Genetics and Molecular Biology, 30, 1, 225-229 (2007)

www.sbg.org.br

Research Article

Comparative genome analysis of proteases, oligopeptide uptake and secretion systems in *Mycoplasma* spp

Charley Christian Staats, Juliano Boldo, Leonardo Broetto, Marilene Vainstein and Augusto Schrank

Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

Abstract

Mycoplasmas are very fastidious in their nutritional requirements for *in vitro* growth and have limited biosynthetic capacity, a reflection of their reduced genomes. As a result, these bacteria depend upon external metabolites for nutrition and growth and have developed dependence on their hosts for survival and maintenance. Protein degradation and peptide importation play an important role in *Mycoplasma* spp. nutrition, and proteases can play a role in host adaptation and pathogenicity. Here, we present a general survey on the genes involved in protein degradation, secretion and importation, comparing all available Mollicute genomes.

Key words: Mycoplasma, proteases, minimal genomes, secretion systems.

Received: April 4, 2006; Accepted: October 5, 2006.

Introduction

Mycoplasmas are considered the smallest cells capable of propagation in cell-free medium, and some species are pathogenic to humans, animals and plants. In spite of their reduced genomes, due to significant gene loss through evolution/adaptation, mycoplasmas are a very successful group of organisms, as judged by their large number of species and habitats (Razin *et al.*, 1998). The `minimum cell' life style of mycoplasmas became possible by the adoption of a parasitic mode of life, exploiting nutrients not synthesized by themselves, and evolving systems to invade and to persist in their hosts (van Ham *et al*., 2003).

Over the last few years, the genomes of nine Mycoplasma species were sequenced, reinforcing comparative genome studies that allow a better understanding of their metabolism and the relations with their hosts. Mycoplasmas evolved from gram-positive bacteria and, through evolution, lost the cell wall and many metabolic pathways for the synthesis of macromolecule building blocks. Mycoplasmas possess no complete routes for amino acids synthesis and degradation, implying that these monomers must be acquired either from their hosts or from a culture medium, depending upon membrane transporters (Vasconcelos *et al*., 2005). Exogenous peptides are an important source of amino acids. Indeed, bacteria have evolved peptide transport systems that also assist in responses to environmental changes, mediating functions such as quorum sensing, sporulation, pheromone transport, and chemotaxis (Wang *et al*., 2004).

Despite the presence of a complete set of genes responsible for essential cell activities such as replication, transcription and translation, genes involved in post-translational protein modifications are not readily disclosed by the annotation of the mycoplasma genomes. Some of these processes, such as protein maturation and localization, are intrinsically dependent on proteases. Microbial proteases may also play important roles in pathogenicity and nutrition.

Bacterial development is also dependent on the secretion of proteins with a plethora of functions. One of the major transport routes, the so-called Sec pathway, is conserved in all domains of life and is the only system found in mycoplasmas by genome surveys (Stephenson, 2005).

In this work, we present a general survey on the genes involved in protein metabolism, based on the available mycoplasma genomes.

Material and Methods

The complete genome sequences of the *Mycoplasma* spp. used in this work were retrieved from the NCBI data base (http://www.ncbi.nlm.nih.gov), as available in October, 2005. Primary searches were conducted using BLAST search tools (Altschul *et al.*, 1990), or based on annotated genome files. The search for *Opp* (oligopeptide transport genes) and secretion systems was conducted using InterPro entries for *Bacillus subtilis* components (http://www.ebi.

Send correspondence to Augusto Schrank. Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Caixa Postal 15005, 91501-970 Porto Alegre, RS, Brazil. E-mail: aschrank@cbiot.ufrgs.br.

ac.uk/interpro, Mulder *et al*., 2005). The classification and analysis of proteases were done according to the MEROPS peptidase database (http://merops.sanger.ac.uk, Rawlings *et al*., 2004).

Results and Discussion

Oligopeptide importing

Mycoplasma genomes possess a diversity of ABC transporters predicted to be involved in the uptake of several inorganic and organic substrates. One class of ABC transporters, the peptide/opine/nickel uptake transporter family (3.A.1.5.1), is involved in oligopeptide uptake with high affinity for tripeptides (Transport Classification Database, http://www.tcdb.org).

