MINIREVIEW

Mycobacterial chaperonins: the tail wags the dog

Camilo A. Colaco[1] & Alistair MacDougall[2]

[1]ImmunoBiology Limited, Babraham Research Campus, Cambridge, UK; and [2]The European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge, UK

Correspondence: Camilo A. Colaco, ImmunoBiology Limited, Babraham Research Campus, Babraham, Cambridge CB22 3AT, UK. Tel.: +44 1223 496116; fax: +44 1223 496011; e-mail: camilo.colaco@immbio.com

Received 19 June 2013; revised 30 August 2013; accepted 11 September 2013. Final version published online 7 October 2013.

DOI: 10.1111/1574-6968.12276

Editor: Peter Lund

Keywords
mycobacteria; chaperonins; GroEL; cpn60; actinomycetes.

Abstract

Molecular chaperones are defined as proteins that assist the noncovalent assembly of other protein-containing structures *in vivo*, but which are not components of these structures when they are carrying out their normal biological functions. There are numerous families of protein that fit this definition of molecular chaperones, the most ubiquitous of which are the chaperonins and the Hsp70 families, both of which are required for the correct folding of nascent polypeptide chains and thus essential genes for cell viability. The *groE* genes of *Escherichia coli* were the first chaperonin genes to be discovered, within an operon comprising two genes, *groEL* and *groES*, that function together in the correct folding of nascent polypeptide chains. The identification of multiple *groEL* genes in mycobacteria, only one of which is operon-encoded with a *groES* gene, has led to debate about the functions of their encoded proteins, especially as the essential copies are surprisingly often not the operon-encoded genes. Comparisons of these protein sequences reveals a consistent functional homology and identifies an actinomycete-specific chaperonin family, which may chaperone the folding of enzymes involved in mycolic acid synthesis and thus provide a unique target for the development of a new class of broad-spectrum antimycobacterial drugs.

Mycobacteria are aerobic acid-fast bacteria, ubiquitous in the environment, which belong to the phylum *Actinobacteria*. More than 125 mycobacterial species have now been identified, about a third of which are potentially pathogenic to humans. These include pathogens of global importance such as *Mycobacterium tuberculosis* and *Mycobacterium leprae*, as well as a diverse group of nontuberculous mycobacteria (Wilson, 2008). The global burden of TB was estimated by the WHO in 2011 as 8.7 million new cases and an annual mortality of 1.4 million deaths, a third of which are in HIV-positive individuals where the emergence of multidrug-resistant strains is of particular concern (WHO, 2012). Despite significant progress in the control of leprosy, a quarter of a million people are still identified with leprosy yearly (WHO, 2010) and the increasing incidence of *Mycobacterium ulcerans* infections has led to the inclusion of Buruli ulcers as a recognised neglected disease target (First WHO Report on Neglected Tropical Diseases, 2010). The global burden of infectious diseases caused by mycobacteria highlights the importance of developing effective tools for the diagnosis and prevention of mycobacterial infections (Wilson, 2008; First WHO Report on Neglected Tropical Diseases, 2010; WHO, 2012). The application of molecular biological techniques provided a huge step forward in the identification of mycobacterial antigens for use in potential diagnostics and vaccines (Wilson, 2008; First WHO Report on Neglected Tropical Diseases, 2010). One of the first mycobacterial antigens to be identified using these techniques was the major 65-kDa antigen of *M. tuberculosis* (Young *et al.*, 1987), which was initially discovered as an immunodominant antigen in both humoral and cell-mediated immune responses in TB and leprosy (Young *et al.*, 1987, 1988). The subsequent demonstration that the 65-kDa antigen was homologous to the heat shock protein GroEL of *Escherichia coli* led to its common nomenclature as Hsp65 in TB studies (Shinnick *et al.*, 1988; Young *et al.*, 1988) and numerous studies on the protein and encoding gene as potential diagnostics and vaccines (Silva, 1999). However, the demonstration of the function of *E. coli* GroEL as an essential molecular chaperone responsible for the correct folding of key housekeeping genes suggested that Hsp65 is a member of the family of protein chaperonins (Hemmingsen *et al.*, 1988).

