Update TRENDS in Microbiology Vol.15 No.8

Research Focus

Actinomycete-like proteasomes in a Gram-negative bacterium

René De Mot

Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, Faculty of Bioscience Engineering, Katholieke Universiteit Leuven, B-3001 Heverlee-Leuven, Belgium

Cultivation-independent proteogenomic exploration of mine-drainage biofilm has revealed proteasomes in Gram-negative bacteria of the *Nitrospirae* phylum (*Leptospirillum* group II) dominating this acidophilic community. Most probably, the proteasome genes were acquired from actinobacteria, the only eubacteria previously known to contain proteasomes. In addition, this study shows that the proteasome and the evolutionarily related ATP-dependent protease HslVU (also known as ClpQY) are not mutually exclusive in prokaryotes.

Proteogenomic mining digs up new proteasomes

The steadily increasing number of genome sequences determined for microorganisms isolated from diverse environments (http://www.genomesonline.org) represents only a small fraction of the huge genetic diversity present in the microbial world. The metagenomic approach provides access to DNA sequences from the vast number of organisms that are not (yet) cultivated. When applied to a simple community, large-scale environmental shotgun-sequencing has the potential to reconstruct near-complete genomes, as shown for a biofilm thriving on sulfuric acid-containing, iron-rich mine drainage conduits [1]. Relying on such metagenomic data, mass-spectrometry-based proteomics enabled exploration of *in situ* production of predicted proteins from the uncultivated *Leptospirillum* group II bacteria (*Nitrospirae* phylum) dominating this biofilm [2,3]. Here, several *Leptospirillum* proteins detected in this harsh environment are identified as proteasome components. As this represents the first evidence for the occurrence of this proteolytic system in a Gram-negative bacterium, its features are compared with those of proteasomes from other prokaryotes.

Prokaryotic proteasomes

Tightly controlled intracellular proteolysis by the ubiquitin–proteasome system has a key role in the physiology of eukaryotic cells. Energy dependence of this process stems from ubiquitin tagging of target proteins and from the activity of the 26S proteolytic machinery, involving ATP-dependent unfolding of substrates by the 19S complex before their degradation inside the associated 20S nanocompartment [4]. Although ubiquitin is absent from prokaryotes, archaea encode 20S proteasomes that can function together with a hexameric ATPase complex reminiscent of the base of the 19S complex, containing six ATPases of the AAA family [5]. In the *Bacteria* kingdom, 20S proteasomes have been biochemically characterized for species of the genera *Rhodococcus* [6], *Streptomyces* [7], *Frankia* [8] and *Mycobacterium* [9], all of which are members of the order *Actinomycetales* (high-GC Gram-positive bacteria; class *Actinobacteria*). The barrel-like structure with four heptameric rings assembled from two types of subunits (*prcA*-encoded α subunit and *prcB*-encoded β subunit) has been resolved for the *Rhodococcus* [10] and *Mycobacterium* complexes [11]. The hexameric ATPase ARC (AAA ATPase forming ring-shaped complexes) [12], designated Mpa (*Mycobacterium* proteasomal ATPase) in *Mycobacterium* [13], is thought to act together with the 20S proteasome in actinomycetes, but direct evidence for this is lacking. Mutants devoid of Mpa or another putative component of the mycobacterial proteasome system, PafA (proteasome accessory factor), are hyper-susceptible to acidified nitrite and severely attenuated in a mouse model of infection [13–15].

Proteasomal gene organization in actinobacteria

Inspection of finished actinobacterial genomes representing 47 different species reveals that the genes encoding Mpa or ARC and PafA homologues, if present, are typically found upstream (Mpa or ARC) and downstream (PafA) of the 20S proteasome structural operon *prcBA*. A *pafA*-paralogous gene (tentatively designated *pafX*) and a small conserved open reading frame (*orfX*) are located between *mpa* and *prcBA* (Figure 1). This 'core' cluster of convergent genes (*mpa-pafX-orfX-prcB-prcA-pafA*) is present in *Arthrobacter* species. In other actinomycetes, this region carries additional, mostly unrelated genes of unknown function downstream of *mpa* or *prcA*. Remarkably, *prcB-prcA* genes are absent from the core cluster (and the genome) in *Actinomyces odontolyticus* and in species of *Corynebacterium* and *Bifidobacterium*. It has therefore been proposed that Mpa-like proteins might also exert proteasome-independent functions [16]. It cannot be excluded that this deviant gene organization reflects reductive evolution. Conversely, the 20S proteasome in *Propionibacterium acnes* seems to operate without the accessory proteins found in most actinomycetes.