The known genomic organization of the *Opp* operon in Mollicutes is shown in Table 1. Proteins encoded by this operon are anchored in the cell membrane; they transport oligopeptides from the extracellular milieu and represent an important form of nutrition (Detmers *et al*., 1998). *Mesoplasma florum* is the only species with no sequences related to the Opp system, as far as predicted by the annotation. The remaining genomes vary from one to two copies of the operon and also scattered single cistron copies. This distribution does not follow the division *hominis*/*pneumoniae* clades. The complete Opp operon (*OppABCDF*) was annotated only in *M. mycoides* and in *Phytoplasma*, and is present in two copies. It is noteworthy that in one of the *M. mycoides* operons the cistron order is altered from *ABCDF* to *BCDFA.* Moreover, in both species there are scattered copies of single components elsewhere in the genome. Two copies of the incomplete operon (lacking *OppA*) are present in three species (*M. pulmonis*, *M. penetrans* and *M. hyopneumoniae*). The same incomplete operon is present, as single copies, in all five genomes; however, in *M. synoviae*, the cistron order is different, and in *M. mobile* an extra copy of OppF is present.

The function of OppA as substrate-binding protein (oligopeptide recognition) is well recognized in bacteria (LeDeaux *et al*., 1997; Detmers *et al*., 1998). In *Mycoplasma hominis* (genome sequence not available), OppA functions as the P100 adherence-associated lipoprotein, and the operon is organized as *OppABCDF* (Henrich *et al*., 1999). Therefore, an important role for OppA in oligopeptide uptake could be expected in other Mollicutes. However, the OppA gene was found only in two Mollicute genome sequences (Table 1). This raises the question if OppA is really a necessary component of the oligopeptide uptake systems in these bacteria. Nevertheless, the low conservation of this protein could hinder its annotation. In addition, the habitat broadness of Mollicutes could result in strong selection/adaptation, especially for proteins involved in the recognition (binding) of oligopeptides, expected to be variable in different habitats. Also, lipoproteins are among the most prominent components of mycoplasma cell membranes (Razin *et al*., 1998), and the substrate recognition function of OppA could be fulfilled by one of these proteins.

Table 1 - Genomic organization of Opp uptake system in mycoplasmas. The five components are disposed in clusters or as isolated ORFs in the genome.

Microorganism	Clusters		Scattered	
	I	II		
Mycoplasma genitalium	*OppBCDF*	-[a]	-	-
Mycoplasma pulmonis	*OppBCDF*	*OppBCDF*	-	-
Mycoplasma penetrans	*OppBCDF*	*OppBCDF*	-	-
Mycoplasma pneumoniae	*OppBCDF*[b]	-	-	-
Ureaplasma urealyticum	*OppBCDF*	-	-	-
Mycoplasma gallisepticum	*OppBCDF*	*OppCDF*	*OppF*	*OppF*
Phytoplasma sp.	*OppABCDF*	*OppABCDF*	*OppF*	*OppA*
Mycoplasma mycoides	*OppBCDFA*	*OppABCDF*	*OppF*	*OppF*
Mycoplasma mobile	*OppBCDF*[c]	-	*OppF*	-
Mycoplasma hyopneumoniae[d]	*OppBCDF*	*OppBCDF*	-	-
Mycoplasma synoviae	*OppBCFD*	-	-	-

[a]not found. [b]*OppC* also known as *AmiD*. [c]*OppD* annotated as *pgk*, but contains the classical ATP-binding domain of the opp family. [d]strains J, 7448 [P], and 232.

