it does when diluted from denaturant in the classical refolding experiment of Anfinsen [4]. Small chaperones such as DnaK/J also exist that do not form a cage, but simply cycle on and off exposed hydrophobic sequences on partly folded chains, thus reducing the time

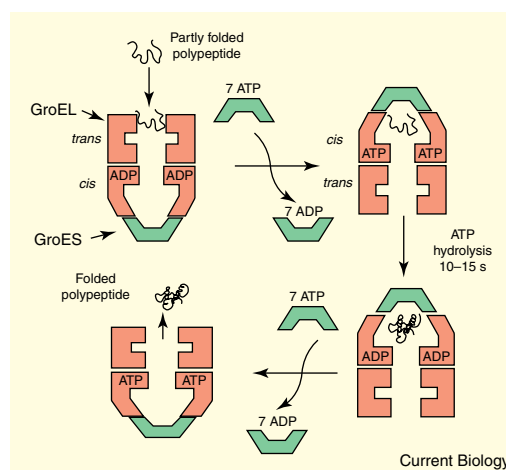
available for them to aggregate with one another [5].

Most studies on chaperones use *in vitro* refolding systems to define their properties, but what matters is not what chaperones can be persuaded to do in such artificial environments but what they actually do inside the cell. The identities of those proteins that bind to GroEL/ES *in vivo* were determined by lysing *E. coli* cells in the presence of glucose and hexokinase to convert cellular ATP to ADP. This removal of ATP traps the encapsulated proteins within the cage (Figure 1). Peptide analysis by liquid chromatography and tandem mass spectrometry identified about 250 proteins. These 250 proteins are nearly all cytosolic, eight being proteins of the periplasm and outer membrane. Each of these 250 proteins was assigned to one of three classes of GroEL substrate, defined by their dependence on GroEL/ES deduced from *in vitro* refolding assays.

Class I substrates, such as enolase and glyceraldehyde-3-phosphate dehydrogenase, have a low tendency to aggregate upon refolding from denaturant, and show only small increases in yield when either GroEL/ES or the small chaperones DnaK and DnaJ are added. Class II substrates, such as glutamate decarboxylase α , have a high tendency to

Figure 1. Mechanism of GroEL/ES action.

The Anfinsen cage folding cycle. A single partly folded polypeptide chain binds to hydrophobic residues exposed at the end of GroEL uncapped by GroES (labelled *trans*). Binding of ATP and GroES to that end triggers release of the chain into the newly created Anfinsen cage (labelled *cis*). The chain has 10–15 seconds to fold inside this cage before ATP hydrolysis allows ATP to bind to the *trans* ring and triggers release of GroES from the *cis* ring [9]. If the chain has not internalised its hydrophobic residues in that time it is likely to be rebound by the same ring [10].



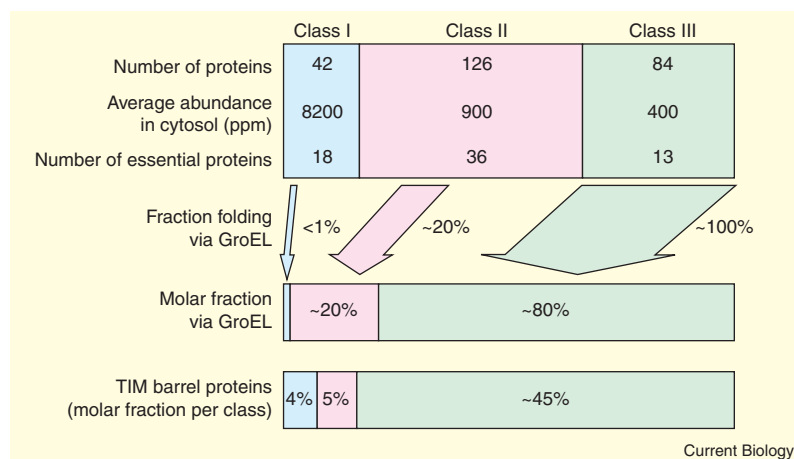


Figure 2. Properties of the three classes of protein that bind to GroEL/ES *in vivo* [1].

aggregate and fail to refold until both GroEL and GroES are added to the refolding buffer, unless the temperature of refolding is lowered from the standard 37°C to 25°C, which allows some spontaneous refolding. But class II substrates will also use the DnaK/J system for refolding, indicating that while encapsulation in the cage occurs with these proteins, it is not essential to prevent them from aggregating. In contrast, class III substrates depend stringently upon the GroEL/ES system to refold at 37°C; this class will not use the DnaK/J system to refold, even though the latter chaperones bind to them. Class III substrates include S-adenosyl methionine synthetase and dihydropicolinate synthase.

