environmental microbiology reports

Environmental Microbiology Reports (2017) **00**(00), 00–00 doi:10.1111/1758-2229.12566

# A matter of hierarchy: activation of orfamide production by the post-transcriptional Gac-Rsm cascade of *Pseudomonas protegens* CHA0 through expression upregulation of the two dedicated transcriptional regulators

**Patricio Martín Sobrero,[1†] Andrés Muzlera,[1†] Julieta Frescura,[1] Edgardo Jofré[2] and Claudio Valverde[1*]**

[1]*CONICET, Departamento de Ciencia y Tecnología, Laboratorio de Bioquímica, Microbiología e Interacciones Biológicas en el Suelo. Roque Sáenz Peña 352, Bernal B1876BXD, Universidad Nacional de Quilmes, Buenos Aires, Argentina.*

[2]*CONICET, Facultad de Ciencias Exactas Físico-Químicas y Naturales, Departmento de Ciencias Naturales. Ruta Nacional 36 Km 601, Universidad Nacional de Río Cuarto, 5800 Río Cuarto, Córdoba, Argentina.*

## Summary

**In this work, we surveyed the genome of *P. protegens* CHA0 in order to identify novel mRNAs possibly under the control of the Gac-Rsm cascade that might, for their part, serve to elucidate as-yet-unknown functions involved in the biocontrol of plant pathogens and/or in cellular processes required for fitness in natural environments. In view of the experimental evidence from former studies on the Gac-Rsm cascade, we developed a computational screen supported by a combination of sequence, structural and evolutionary constraints that led to a dataset of 43 potential novel mRNA targets. We then confirmed several mRNA targets experimentally and next focused on two of the respective genes that are physically linked to the orfamide biosynthetic gene cluster and whose predicted open-reading frames resembled cognate LuxR-type transcriptional regulators of cyclic lipopeptide clusters in related pseudomonads. In this report, we demonstrate that in strain CHA0, orfamide production is stringently dependent on a functional Gac-Rsm cascade and that both mRNAs encoding transcriptional regulatory proteins are under direct translational control of the RsmA/E proteins. Our results have thus revealed a hierarchical control over the expression of orfamide biosynthetic genes with the final transcriptional control subordinated to the global Gac-Rsm post-transcriptional regulatory system.**

Received 24 March, 2017; accepted 21 June, 2017. *For correspondence. E-mail cvalver@unq.edu.ar; Tel. +54-11-4365-7100 ext. 5638; Fax +54-11-4365-7132.

†These authors contributed equally to this work.



## Introduction

In *Pseudomonas* spp., the ubiquitous and global regulatory two-component system GacS-GacA controls sensitive traits like cell–cell communication, secondary metabolite production and motility (Lapouge *et al.*, 2008). In *Pseudomonas protegens* Pf-5, the global regulation by GacA extends to more than 10% of the annotated genes, whereas in *P. aeruginosa* M18 that system influences the expression of some 15% of the encoded genes (Hassan *et al.*, 2010; Wei *et al.*, 2013). Such a striking impact on global gene expression operates mainly at the post-transcriptional level through the action of GacA-responsive small, noncoding and *trans*-acting regulatory RNAs (sRNAs) of the Csr/Rsm family (Lapouge *et al.*, 2008). The entire global regulatory pathway, known as the Gac-Rsm cascade in the genus *Pseudomonas*, is also present in other members of the γ-proteobacteria, thus tracing the ancestral and pivotal role of that cascade in gene regulatory networks (Lapouge *et al.*, 2007).