Proteases

Bacterial development is dependent on a plethora of proteolytic activities involved in diverse functions, such as protein homeostasis, pathogenicity and nutrient acquisition. Mycoplasma genomes analysis revealed a complex distribution of these enzymes (Table 2). ATP-dependent proteases, such as Lon and FtsH, that degrade aberrant proteins, were found in all genomes analyzed here. Lon is a DNA-binding protease that degrades regulatory and abnormal proteins and has both the proteolytic and the ATPase domains. FtsH is a membrane protease that degrades membrane and cytoplasmic proteins. However, other important proteases involved in abnormal protein degradation, such as ClpPX and HslUV, were not annotated in five Mycoplasma species analyzed previously. This wider *in silico* survey of 12 Mollicute genomes supports the hypothesis outlined by Wong and Houry (2004) that the protein homeostatic process in these organisms has shifted through evolution towards favoring protein degradation rather than protein folding. Lon-defective *Escherichia coli* mutants remain phenotypically stable when overproducing HslU and HslV proteases, denoting a probable substrate overlapping among these proteins under certain physiological conditions (Wu *et al.*, 1999). This suggests that the lack of the HslUV system in mycoplasmas could be surpassed by the presence of Lon. FtsH is an endopeptidase, dependent on ATP and Zn^{2+}, that degrades abnormally-folded proteins

Table 2 - Distribution of putative proteases in mycoplasmas.

	Microorganism[a]	Mge	Mpu	Mpe	Mpn	Mga	Mmy	Mmo	Mhy	Msy
Group	Peptidase									
A	Signal peptidase I	-[b]	+[c]	-	-	+	-	-	+	+
B	Prolipoprotein signal peptidase II (SpaseII)	+	+	+	+	+	+	-	+	+
C	C1A subfamily peptidase	-	-	-	-		-	-	-	+
D	Family C39 (SunT protein) - ABC trans., Leucyl peptidase Heat shock ATP-dependent protease (Lon) Methionine aminopeptidase (map) Xaa-proaminopeptidase (pepP) o-sialoglycoprotein endopeptidase FtsH peptidase Putative glycoprotein endopeptidase	+	+	+	+	+	+	+	+	+
E	ATP-dependent serine protease-binding protein	+	-	-	-	-	+	+	+	-
F	Aminopeptidase	-	+	-	-	-	-	-	-	-
G	Subtilisin-like serine protease	-	+	-	-	-	-	-	+	+
H	Proline dipeptidase	-	+	+	-	+	+	-	+	+
I	Oligoendopeptidase F (pepF)	+	+	+	-	-	+	+	+	+
J	Prolyl aminopeptidase	+	-	+	+	-	-	-	-	-
K	Aminopeptidase C Collagenase Zn-dependent protease	-	-	+	-	+	-	-	-	-
L	Endopeptidase O	-	-	+	-	-	+	-	-	+

[a]Abbreviations: Mge, *Mycoplasma genitalium*; Mpu, *Mycoplasma pulmonis*; Mpe, *Mycoplasma penetrans*; Mpn, *Mycoplasma pneumoniae*; Mga, *Mycoplasma gallisepticum*; Mmy, *Mycoplasma mycoides*; Mmo, *Mycoplasma mobile*;. Mhy *Mycoplasma hyopneumoniae*; Msy, *Mycoplasma synoviae.* [b]not present. [c]present.

and the proteolysis products that otherwise cause cellular abnormalities. The absence of the HslUV system and the presence of Lon and FtsH appear to be conserved among mycoplasmas, except for *M. florum* that does not possess FtsH. The same applies to the absence of ClpPX, observed in Mollicute genomes except the Onion Yellow *Phytoplasma*, which possesses ClpX (Table 2).

Protease secretion in order to obtain peptides from the milieu is a common feature of bacteria (Morales *et al*., 2001). Most subtilisin-like and other serine proteases are secreted endopeptidases with little specificity for their substrates. The subtilisin-like serine proteases in the mycoplasma genomes belong to the subfamily S8A, which are endopeptidases. It is important to note that the gram-negative bacterium *Dichelobacter* has one subtilisin-like serine protease directly implicated in pathogenesis. Microbial pathogens often utilize secreted proteases as virulence factors, which may contribute largely to their pathogenicity. These proteases participate in tissue destruction, inactivation of host defense molecules, activation of key regulatory proteins or peptides and in nutrient acquisition. Some bacterial proteases can also activate bacterial toxins, thus triggering toxigenic pathogenesis. These proteases are also capable of degrading immunoglobulins and components of the complement system assisting infection propagation. Microbial proteases are very critical in enhancing pathogenesis of many severe diseases. However, only three out of the nine mycoplasmas analyzed here possess putative genes coding for serine proteases (Table 2).