Proteasomes in *Leptospirillum*

The current view of eubacterial proteasomes being confined to actinobacteria is challenged by the results of

Corresponding author: De Mot, R. (rene.demot@biw.kuleuven.be).
Available online 22 June 2007.

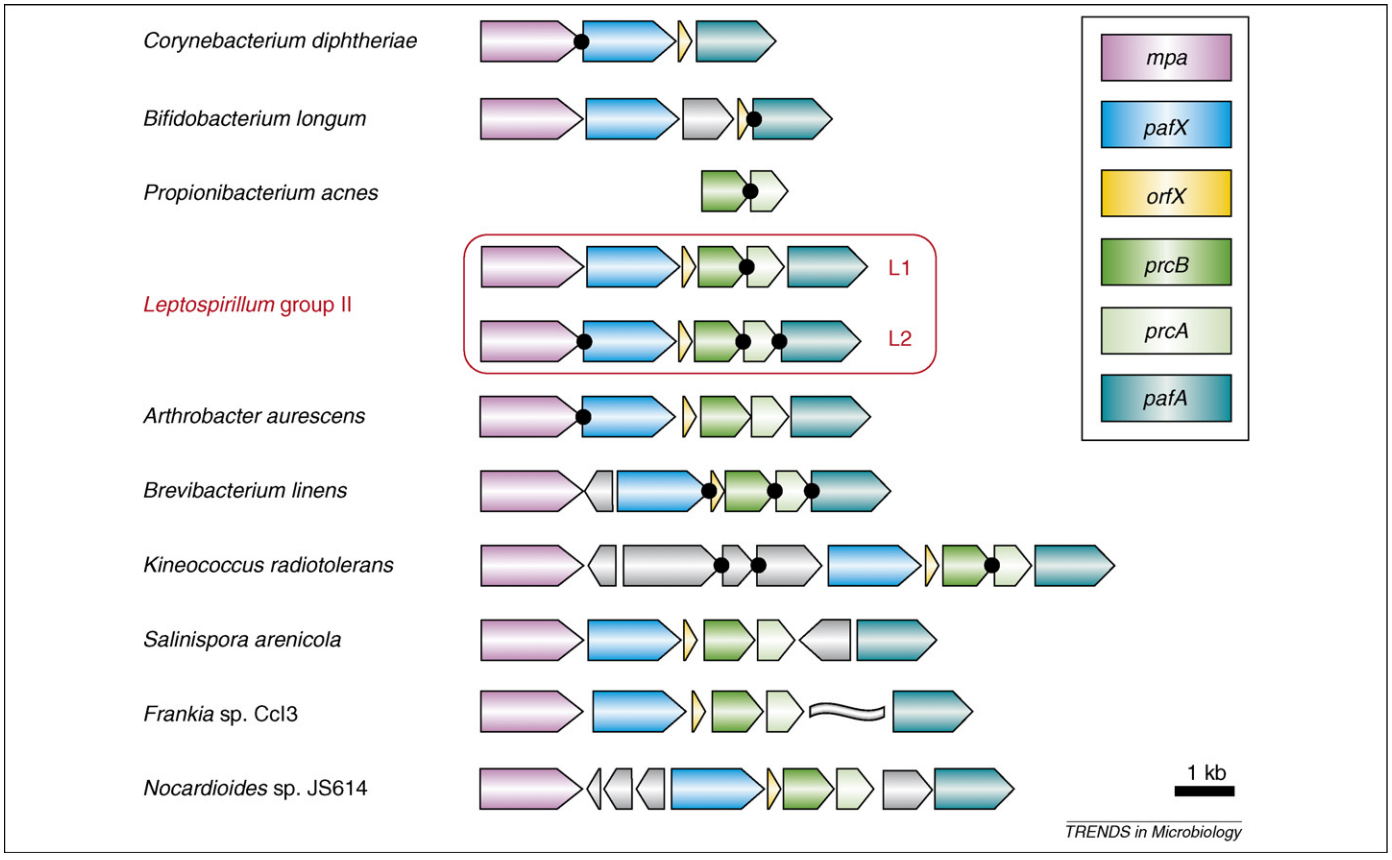

Figure 1. Comparative organization of (putative) proteasome genes in *Leptospirillum* group II composite genomes [3] and representative actinobacteria. The L1 cluster (59.8% G+C) represents genes 76→71 from scaffold UBA_8241 and the orthologues 49→44 from scaffold 5way_CG_21 (96–99% amino acid identity). The scaffold UBA_8062 genes 10→5 and the 5way_CG_21 orthologues 13→18 (88–97% amino acid sequence identity) are designated cluster L2 (53.7% G+C). Pair-wise amino acid sequence identities between predicted gene products from L1 and L2 vary between 26–34%. For the best matches with actinomycete genes, amino acid sequence identities are in the range of 40–55% for L1 and 26–39% for L2 polypeptides, respectively. No other actinomycete-like genes are present in the respective flanking regions. The L1-containing scaffolds also carry an unlinked *hslVU* operon: UBA_8241 genes 165–164 and 5way_CG_21 genes 92–93 (100% amino acid identity; not shown). Similar colours represent apparent orthologues in different species. Other genes are shown in grey. Partially overlapping reading frames, suggesting translational coupling, are marked with black circles. The unlinked genes in *Frankia* are separated by a wavy line (~15 kb). The gene cluster *mpa-pafX-orfX-prcB-prcA-pafA* and variants thereof are found in most *Actinomycetales* genera (including species of *Acidothermus, Janibacter, Mycobacterium, Nocardia, Rhodococcus, Saccharopolyspora, Streptomyces,* and *Thermobifida*; not shown), except in *Corynebacterium* species and *Actinomyces odontolyticus* (lacking the *prcB-prcA* genes), *Propionibacterium acnes* (carrying