FEMS MICROBIOLOGY LETTERS

The chaperonins are a group of molecular chaperones related by homology to the GroEL proteins of *E. coli* (Hemmingsen *et al.*, 1988; Hartl & Hayer-Hartl, 2002). They usually form oligomers of *c.* 800 kDa, made up of two heptameric rings of 60-kDa subunits, each with an apical, an intermediate and an equatorial domain that together enclose a central cavity in which client proteins fold (Hemmingsen *et al.*, 1988; Hartl & Hayer-Hartl, 2002). Client proteins bind to the apical domains and chaperonin function requires a heptameric cochaperonin (GroES in *E. coli*) which binds the same regions of the chaperonin as the client proteins and displaces these into the cavity, where they fold without interacting with other proteins with which they might aggregate (Hartl & Hayer-Hartl, 2002). The chaperonin folding cycle requires binding and hydrolysis of ATP, and networks of allosteric interactions within and between the two rings are needed to complete the cycle (Hartl & Hayer-Hartl, 2002). In *E. coli*, the *groEL* and *groES* genes form part of a single operon and homologous *groEL*/S operons have now been described as essential genes in all phyla and kingdoms; these genes have been ascribed the names *cpn60* and *cpn10* (Coates *et al.*, 1993; Lund, 2001). However, *c.* 30% of bacterial genomes have two or more chaperonin genes raising the question of what purpose the additional copies serve, the most likely possibility being that one copy retains the essential chaperone function, while the others diverge to take on different roles (Lund, 2001).

The *Mycobacteria* were the first bacteria shown to have multiple chaperonins (Kong *et al.*, 1993; Lund, 2001). In *M. tuberculosis* there are two chaperonin genes, one (*cpn60.1*) in an operon with the cochaperonin gene *cpn10* and the other (*cpn60.2*) elsewhere on the chromosome (Kong *et al.*, 1993). The latter encodes Hsp65 and its nomenclature as *cpn60.2* genes reflect its distinct non-operon-encoded genomic localisation. Surprisingly, however, deletion studies in *Mycobacterium smegmatis*, *M. tuberculosis* and *Mycobacterium bovis BCG* have shown that *cpn60.2*, and not *cpn60.1*, encodes the essential chaperonin, despite the latter being operon-encoded with *cpn10* as in *E. coli* (Ojha *et al.*, 2005; Hu *et al.*, 2008; Wang *et al.*, 2011). This has led to some debate about the functional equivalence of the mycobacterial *cpn60* and the *groEL* genes (Lund, 2009). This controversy has not been resolved by the conflicting results obtained from studies on the oligomerisation of recombinant products of the different *cpn60* genes and the crystal structures of their gene products (Qamra & Mande, 2004; Qamra *et al.*, 2004; Lund, 2009). More recently, Lund and colleagues have addressed the questions posed by the presence of multiple Cpn60 proteins and their state of oligomerisation by undertaking a detailed genetic and biophysical characterisation of the chaperonins from *M. tuberculosis* and *M. smegmatis* (Fan *et al.*, 2012). These studies present evidence supporting the evolution of novel function for the *cpn60.1* genes and show that the *cpn60.2*-encoded proteins are highly likely to function as oligomers *in vivo* as they assemble into oligomers in the presence of high salt and nucleotides. They also show that Cpn60.2 from both *M. tuberculosis* and *M. smegmatis* is able to replace GroEL in *E. coli*, when expressed with either the cochaperonin GroES or the cognate cochaperonin Cpn10. However neither Cpn60.1 nor Cpn60.3, a third chaperonin homologue found in *M. smegmatis*, was able to complement GroEL in *E. coli*. These studies also addressed the question of oligomerisation using a number of biophysical techniques and confirmed earlier structural studies showing that, under normal physiological conditions, the purified chaperonins are largely monomers or dimers (Qamra *et al.*, 2004; Fan *et al.*, 2012). However, as monomeric GroEL is nonfunctional (Hartl & Hayer-Hartl, 2002), they examined oligomer formation under a range of conditions and showed oligomerisation in the presence of high concentrations of ammonium salts and either ATP or ADP. Under these conditions, the ATPase activity of the chaperonins increased and the oligomers formed had molecular masses consistent with the typical GroEL tetra-decameric structure of a double ring with seven subunits each. Finally, they showed that substitution of the 22 amino acids at the N-terminus of *cpn60.2* with the equivalent sequences from GroEL led to significant stabilisation of the oligomer, a result consistent with previous studies demonstrating the importance of the N-terminal tail in GroEL assembly (Horovitz *et al.*, 1993). As the N-terminal 60 residues, which include the ATP binding site, are undefined in the crystal structure of dimeric Cpn60.2 (Qamra & Mande, 2004), it is likely that the binding of nucleotide may also assist in stabilising the functional oligomers of this chaperonin *in vivo* (Fan *et al.*, 2012).