In the biocontrol strain *P. protegens* CHA0 the Gac-Rsm cascade upregulates the expression of genes required for production of all secondary metabolites that have been identified as relevant to the biocontrol of fungal and oomycete plant pathogens and to insect toxicity, as well as to escaping predation by bacterivorous protozoa and nematodes, for example, 2,4-diacetylphoroglucinol (DAPG), pyoluteorin, HCN and exoprotease A (Haas and Defago, 2005; Jousset *et al.*, 2006; Romanowski *et al.*, 2011; Flury *et al.*, 2017). The histidine kinase

GacS responds to an unknown cellular signal that accumulates upon leaving the exponential phase of growth (Heeb *et al*., 2002). As a consequence, the response regulator GacA becomes active and induces expression of the three molecular mimic sRNAs RsmX, RsmY and RsmZ (Kay *et al*., 2005), which transcripts exploit their molecular-mimicry decoying capabilities to compete with mRNAs for ligation to the translational repressor RNA binding proteins RsmA and RsmE (Reimmann *et al*., 2005), thus inactivating those proteins so as to permit ongoing translation of the products of the aforementioned biocontrol genes. Genetic and structural data accumulated in studies of the Gac-Rsm system of *P. protegens* CHA0 provided insights into the nucleotide-sequence requirement and topology amenable to the molecular decoying of the RsmA/RsmE proteins by those sRNAs and target mRNAs. Both RNA-binding proteins display a high susceptibility to binding to the sequence motif ANGGA exposed in the loop of short hairpin turns (Lapouge *et al*., 2007; Schubert *et al*., 2007). If those same motifs are near to or they overlap the ribosome-binding site (RBS) of an mRNA, binding of either of the RsmA/E proteins will interfere with the translation initiation. Thus, unidentified mRNAs under translational control by the proteins of the Gac-Rsm cascade should contain the RsmA/E recognition motif in the periphery of its RBS, particularly so with respect to an ANGGA segment as part of a stem-loop structure exposing the trinucleotide GGA (Valverde *et al*., 2004; Lapouge *et al*., 2013). Moreover, if such post-transcriptional control is instrumental for bacterial fitness in a set of phylogenetically related species, we could predict that the sequence-structure decoy required for Gac-Rsm control has to be evolutionarily conserved in functional homolog mRNA sequences and could serve as well as a reliable criterion for identifying candidate genes that might be under Gac-Rsm regulation.

The production of cyclic lipopeptides (CLPs) – they considered as relevant agents effecting the biocontrol of different plant pathogens (D'Aes *et al*., 2010)- has been characterized at the biochemical and genetic level in several *Pseudomonas* species. The CLPs are bimodular molecules, with a variable fatty acid tail linked to a lactone-cyclized oligopeptide ring. The peptide module comprises 8 to 25 amino acids, and is assembled by nonribosomal peptide synthetases (NRPSs). Variations in the chemical nature of the fatty-acid tail and the primary structure of the oligopeptide have given rise to and thus characterize a number of CLP families (Raaijmakers *et al*., 2010). The vast chemical diversity of these biomolecules, along with their surfactant and antifungal activity, prompted their consideration as novel bioproducts with applications in agronomy for the control of plant pests (Ongena and Jacques, 2008). Of interest to us is that the gene encoding the NRPSs are generally found in defined clusters together with other genetic elements involved in the regulation of expression and exportation of CLPs (Raaijmakers *et al*., 2010). The main CLP produced by *P. protegens* CHA0 and by that strain's close phylogenetic relative *P. protegens* Pf-5 is orfamide A (Gross *et al*., 2007). Orfamide A induces the lysis of zoospores from the oomycetes *Phytophtora ramorum* and *Pythium ultimum*, inhibits formation of appressoria in the fungal pathogen *Magnaporthe oryzae*, and it also contributes to the insect oral toxicity of strain CHA0 with respect to the cabbage moth *Plutella xylostella* (Ma *et al*., 2016; Flury *et al*., 2017). Even though the involvement of sRNA elements in the control of CLP production has been reported in other pseudomonads (Song *et al*., 2015), the operation of post-transcriptional regulation over the expression of the orfamide A biosynthetic genes in strain CHA0 has not been yet characterized.

In the experiments reported here, a computational prediction of mRNAs that may be subject to direct control of the Gac-Rsm system in the biocontrol model strain *P. protegens* CHA0 led us to identify a pair of genes encoding LuxR-type transcriptional regulators flanking the biosynthetic operon for the CLP orfamide A, which alleles were subsequently validated experimentally as novel authentic targets of the Gac-Rsm cascade.