only the *prcBA* operon), and some species devoid of all these genes (*Tropheryma whipplei, Clavibacter michiganensis, Leifsonia xyli*). The proteasome system is also absent from *Collinsella aerofaciens* (order *Coriobacteriales*), *Rubrobacter xylanophilus* (order *Rubrobacterales*), and an unclassified marine actinobacterium (strain PHSC20C1). *Bifidobacterium* species (order *Bifidobacteriales*) have a gene organization similar to *Actinomyces* and *Corynebacterium*. The *Leptospirillum* group II genomes do not contain homologues of the genes (*pafB, pafC*) that constitute an operon with *pafA* in *M. tuberculosis* [25], and probably also in other mycobacteria and the phylogenetically related rhodococci and corynebacteria. Finished actinobacterial genomes were searched at NCBI-Microbial genomes (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) and the Sanger Institute (http://www.sanger.ac.uk/).

Banfield and co-workers [2,3]. In an acid mine-drainage biofilm sampled at the Richmond Mine (Iron Mountain, CA, USA), peptides were identified by tandem mass spectrometry that could be assigned to specific proteins of the major biofilm organism *Leptospirillum* group II. Protein identification relied on the proteomes predicted from two near-complete genomic datasets assembled from sequence data for communities sampled at two different locations (UBA, 5-way CG). These near-clonal composite genomes are largely syntenous with orthologues sharing, on average, ~95% identity [3]. The acidophilic biofilm contains proteins derived from two distinct *Arthrobacter*-like proteasome gene clusters (Figure 1). All proteins except one (PafA) from the L1 cluster were identified for the CG type, along with one of its UBA-type proteins (PafX). Conversely, no CG-type proteins, but all UBA-type proteins (except OrfX), were found for the L2 cluster.

Two cases of inter-phylum gene transfer?