These results conclusively demonstrate that the mycobacterial Hsp65, or Cpn60.2, is the structural and functional equivalent of the *E. coli* GroEL and is responsible for the correct folding of essential housekeeping genes as also suggested by the deletional analysis of mutants of *M. smegmatis*, *M. tuberculosis* and *M. bovis BCG* (Ojha *et al.*, 2005; Hu *et al.*, 2008; Wang *et al.*, 2011). This, however, leaves open the intriguing question of the function of the nonessential Cpn60.1, particularly as the recent structural study of *M. tuberculosis* Cpn60.1 suggests that it may indeed act as a conventional chaperonin (Sielaff *et al.*, 2011) in marked contrast to earlier gene deletion studies that proposed a more specialised role in aiding biofilm formation and nonplanktonic growth (Ojha *et al.*, 2005; Hu *et al.*, 2008; Wang *et al.*, 2011). A possible resolution of this apparent paradox is if the two *cpn* genes code for

chaperonins that are involved in the folding of two distinct classes of cellular proteins, with Cpn60.1 chaperoning the folding of a class of nonessential proteins. This possibility is supported by a comparison of the two Cpn60 sequences and, in particular, their more divergent C-terminal domains. The canonical *E. coli groEL* gene encodes a chaperonin with a C-terminal tail rich in glycine and methionine residues that is also seen in the mycobacterial *cpn60.2* sequence, but not in the *cpn60.1* gene, which has a distinct histidine-rich C-terminal tail instead (Fig. 1; Kong *et al.*, 1993; Lund, 2001; Lund, 2009). This sequence difference raises the possibility that the C-terminal tail may be characteristic of the functional equivalents of the *E. coli* GroEL, such as the mycobacterial Cpn60.2, and suggests that the glycine/methionine tail may be used to identify those chaperonins that mediate the folding of essential housekeeping genes. This suggestion is supported by several studies across a number of bacteria that contain multiple chaperonin genes, where deletion studies have revealed that only one of these genes appears to be essential for viability (Lund, 2001, 2009). In all these, the essential *cpn60* genes encode proteins with a glycine- and methionine-rich C-terminal tail (Fig. 1 and C. Colaco, unpublished data). Moreover, it should also be noted that the third Cpn60 sequence found in some mycobacteria, such as *M. smegmatis*, has a distinct C-terminal tail that is neither glycine-/methionine-rich nor histidine-rich (C. Colaco, unpublished data).