## Results and discussion

### In silico *prediction of novel Gac-Rsm mRNA targets in* P. protegens *CHA0 and experimental validation*

As stated above, the Gac-Rsm cascade of *P. protegens* CHA0 controls the expression of genes involved in production of secondary metabolites for the biocontrol of fungal phytopathogens and for escaping predation by bacterivorous protozoa and nematodes (Jousset *et al*., 2006; Lapouge *et al*., 2008; Romanowski *et al*., 2011). Genes under direct control by the Gac-Rsm system include the *hcn* operon (for synthesis of hydrogen cyanide), *aprA* (for the extracellular metalloprotease A), *phlACBDE* (for the antibiotic DAPG) and *pltA* (part of the *plt* operon for the antibiotic pyoluteorin) (Lapouge *et al*., 2013). Evidence has been garnered from related pseudomonads that the Gac-Rsm cascade has a much broader influence on gene expression than previously estimated (Valverde and Haas, 2008; Hassan *et al*., 2010; Wei *et al*., 2013); likewise, the molecular details of the RNA motifs that are recognized by RsmE protein (Lapouge *et al*., 2007; Schubert *et al*., 2007; Lapouge *et al*., 2013) suggest that other mRNAs might exist that could constitute direct targets of RsmE as well as of RsmA proteins. We, therefore, carried out an *in silico* survey to encounter novel targets of the Gac-Rsm

cascade in the genome of *P. protegens* CHA0 (Jousset *et al*., 2014), which search was based on the experimental evidence available for this system (Schubert *et al*., 2007). The *in silico* pipeline (Supporting Information Fig. 1) took into consideration the sequence, structural and evolutionary features of the putative RNA recognition sites preferred by the RsmA/E proteins. The prediction was focused on a 300-bp DNA sequence around every annotated open-reading frame (ORF), ranging from 250 bp upstream to 50 bp downstream from the start codon. The reasons for this selection were: (1) RsmA/E-targeted binding sites distant from the RBS cannot be excluded (Lapouge *et al*., 2007); (2) mRNAs usually exhibit a relaxed structure that facilitates ribosome binding and translation along the first 40–50 bp within the coding region (Kudla *et al*., 2009); and (3) possible misannotation of the start codon during automatic genome curation. All 300-bp DNA sequences retrieved were queried for the presence of the RsmE decoy 5′-WCANGGANGW-3′ (Schubert *et al*., 2007). The high degree of identity between RsmA and RsmE and the successful functional exchange reported for these two proteins (Reimmann *et al*., 2005), prompted us to consider that the search may also identify mRNA targets for RsmA. One disadvantage of this approach is that putative RsmA/E targets located within the coding region or in the 3′-UTR of mRNAs (Holmqvist *et al*., 2016) might have been disregarded.

The search resulted in a total of 116 putative DNA fragments with matching sequence motifs (Supporting Information Fig. 1a and Supporting Information Table 1). Almost 2/3 of the hits contained the motif within the 5′-UTR, 25% overlapped the start codon and fewer than 10% of the hits had the motif within the first 50 bp of the coding region (Supporting Information Fig. 1b). With regards to the central pentamer of the query string (ANGGA), ACGGA was the rarest combination ($<$15%), with AAGGA, AGGGA and AUGGA being present at comparable frequencies (ranging 25–33%) (Supporting Information Fig. 1c). Within the 5′-UTRs, most of the motifs overlapped the Shine–Dalgarno sequence (Supporting Information Fig. 1d). The primary list was narrowed down upon application of the following three filter criteria: (1) the putative RsmA/E-binding site must be located at or near the Shine–Dalgarno sequence; (2) the site sequence must fold into a stem-loop structure exposing the trinucleotide GGA in the loop; (3) the sequence motif must be evolutionarily conserved in the 5′-UTRs of the orthologous genes in phylogenetically related pseudomonads. Upon application of these filters, we arrived at a list of 45 putative mRNA targets for RsmA/E proteins (Supporting Information Table 2) that included two well characterized Gac-Rsm mRNA targets, *hcnA* and *aprA* (Blumer *et al*., 1999), thus supporting our predictive strategy and the resulting output (Supporting Information Table 2). Two of the known Gac-Rsm targets (*phlA* and *pltA*) were not detected by our approach, because of a slight divergence in the decoy present in these two 5′-UTRs from the RsmE-consensus recognition motif (Lapouge *et al*., 2013).