In addition to the remarkably conserved gene organization and sequence homology (Figure 1), a specific feature of the β-subunit precursor proteins strongly points to a probable actinobacterial rather than archaeal origin of the *Leptospirillum* proteasomes. Proteasomal β-type-subunit precursors are processed by autolytic removal of their propeptide to expose the N-terminal active-site residue threonine. Compared with the short archaeal β-subunit propeptides, the propeptides in actinomycetes are considerably longer and promote assembly by interaction with α subunits [10]. This feature seems to be retained in the *Leptospirillum* PrcB proteins, including the consensus cleavage site (G/A)T with the catalytic threonine and the crucial SF sequence ('central box') conserved in all actinomycetal PrcB proteins (data not shown). Probably, a similar 20S proteasome assembly pathway is active in *Leptospirillum*.

Both *Leptospirillum* ATPases contain the 'second region of homology' present in proteasomal ATPases [5] but they differ at their C termini. Only the Mpa or ARC homologue from cluster L1 contains the extended C terminus ('box II') present in most actinomycetal homologues [16]. This extension is absent from the corresponding ATPases in actinomycetes that lack the *prcB-prcA* genes, namely *Bifidobacterium* and *Corynebacterium* species [16], and *A. odontolyticus*. The invariant penultimate tyrosine residue from box II is essential for at least part of the proteasome-related functions of Mpa in *Mycobacterium tuberculosis* [17]. This indicates that the ATPases from the two *Leptospirillum* clusters might exert dedicated functions.

Although a proteasome-related function of PafA and PafX is inferred from phenotypic characterization of the corresponding mutants [17,18], their biochemical activity is unknown. PafA and the paralogous PafX constitute a family of proteins composed of the DUF275–DUF245 domains (http://www.sanger.ac.uk/Software/Pfam) that were not detected outside actinomycetes until their identification in *Leptospirillum* group II organisms. Likewise, the small OrfX (of the DUF797 family) was previously considered to be an actinomycete-specific protein.

These similarities, the rather low sequence identity between proteins from both *Leptospirillum* clusters and the difference in G+C content (~6%) indicate that these proteasome gene sets were acquired independently by horizontal gene-transfer events from actinobacteria rather than having evolved following gene duplication. Actinobacteria of the order *Acidimicrobiales* present in the *Leptospirillum*-dominated niche [19] are potential sources but genome sequence data and, hence, information about proteasome content is lacking for these organisms. Although most actinomycete genomes encode only one set of proteasome genes, the *Streptomyces avermitilis* and *Salinispora tropica* genomes carry a second unlinked *prcB* gene, and *Rhodococcus erythropolis* NI86/21 co-expresses two proteasome operons [6], which is reminiscent of the *Leptospirillum* proteasome profile. Such expanded proteasome gene content, which is also obvious for some archaebacteria, probably serves to diversify functional capacity and increase regulatory versatility of proteasome-mediated proteolysis in prokaryotes.

Evolutionary aspect

The presence of other *Leptospirillum* ATP-dependent proteolytic activities (Lon, FtsH) can be inferred from the peptides detected in the biofilm community. The scaffolds with the L1 cluster also carry an *hslVU* operon. Although Banfield and co-workers [3] did not detect this using community proteomics, expression of this operon, together with *lon*, was demonstrated by metatranscriptomics of a similar environment (Tinto River site, Spain) using a *Leptospirillum ferrooxidans* microarray [20]. Until now, the ATP-dependent protease system HslV(ClpQ protease)-HslU(ClpY ATPase), although widely distributed in bacteria, was considered to be absent from prokaryotes containing proteasomes [21]. Notably, the similar folding of HslV and proteasome subunits is indicative of an evolutionary relationship between these compartmentalized proteases, which both belong to the N-terminal nucleophile (Ntn) hydrolases [22]. The hypothesis that the more-complex 20S proteasome barrel evolved from the 'simpler' dodecameric HslV has been advocated as a major criterion for rooting the tree of life [23].