This raises the more enigmatic question of the cellular functions of the client proteins chaperoned by Cpn60.1 and the possible significance of the histidine-rich C-terminal tail in selecting these polypeptide substrates. In GroEL, the C-terminal tail is highly flexible and thus undefined in the crystal structures (Hartl & Hayer-Hartl, 2002; Machida *et al.*, 2008). However, a detailed genetic analysis of the final 23 residues assessing the ability of C-terminal-truncated, double- and single-ring mutants to assist the refolding of rhodanese and malate dehydrogenase showed that this domain defines the environment within the central cavity and in particular its hydropathicity, features that would impact on both the size and nature of the substrate protein folded by the chaperonin (Tang *et al.*, 2006; Machida *et al.*, 2008). This is consistent with a role for the mycobacterial Cpn60.1 chaperonins in the folding of a distinct class of proteins, possibly unique to mycobacteria or actinomyces. Although a distinct DNA-bound function in the assembly of the nucleoid has recently been proposed for Cpn60.1 (Basu *et al.*, 2009) this is unlikely to involve the C-terminal tail sequence, as the mitochondrial Hsp60 chaperonin for

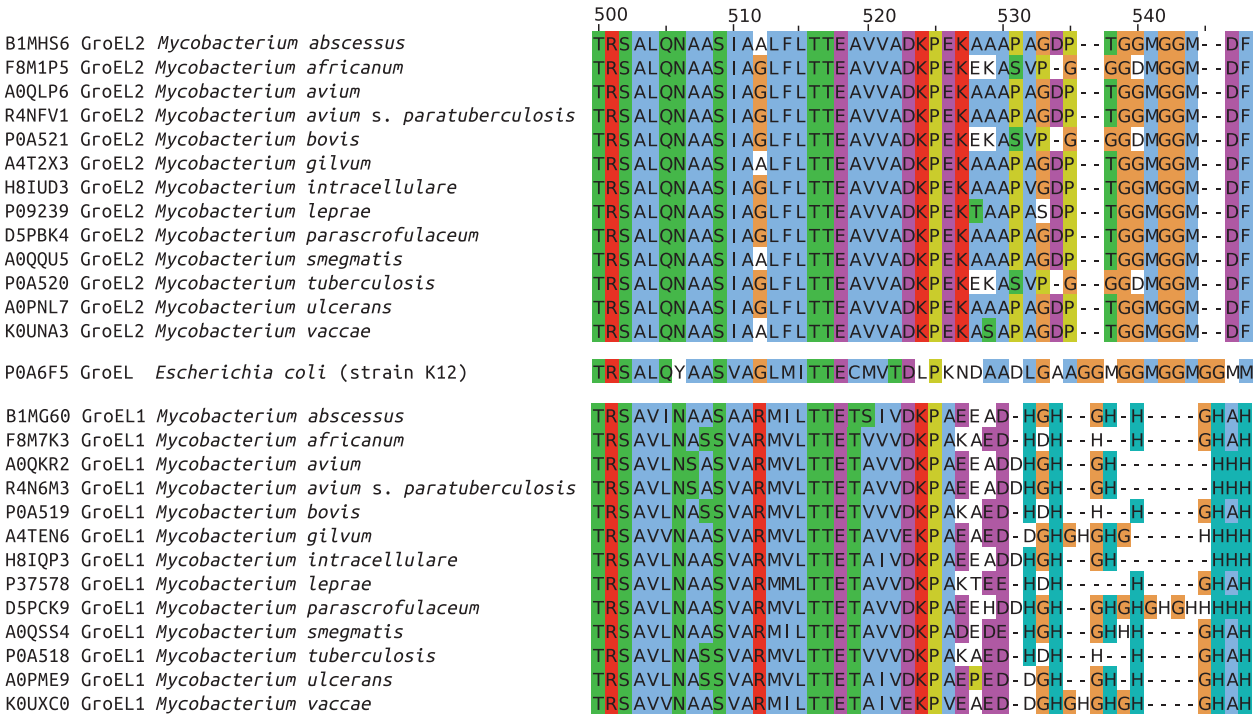

Fig. 1. C-terminal sequence alignment of mycobacterial GroELs with the *Escherichia coli* homologue reveals two protein families, one of which shares a glycine-rich domain with the *E. coli* GroEL and the other showing a distinct histidine-rich sequence. Sequences were retrieved from UniProtKB and aligned using T-Coffee within the Jalview alignment viewer. The amino acids are marked up with the default ClustalX colour scheme.