Does this *in vitro* classification of GroEL/ES dependence apply to the *in vivo* situation? This point was tested by lowering by genetic means the concentration of GroEL/ES in the cell. Wild-type cells contain about 2000 GroEL oligomers per cell, but remarkably this number can be reduced to about 100 before an effect on cell viability is noted [6]. In the new experiments [1], the GroEL concentration dropped by 90% within 3 hours, but cell growth continued for about 8 hours. Enolase and glutamate carboxylase alpha remained soluble during these 8 hours, but three class III substrates tested either aggregated or were degraded. The next problem was to assign

each of the 250 proteins binding to GroEL/ES *in vivo* to one of the three classes.

Under the growth conditions employed, the doubling time of the cells is 30–40 minutes, while the average half time for GroEL/ES-assisted folding *in vivo* is about 60 seconds [7]. Moreover the concentration of GroEL/ES is about 10% of the concentration of ribosomes. It follows that more than 3% of any given protein should be bound to GroEL/ES if that protein is absolutely dependent upon GroEL/ES for folding. It is possible by labelling the amino acid pool with stable isotopes to use mass spectrometry to measure the fraction of each protein that is bound to GroEL *in vivo* [8]. A total of 84 proteins were assigned to class III by this procedure, and these include 13 proteins essential for growth on glucose and minimal media. Class I and class II proteins are defined as those where less than 0.02%, and between 0.1% and 2.6%, respectively, are bound to GroEL/ES *in vivo*.

The properties of these three classes of GroEL substrate are summarised in Figure 2. Class III substrates are of low to intermediate abundance in the cell but occupy about 80% of the GroEL capacity. Class II substrates are more abundant and occupy the remaining 20% capacity. Thus less than 5% of all the proteins in *E.coli* require the GroEL/ES system absolutely in order to avoid aggregation.

This low value suggests that the *E.coli* proteome has a high degree of folding robustness, presumably because there is extensive functional redundancy among the chaperone complement as a whole. But why do class III substrates depend upon GroEL/ES for folding and how are they selected for binding?

The apical domains of GroEL are known to bind to both extended hydrophobic β -strands and to amphiphilic α -helices but such structures are not noticeably enriched in class III substrates compared to class I and II substrates. A homology-based protein fold assignment of the class III substrates reveals that they are enriched in the $(\beta\alpha)_8$ triosephosphate isomerase (TIM) barrel domain (Figure 2). But the wide distribution of this domain, and the fact that it includes the class I substrate enolase, shows it cannot be the sole criterion for dependence on GroEL/ES. The tendency of class III substrates to aggregate suggest that they expose hydrophobic regions during folding for longer than the other substrates because their folding pathways encounter kinetic traps that take time to be overcome. *In vitro* refolding competition experiments confirm that class III substrates have a higher affinity for GroEL than the other classes [1]. Thus it is likely that class III substrates are characterised by the relative persistence of exposed hydrophobic regions during their folding and not by the structure of the final fold. The satisfying conclusion is that this sophisticated chaperonin machine selects just those properties of its substrates that make them prone to aggregation, supporting the view that this family of molecular chaperone has evolved to reduce aggregation [2].

References

1. Kerner, K.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.-C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 122, 209–220.
2. Ellis, R.J. (2005). Chaperone function: the orthodox view. In *Molecular Chaperones and Cell Signalling*. A.G. Pockley and B.

- Henderson, eds. (Cambridge: Cambridge University Press), pp. 3–21.
3. Ellis, R.J. (1996). Revisiting the Anfinsen cage. *Fold. Des.* 1, R9–15.
 4. Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223–230.
 5. Hartl, F.U., and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295, 1852–1858.
 6. Ivic, A., Olden, D., Wallington, E.J., and Lund, P.A. (1997). Deletion of *Escherichia coli* groEL is complemented by a *Rhizobium leguminosarum* homologue at 37°C but not at 43°C. *Gene* 194, 1–8.
 7. Ewalt, K.L., Hendrick, J.P., Houry, W.A., and Hartl, F.U. (1997). In vivo observation of polypeptide flux through the bacterial chaperonin system. *Cell* 90, 491–500.
 8. Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labelling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376–386.
 9. Farr, G.W., Fenton, W.A., Chaudhuri, T.K., Clare, D.L., Saibil, H.R., and Horwich, A.L. (2003). Folding with and without encapsulation by *cis*- and *trans*-only GroEL–GroES complexes. *EMBO J.* 22, 3220–3230.
 10. Martin, J., and Hartl, F.U. (1997). The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proc. Natl. Acad. Sci. USA* 94, 1107–1112.