Concluding remarks and future perspectives

The study by Banfield and co-workers [3] illustrates the potential of integrated cultivation-independent approaches for exploring poorly characterized microbial niches. High-resolution proteomics of an extreme environment previously characterized by metagenomics has provided proof of extensive recombinational genetic exchange between genomes of the dominant *Leptospirillum* population. Because actinomycete proteasomes have been implicated in defence against environmental stress [18,24], the presence of two distantly related proteasome gene clusters of probable actinomycete origin in *Leptospirillum* probably reflects the contribution of horizontal gene transfers to optimized adaptation in a niche that triggers high-level expression of stress-responsive genes [2]. It seems that the large phylogenetic distance between *Nitrospirae* and actinomycetes did not constitute a barrier for acquisition of a multi-subunit molecular machine such as the proteasome. The autocatalytic nature of the 20S proteasome assembly pathway, which enables production of functional 28-membered complexes upon co-expression of actinomycete-subunit genes in a Gram-negative host (*Escherichia coli*), might have facilitated such horizontal transfer. This property might enable biochemical characterization of recombinant *Leptospirillum* proteasomes as a first step in elucidating the physiological role of the proteasome in these extremophilic Gram-negative bacteria (Box 1). If confirmed to contain proteasome genes, some of the *Leptospirillum* group II biofilm isolates characterized by multi-locus sequence typing [3] might serve as sources of native 20S proteasomes and accessory proteins. Combined with data from ongoing genomic sequencing of *Nitrospirae* members, this will shed more light on the distribution of the proteasome system in a poorly characterized phylum of Gram-negative bacteria involved in important environmental processes such as nitrite oxidation (*Nitrospira*), thermophilic sulfate reduction (*Thermodesulfovibrio*) and acidophilic iron oxidation (*Leptospirillum*), and add relevant insight into the evolution of the proteasome in general.

Box 1. Unanswered questions about *Leptospirillum* proteasomes

- Are these novel proteasomes genes from as yet uncultivated bacteria amenable to recombinant expression for characterization of their catalytic activity, structure and assembly?
- Which actinobacterial species are the most likely source(s) of the proteasome genes, and are these organisms colonizing the same niche?
- Does acquisition of a proteasome system provide a competitive advantage in the extreme environment dominated by *Leptospirillum*?
- How widespread are proteasomes among Leptospirilla and other (extremophilic) *Nitrospirae* members?
- If effectively present in a single *Leptospirillum* strain, do both proteasome systems exert dedicated functions or do they display some level of functional redundancy? How is their activity coordinated with proteolysis by the proteasome-like HslVU?

Acknowledgement
This work was supported by a grant (Onderzoeksproject G.0192.02N) from the Fund for Scientific Research (FWO-Vlaanderen).