which nucleotide binding has also been reported does not have a histidine-rich C-terminal tail (Kaufman *et al.*, 2003; Basu *et al.*, 2009). A database search with the histidine-rich C-terminal sequence of Cpn60.1 reveals highly homologous proteins across all mycobacterial species, as well as *Corynebacteria*, *Nocardia* and *Rhodococcus* (C. Colaco, unpublished data). A common feature of all these *Actinobacteria* is their synthesis of a complex cell wall containing mycolic acid derivatives, and this suggests the intriguing possibility that the biological role of the mycobacterial Cpn60.1 may be to chaperone the folding of key enzymes involved in the synthesis of mycolic acid. Such a role for Cpn60.1 is also consistent with the defects in mycolates and biofilm formation observed in the *cpn60.1* knockouts in *M. smegmatis*, where the protein was also found to be associated with *KasA* and SMEG4308, both key enzymes implicated in biofilm formation and involved in fatty acid synthesis (Tang *et al.*, 2006; Kumar *et al.*, 2009). In this respect, it is interesting to note that the oligomerisation of Cpn60.1 has been shown to be facilitated by phosphorylation (Canova *et al.*, 2009), which is thought to be mediated by the serine threonine protein kinases that have also been implicated in biofilm formation (Gopalaswamy *et al.*, 2008).

Finally, as KasA has been identified as an important drug target for the development of new drugs against TB (Brown *et al.*, 2009), the most interesting implication of the suggested role of Cpn60.1 is that this novel mycobacterial chaperonin may present an upstream target for drug development. Thus, therapeutics that target Cpn60.1 would both inhibit other enzymes involved in the synthesis of mycolates and potentially also be effective in the treatment of all pathogenic mycobacterial strains including not just *M. tuberculosis*, but also *M. leprae* and *M. ulcerans*.

References

Basu D, Khare G, Singh S, Tyagi A, Khosla S & Mande SC (2009) A novel nucleoid-associated protein of *Mycobacterium tuberculosis* is a sequence homolog of GroEL. *Nucleic Acids Res* **37**: 4944–4954.

Brown AK, Taylor RC, Bhatt A, Futterer K & Besra GS (2009) Platensimycin activity against mycobacterial beta-ketoacyl-ACP synthases. *PLoS ONE* **4**: e6306.

Canova MJ, Kremer L & Molle V (2009) The *Mycobacterium tuberculosis* GroEL1 chaperone is a substrate of Ser/Thr protein kinases. *J Bacteriol* **191**: 2876–2883.

Coates AR, Shinnick TM & Ellis RJ (1993) Chaperonin nomenclature. *Mol Microbiol* **8**: 787.

Fan M, Rao T, Zacco E, Ahmed MT, Shukla A, Ojha A, Freeke J, Robinson CV, Benesch JL & Lund PA (2012) The unusual mycobacterial chaperonins: evidence for *in vivo* oligomerization and specialization of function. *Mol Microbiol* **85**: 934–944.

First WHO Report on Neglected Tropical Diseases (2010) Available at: whqlibdoc.who.int/publications/2010/9789241564090_eng.pdf.

Gopalaswamy R, Narayanan S, Jacobs WR Jr & Av-Gay Y (2008) *Mycobacterium smegmatis* biofilm formation and sliding motility are affected by the serine/threonine protein kinase PknF. *FEMS Microbiol Lett* **278**: 121–127.

Hartl FU & Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**: 1852–1858.

Hemmingsen SM, Woolford CS, van der Vies M, Tilly K, Dennis DT, Georgopoulos CP, Hendrix RW & Ellis RJ (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**: 330–334.

Horovitz A, Bochkareva ES & Girshovich AS (1993) The N-terminus of the molecular chaperonin GroEL is a crucial structural element for its assembly. *J Biol Chem* **268**: 9957–9959.