Intracellular peptidases were also found in the present genome survey (Table 2). Bacterial leucyl aminopeptidases are involved in processing and maintaining a regular turnover of intracellular proteins and peptide breakdown products generated by intracellular proteases (Jenal and Hengge-Aronis, 2003). Methionyl aminopeptidase (Map) degrades “Ala/Ser-Pro” dipeptides, avoiding their accumulation, which could become toxic to the cell. Furthermore, oligopeptidase F (functions) acts in the degradation of intracellular oligopeptides generated from cell protease activity and can also cleave signal peptides. The main role of Map is to remove the initial methionine of many proteins during translation. The Xaa-Pro aminopeptidases, that hydrolyse “Xaa-Pro” dipeptides, and the prolyl dipeptidases, that release N-terminal residues from peptides, preferably (but not exclusively) a proline , were annotated in mycoplasma genomes (Table 2). The *o*-sialoglycoprotein endopeptidases cleave heavily *o*-sialoglycosylated proteins. These enzymes do not possess identifiable signal peptides and, therefore, are probably not secreted, establishing themselves as intracellular proteins. The involvement of these enzymes in pathogenicity remains to be demonstrated.

Protein secretion

The first description of secretion systems in mycoplasmas came with the genomes of *Mycoplasma genitalium* (Fraser *et al*., 1995) and *M. pneumoniae* (Himmelreich *et al*., 1996). These authors referred to the incomplete Sec translocation machinery in mycoplasmas compared to *Haemophilus influenzae* and *E. coli*, as a result of the lower complexity of the mycoplasma cell membrane. Now, the characterization of secretion systems in phylogenetically closer bacteria (Tjalsma *et al*., 2000; van Wely *et al*., 2001) and the availability of several Mollicute genomes allow a more accurate comparison. A survey of mycoplasma genomes, based on similarity searches using conserved domains of the Sec translocation proteins, yielded near-complete secretion systems in all individuals (Table 3). The absence of identifiable SecB protein is in agreement to the Sec translocation machinery of *B. subtilis*, in which other chaperone-like protein(s) would function as SecB-like protein (Tjalsma *et al*., 2000). The most intriguing fact is that some mycoplasmas are devoid of annotated pore-forming transmembrane proteins SecG or SecE.

Signal peptidases, such as signal peptidase I (Spase I) and signal peptidase II (Spase II), are responsible for cleaving off the hydrophobic N-terminal signal peptide regions of proteins to be exported or held in specific parts of the cell, such as the cytoplasmic membrane. Several types of Spase I from gram-negative and gram-positive bacteria have clear differences concerning gene size, gene copy number and substrate specificity, although there are substantial sequence similarities, as indicated by six different regions with conserved amino acids. Besides that, Spases type I are essential for cell life. The distribution of type I and type II Spases was found to be quite different in the mycoplasmas analyzed here (Table 2). Signal peptidases are thought to be fundamental in the pathogenicity of mycoplasmas, since their activities were found to be involved in the processing of a cilium adhesin from *M. hyopneumoniae* (Djordjevic *et al*., 2004) and *M. pneumoniae* and hence its their role in pathogenicity.

Table 3 - Sec translocon pathway secretion system components among known mycoplasma genomes.