Department of Biological Sciences,
University of Warwick, Coventry CV4
7AL, UK.
E-mail: jellis@bio.warwick.ac.uk

DOI: 10.1016/j.cub.2005.08.025

Plant Development: Spacing out Stomatal Pores

How do plants generate the optimal spacing of stomatal pores on their surfaces to prevent excessive water loss, whilst allowing efficient gas exchange? New research into the *ERECTA* family of receptor-like kinases has provided an important link in the cell–cell signalling pathways controlling this process.

Gwyneth C. Ingram

The growth and differentiation of the plant epidermis poses an interesting developmental problem: how do you generate a uniform distribution of specialised cell types, such as stomatal pore complexes, in a sea of cells which are constantly dividing in response to both developmental and environmental stimuli? Studies of stomatal pore complex development have revealed simple rules for ensuring optimum spacing whilst allowing adaptation to different cell division patterns [1,2]. During protodermal proliferation, a subset of cells makes a developmental switch to give stomatal lineages. These so-called meristemoid mother cells undergo asymmetric divisions to give small triangular meristemoids, whilst their neighbours continue to divide symmetrically and generate pavement cells (the basic building blocks of the plant epidermis).

Meristemoids usually undergo a further one to three rounds of asymmetric division, depending on where and when they arose during development. Each division generates a meristemoid and a larger daughter cell with meristemoid mother cell competence — the ability to

divide asymmetrically and generate a secondary stomatal complex. In the mature organ the larger products of asymmetric division differentiate with pavement cell characteristics. In contrast, meristemoids eventually undergo a second developmental switch to give so called guard mother cells which divide once more, this time symmetrically, to give the two guard cells of the stomatal pore (Figure 1A).

During this process, the orientation of asymmetric divisions and the timing of the meristemoid–guard mother cell switch are regulated by pre-existing stomata so that differentiated stomatal pores are almost always separated by at least one pavement-like cell and very rarely touch each other. This so-called one cell rule, and the relatively even spacing of cells taking the meristemoid mother cell pathway earlier in development, indicate that the exchange of positional information from cell to cell plays an important role in organising stomatal distribution.

Three loci have been identified which play critical roles in repressing stomatal development and implementing the one cell rule: *STOMATAL DENSITY AND DISTRIBUTION1* (*SDD1*), which

encodes a putative subtilisin-related extracellular protease [3]; *TOO MANY MOUTHS* (*TMM*), which encodes a leucine rich repeat (LRR) containing receptor-like protein [2,4]; and *YODA* (*YDA*), which encodes a mitogen activated protein (MAP) kinase kinase kinase [5]. Loss-of-function-mutations in these genes cause increased stomatal index and stomatal clusters. From the results of genetic studies, models have been proposed in which a receptor complex composed of *TMM* and an unidentified receptor-like kinase is activated on perception of a ligand either modified or generated by the activity of *SDD1*. *YDA* was proposed to act in a MAP kinase cascade controlled by the activity of the *TMM*–receptor-like kinase complex, although this had not been conclusively demonstrated [5,6] (Figure 1B).

In an intriguing new twist to this tale, Shpak and colleagues [7] have uncovered new roles in stomatal development for the three members of the *ERECTA* class of LRR receptor-like kinase encoding loci: *ERECTA* (*ER*), *ERECTA-LIKE1* (*ERL1*) and *ERECTA-LIKE2* (*ERL2*). Previous studies [8,9] showed that these three genes work together redundantly in promoting proliferative cell divisions in the cortex. Loss-of-function of all three genes leads to a dramatic reduction in plant and organ size, and defects in the differentiation of floral organs. The new study [7] has revealed an additional phenotype for the triple mutant: increased stomatal index and the production of high density stomatal clusters. This is