Hu Y, Henderson B, Lund PA, Tormay P, Ahmed MT, Gurcha SS, Besra GS & Coates AR (2008) A *Mycobacterium tuberculosis* mutant lacking the groEL homologue cpn60.1 is viable but fails to induce an inflammatory response in animal models of infection. *Infect Immun* **76**: 1535–1546.

Kaufman BA, Kolesar JE, Perlman PS & Butow RA (2003) A function for the mitochondrial chaperonin Hsp60 in the structure and transmission of mitochondrial DNA nucleoids in *Saccharomyces cerevisiae*. *J Cell Biol* **163**: 457–461.

Kong TH, Coates AR, Butcher PD, Hickman CJ & Shinnick TM (1993) *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *P Natl Acad Sci USA* **90**: 2608–2612.

Kumar CM, Khare G, Srikanth CV, Tyagi AK, Sardesai AA & Mande SC (2009) Facilitated oligomerization of mycobacterial GroEL: evidence for phosphorylation-mediated oligomerization. *J Bacteriol* **191**: 6525–6538.

Lund PA (2001) Microbial molecular chaperones. *Adv Microb Physiol* **44**: 93–140.

Lund PA (2009) Multiple chaperonins in bacteria-why so many? *FEMS Microbiol Rev* **33**: 785–800.

Machida K, Kono-Okada A, Hongo K, Mizobata T & Kawata Y (2008) Hydrophilic residues 526 KNDAAD 531 in the flexible C-terminal region of the chaperonin GroEL are critical for substrate protein folding within the central cavity. *J Biol Chem* **283**: 6886–6896.

Ojha A, Anand M, Bhatt A, Kremer L, Jacobs WR & Hatfull GF (2005) GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell* **123**: 861–873.

Qamra R & Mande SC (2004) Crystal structure of the 65-kilodalton heat shock protein, chaperonin 60.2, of *Mycobacterium tuberculosis*. *J Bacteriol* **186**: 8105–8113.

Qamra R, Srinivas V & Mande SC (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.

Shinnick TM, Vodkin MH & Williams JC (1988) The *Mycobacterium tuberculosis* 65-kilodalton antigen is a heat shock protein which corresponds to common antigen and to the *Escherichia coli* GroEL protein. *Infect Immun* **56**: 446–451.

Sielaff B, Lee KS & Tsai FTF (2011) Structural and functional conservation of *Mycobacterium tuberculosis* GroEL paralogs suggests that GroEL1 is a chaperonin. *J Mol Biol* **405**: 831–839.

Silva CL (1999) The potential use of heat-shock proteins to vaccinate against mycobacterial infections. *Microbes Infect* **1**: 429–435.

Tang YC, Chang HC, Roeben A, Wischnewski D, Wischnewski N, Kerner MJ, Hartl FU & Hayer-Hartl M (2006) Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* **125**: 903–914.

Wang XM, Lu C, Soetaert K *et al.* (2011) Biochemical and immunological characterization of a cpn60.1 knockout mutant of *Mycobacterium bovis BCG*. *Microbiology* **157**: 1205–1219.

WHO (2010) Global leprosy situation. *Wkly Epidemiol Rec* **85**: 337–348.

WHO (2012) Global tuberculosis report. Available at: http://apps.who.int/iris/bitstream/10665/75938/1/9789241564502_eng.pdf.

Wilson ML (2008) Reducing the global burden of mycobacterial infections: one more piece of the puzzle. *Am J Clin Pathol* **130**: 849–852.

Young DB, Ivanyi J, Cox JH & Lamb JR (1987) The 65 kDa antigen of mycobacteria – a common bacterial protein? *Immunol Today* **8**: 215–219.

Young D, Lathigra R, Hendrix R, Sweetser D & Young RA (1988) Stress proteins are immune targets in leprosy and tuberculosis. *P Natl Acad Sci USA* **85**: 4267–4270.