Gene	SecA, SecY, Ffh, ftsY, scRNA	YajC, SecB	SecDF	SecE	SecG
Microorganism[a]					
Mge	+[b]	-[c]	-	+	-
Mpu	+	-	+	-	-
Mpe	+	-	+	+	+
Mpn	+	-	-	+	+
Uur	+	-	-	+	+
Mga	+	-	-	+	+
Phy	+	-	+	+	-
Mmy	+	-	+	+	-
Mmo	+	-	+	-	-
Mhy*e*	+	-	+	-	+
Msy	+	-	+	-	-

[a]Abbreviations: Mge, *Mycoplasma genitalium*; Mpu, *Mycoplasma pulmonis*; Mpe, *Mycoplasma penetrans*; Mpn, *Mycoplasma pneumoniae*; Uur, *Ureaplasma urealyticum*; Mga, *Mycoplasma gallisepticum*; Phy, *Phytoplasma*, Mmy, *Mycoplasma mycoides*; Mmo, *Mycoplasma mobile*;. Mhy *Mycoplasma hyopneumoniae*; Msy, *Mycoplasma synoviae.* [b]present. [c]not present.

Acknowledgments

C.C.S, J.B. and L.B. are recipients of CAPES predoctoral fellowships. This work was supported by MCT/CNPq and FAPERGS. The authors thank the two anonymous referees for their useful suggestions to improve the manuscript.

References

Akiyama Y and Ito K (2003) Reconstitution of membrane proteolysis by FtsH. J Biol Chem 278:18146-18153.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410.

Catrein I, Herrmann R, Bosserhoff A and Ruppert T (2005) Experimental proof for a signal peptidase I like activity in *Mycoplasma pneumoniae*, but absence of a gene encoding a conserved bacterial type I SPase. FEBS J 272:2892-2900.

Chang PC and Lee YH (1992) Extracellular autoprocessing of a metalloprotease from *Streptomyces cacaoi*. J Biol Chem 267:3952-3958.

Dave JA, Gey van Pittius NC, Beyers AD, Ehlers MR and Brown GD (2002) Mycosin-1, a subtilisin-like serine protease of *Mycobacterium tuberculosis*, is cell wall-associated and expressed during infection of macrophages. BMC Microbiol 2:30. http://www.biomedcentral.com/bmcmicrobiol/

Detmers FJ, Kunji ER, Lanfermeijer FC, Poolman B and Konings WN (1998) Kinetics and specificity of peptide uptake by the oligopeptide transport system of *Lactococcus lactis*. Biochem 37:16671-16679.

Djordjevic SP, Cordwell SJ, Djordjevic MA, Wilton J and Minion FC (2004) Proteolytic processing of the *Mycoplasma hyopneumoniae* cilium adhesin. Infect Immun 72:2791-2802.

Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, *et al*. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397-403.

Fu GK, Smith MJ and Markovitz DM (1997) Bacterial protease Lon is a site-specific DNA-binding protein. J Biol Chem 272:534-538.

Henrich B, Hopfe M, Kitzerow A and Hadding U (1999) The adherence-associated lipoprotein P100, encoded by an opp operon structure, functions as the oligopeptide-binding domain OppA of a putative oligopeptide transport system in *Mycoplasma hominis*. J Bacteriol 181:4873-4978.

Herman C, Prakash S, Lu CZ, Matouschek A and Gross CA (2003) Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH. Mol Cell 11:659-669.

Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li BC and Herrmann R (1996) Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res 24:4420-4449.

Hsu T, Artiushin S and Minion FC (1997) Cloning and functional analysis of the P97 swine cilium adhesin gene of *Mycoplasma hyopneumoniae*. J Bacteriol 179:1317-1323.

Ito K and Akiyama Y (2005) Cellular functions, mechanism of action, and regulation of FtsH protease. Annu Rev Microbiol 59:211-231.

Jenal U and Hengge-Aronis R (2003) Regulation by proteolysis in bacterial cells. Curr Opin Microbiol 6:163-172.

Knight CG, Dando PM and Barrett AJ (1995) Thimet oligopeptidase specificity: Evidence of preferential cleavage near the C-terminus and product inhibition from kinetic analysis of peptide hydrolysis. Biochem J 308:145-150.

Kortt AA and Stewart DJ (1994) Properties of the extracellular acidic proteases of *Dichelobacter nodosus*. Stability and specificity of peptide bond cleavage. Biochem Mol Biol Int 34:1167-1176.