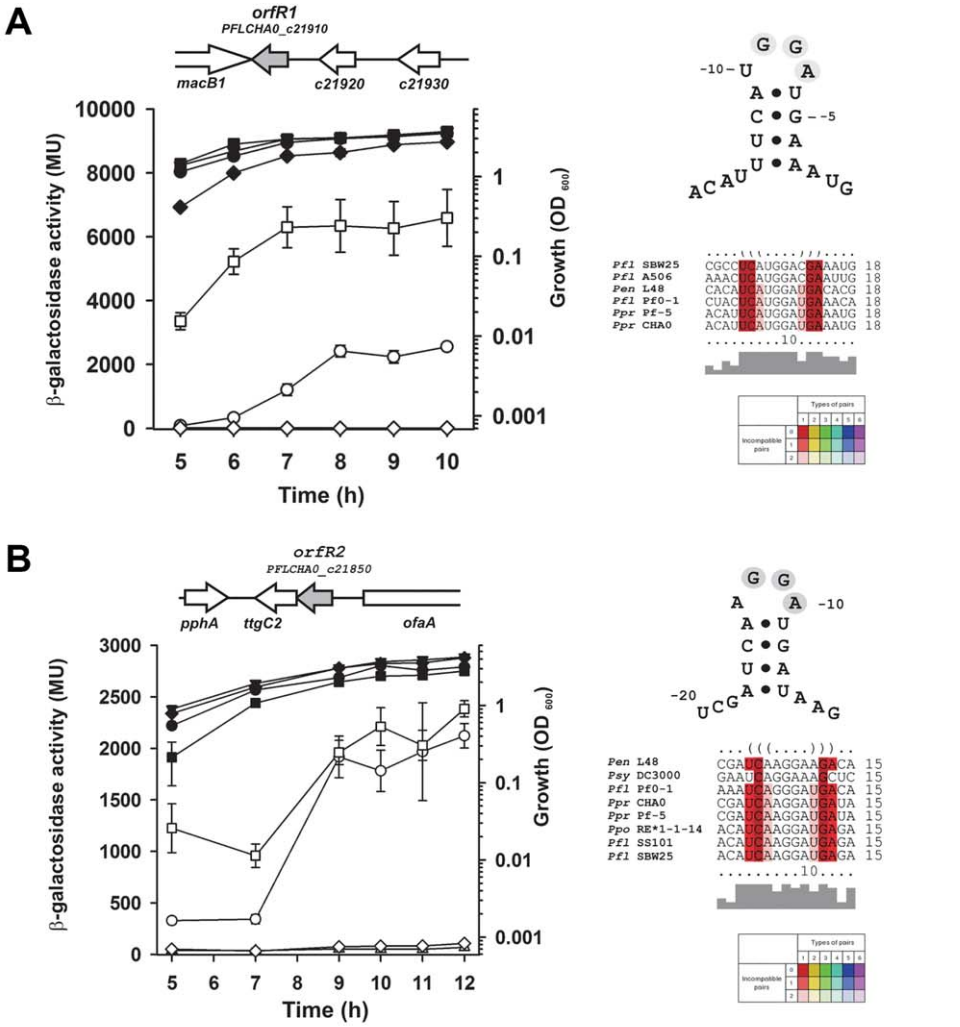
Four putative novel target mRNAs were chosen for experimental confirmation of their regulated expression by the Gac-Rsm system in *P. protegens* CHA0. A '*lacZ* translational-reporter fusion with the 5′-UTR of a *bona-fide* Gac-Rsm target gene must be induced in the transition from exponential to stationary phase in the wild-type strain CHA0; be strongly downregulated in a *gacS*, *gacA* or *rsmX*/*Y*/*Z* mutant background; and be upregulated in an *rsmA*/*E* mutant background (Kay *et al*., 2005; Crespo and Valverde, 2009). Three out of the four genes tested matched the requirements for Gac-Rsm regulation (Supporting Information Fig. 2): *PFLCHA0_c00250* (*dprA*, for the DNA-processing chain A) (Mortier-Barriere *et al*., 2007), *PFLCHA0_c02690* (*pksP*, encoding a small putative acyl-carrier protein) and *PFLCHA0_c21910* (encoding an ortholog of a LuxR-type transcriptional factor). The fourth gene tested, *PFLCHA0_c04880* (*soj1*), involved in the cell-division process (Murray and Errington, 2008), did not contain a typical pattern corresponding to a Gac-Rsm target, possibly because of the lack of a proper secondary structure (Supporting Information Fig. 2c). For reasons that will become clear below, we focussed this study on the newly identified Gac-Rsm target gene *PFLCHA0_c21910*.