References

1 Tyson, G.W. *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37–43
2 Ram, R.J. *et al.* (2005) Community proteomics of a natural microbial biofilm. *Science* 308, 1915–1920
3 Lo, I. *et al.* (2007) Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. *Nature* 446, 537–541
4 Voges, D. *et al.* (1999) The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* 68, 1015–1068
5 Smith, D.M. *et al.* (2006) Proteasomes and their associated ATPases: a destructive combination. *J. Struct. Biol.* 156, 72–83
6 Tamura, T. *et al.* (1995) The first characterization of a eubacterial proteasome: the 20S complex of *Rhodococcus*. *Curr. Biol.* 5, 766–774
7 Nagy, I. *et al.* (1998) The 20S proteasome of *Streptomyces coelicolor*. *J. Bacteriol.* 180, 5448–5453
8 Pouch, M.N. *et al.* (2000) Characterization of the 20S proteasome from the actinomycete *Frankia*. *Mol. Microbiol.* 35, 368–377
9 Lin, G. *et al.* (2006) *Mycobacterium tuberculosis prcBA* genes encode a gated proteasome with broad oligopeptide specificity. *Mol. Microbiol.* 59, 1405–1416
10 Kwon, Y.D. *et al.* (2004) Crystal structures of the *Rhodococcus* proteasome with and without its pro-peptides: implications for the role of the pro-peptide in proteasome assembly. *J. Mol. Biol.* 335, 233–245
11 Hu, G. *et al.* (2006) Structure of the *Mycobacterium tuberculosis* proteasome and mechanism of inhibition by a peptidyl boronate. *Mol. Microbiol.* 59, 1417–1428
12 Wolf, S. *et al.* (1998) Characterization of ARC, a divergent member of the AAA ATPase family from *Rhodococcus erythropolis*. *J. Mol. Biol.* 277, 13–25
13 Darwin, K.H. *et al.* (2005) Characterization of a *Mycobacterium tuberculosis* proteasomal ATPase homologue. *Mol. Microbiol.* 55, 561–571
14 Darwin, K.H. *et al.* (2003) The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science* 302, 1963–1966
15 Lamichhane, G. *et al.* (2006) Deletion of a *Mycobacterium tuberculosis* proteasomal ATPase homologue gene produces a slow-growing strain that persists in host tissues. *J. Infect. Dis.* 194, 1233–1240
16 Zhang, X. *et al.* (2004) The N-terminal coiled coil of the *Rhodococcus erythropolis* ARC AAA ATPase is neither necessary for oligomerization nor nucleotide hydrolysis. *J. Struct. Biol.* 146, 155–165
17 Pearce, M.J. *et al.* (2006) Identification of substrates of the *Mycobacterium tuberculosis* proteasome. *EMBO J.* 25, 5423–5432
18 De Mot, R. *et al.* (2007) Proteome analysis of *Streptomyces coelicolor* mutants affected in the proteasome system reveals changes in stress-responsive proteins. *Arch. Microbiol.* DOI: 10.1007/s00203-007-0243-8
19 Bond, P.L. *et al.* (2000) Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl. Environ. Microbiol.* 66, 3842–3849
20 Parro, V. *et al.* (2007) Analysis of environmental transcriptomes by DNA microarrays. *Environ. Microbiol.* 9, 453–464
21 Gille, C. *et al.* (2003) A comprehensive view on proteasomal sequences: implications for the evolution of the proteasome. *J. Mol. Biol.* 326, 1437–1448
22 Volker, C. and Lupas, A.N. (2002) Molecular evolution of proteasomes. *Curr. Top. Microbiol. Immunol.* 268, 1–22
23 Cavalier-Smith, T. (2006) Rooting the tree of life by transition analyses. *Biol. Direct* 1, 19
24 Butler, S.M. *et al.* (2006) Self-compartmentalized bacterial proteases and pathogenesis. *Mol. Microbiol.* 60, 553–562
25 Festa, R.A. *et al.* (2007) Characterization of the proteasome accessory factor (*paf*) operon in *Mycobacterium tuberculosis*. *J. Bacteriol.* 189, 3044–3050

0966-842X/$ – see front matter
doi:10.1016/j.tim.2007.06.002

Letters

Protozoa as an environmental reservoir of bovine tuberculosis

Shelley G. Rhodes[1], Frans A.A.M. De Leij[2] and Jeremy W. Dale[2]

[1] TB Research Group, Veterinary Laboratories Agency, Surrey, KT15 3NB, UK
[2] Microbial Sciences Group, School of Biomedical and Molecular Sciences, University of Surrey, Surrey, GU2 7XH, UK

An issue for the epidemiology of bovine tuberculosis?

The route of infection of cattle with *Mycobacterium bovis* is not entirely understood. It is suspected that climate has some role in the epidemiology of bovine tuberculosis in Britain [1,2], and local climatic features that enable prolonged survival of *M. bovis* on pasture could be important in determining the survival and subsequent transmission of *M. bovis* to cattle. Soil microorganisms form part of the local environment for *M. bovis*, particularly at badger and cattle watering hole or latrine areas where the organic content of soil is high and, interestingly, where *M. bovis* has recently been detected [3]. The potential importance of soil microorganisms to *M. bovis* ecology has not thus far been addressed, and studies of mycobacterial survival in the environment have been limited. Whether free-living soil protozoa impact upon *M. bovis* is yet to be determined and forms part of an ongoing study.

Consider the protozoa

Protozoa are unicellular organisms that represent one of the oldest forms of animal life. They have become adapted to almost all types of environment, the availability of food and water being the most important factors governing their prevalence in different micro-environments. Few protozoa can synthesize their food from inorganic materials and they therefore depend upon available organic substances such as disintegrating plant or animal material or, relevant here, living microorganisms. Hence, soils that contain high levels of organic matter and bacteria will support an abundance of protozoa. The survival of protozoa under fluctuating environmental conditions is enhanced by the ability to form cysts. Amoebic cysts are resistant to many different environmental stresses such as disinfectants and desiccation.

Corresponding author: De Mot, R. (rene.demot@biw.kuleuven.be).
Available online 18 June 2007.