LeDeaux JR, Solomon JM and Grossman AD (1997) Analysis of non-polar deletion mutations in the genes of the spo0K (opp) operon of *Bacillus subtilis*. FEMS Microbiol Lett 153:63-69.

Maeda H and Yamamoto T (1996) Pathogenic mechanisms induced by microbial proteases in microbial infections. Biol Chem Hoppe Seyler 377:217-226.

Mellors A and Sutherland DR (1994) Tools to cleave glycoproteins. Trends Biotechnol 12:15-18.

Morales P, Fernandez-Garcia E, Gaya P, Medina M and Nunez M (2001) Hydrolysis of caseins and formation of hydrophilic and hydrophobic peptides by wild *Lactococcus lactis* strains isolated from raw ewes' milk cheese. J Appl Microbiol 91:907-915.

Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bradley P, Bork P, Bucher P, Cerutti L, *et al.* (2005) InterPro, progress and status in 2005. Nucleic Acids Res 33:D201-D205.

Myara I, Cosson C, Moatti N and Lemonnier A (1994) Human kidney prolidase-purification, preincubation properties and immunological reactivity. Int J Biochem 26:207-214.

Nagy I, Banerjee T, Tamura T, Schoofs G, Gils A, Proost P, Tamura N, Baumeister W and De Mot R (2003) Characterization of a novel intracellular endopeptidase of the alpha/beta hydrolase family from *Streptomyces coelicolor* A3(2). J Bacteriol 185:496-503.

Paetzel M, Karla A, Strynadka NC and Dalbey RE (2002) Signal peptidases. Chem Rev 102:4549-4580.

Rawlings ND, Tolle DP and Barrett AJ (2004) MEROPS: The peptidase database. Nucleic Acids Res 32:D160-D164.

Razin S, Yogev D and Naot Y (1998) Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev 62:1094-1156.

Stephenson K (2005) Sec-dependent protein translocation across biological membranes: Evolutionary conservation of an essential protein transport pathway. Mol Membr Biol 22:17-28.

Takaya A, Tomoyasu T, Tokumitsu A, Morioka M and Yamamoto T (2002) The ATP-dependent Lon protease of *Salmonella enterica* serovar *Typhimurium* regulates invasion and expression of genes carried on *Salmonella* pathogenicity island 1. J Bacteriol 184:224-232.

Tan PS, Poolman B and Konings WN (1993) Proteolytic enzymes of *Lactococcus lactis*. J Dairy Res 60:269-286.

Tjalsma H, Bolhuis A, Jongbloed JDH, Bron S and van Dijl JM (2000) Signal peptide-dependent protein transport in *Bacillus subtilis*: A genome-based survey of the secretome. Microbiol Mol Biol Rev 64:515-547.

van Ham RC, Kamerbeek J, Palacios C, Rausell C, Abascal F, Bastolla U, Fernandez JM, Jimenez L, Postigo M, Silva FJ, *et al.* (2003) Reductive genome evolution in *Buchnera aphidicola*. Proc Natl Acad Sci USA 100:581-586.

van Wely KHM, Swaving J, Freudl R and Driessen AJM (2001). Translocation of proteins across the cell envelope of Gram-positive bacteria. FEMS Microbiol Rev 25:437-454.

Vasconcelos AT, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, Almeida DF, Almeida LG, Almeida R, Alves-Filho L, *et al*. (2005) Swine and poultry pathogens: The complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. J Bacteriol 187:5568-5577.

Wang XG, Kidder JM, Scagliotti JP, Klempner MS, Noring R and Hu LT (2004) Analysis of differences in the functional properties of the substrate binding proteins of the *Borrelia burgdorferi* oligopeptide permease (opp) operon. J Bacteriol 186:51-60.

Wong P and Houry WA (2004) Chaperone networks in bacteria: Analysis of protein homeostasis in minimal cells. J Struct Biol 146:79-89.

Wu WF, Zhou Y and Gottesman S (1999) Redundant *in vivo* proteolytic activities of *Escherichia coli* Lon and the ClpYQ (HslUV) protease. J Bacteriol 181:3681-3687.

Associate Editor: Ana Tereza Vasconcelos