### *Control of expression of the two mRNAs encoding transcriptional regulators of the orfamide biosynthetic gene cluster by Gac-Rsm*

The expression pattern of a *PFLCHA0_c21910*′-'*lacZ* fusion strictly matched that of a true Gac-Rsm target mRNA (Fig. 1A). The gene *PFLCHA0_c21910* encodes an ortholog of a LuxR-type transcriptional factor that is located downstream from a cluster encoding the biosynthetic proteins for the CLP orfamide A (Fig. 1A) (Ma *et al*., 2016). The production of orfamide has been recently demonstrated in *P. protegens* CHA0 (Ma *et al*., 2016; Flury *et al*., 2017), but that CLP's genetic-regulatory features remain obscure. This lack of information prompted us to investigate the regulation of orfamide production by the Gac-Rsm cascade in strain CHA0. The architecture of the orfamide-gene cluster in strain CHA0 is syntenic to clusters for CLP biosynthesis from other pseudomonads containing two flanking genes for LuxR-type transcription regulators encoded within a single locus, *PFLCHA0_c21910* located downstream and *PFLCHA0_c21850* upstream from the biosynthetic genes (Supporting Information Fig. 3). We will hereafter

refer to these alleles as *orfR1* and *orfR2*, respectively. As our *in silico* search (Supporting Information Fig. 1) did not reveal a putative RsmA/E binding site for *orfR2*, we manually inspected the genomic annotation and realized that the true start codon was at 57 nucleotides (19 codons) downstream from the location originally presumed and, therefore, had been previously automatically misannotated. After this correction, a *bona-fide* RsmA/E binding site (UCAAGGAUGA) was detected at positions −8 to −17 relative to the true start codon (Fig. 1B) and for this reason had been missed in our original search (Supporting Information Fig. 1). The expression pattern of a '*lacZ* translational fusion also indicated a typical activation of the *orfR2* gene during the transition from




C

| | *hcnA* -2.3 Kcal/mol | *aprA* -3.2 Kcal/mol | *orfR1* -3.1 Kcal/mol | *orfR2* -3.1 Kcal/mol |
|---|---|---|---|---|
| GacS activation factor | 25x | 38x | 135x | 36.4x |
| RsmA/E repression factor | 4.3x | 3.3x | 5.2x | 1.0x |
| Gac/Rsm control | 108x | 124x | 704x | 37.2x |

**Fig. 1.** Control of expression of the two genes encoding transcriptional regulators of the orfamide-biosynthetic gene cluster, *orfR1* and *orfR2*, by the Gac-Rsm system of *P. protegens* CHA0. Impact of the *gacS*, *rsmA/rsmE*, and *rsmX/Y/Z* mutations on the expression of a plasmid-borne *orfR1*'-'*lacZ* (A) or *orfR2*'-'*lacZ* (B) translational fusion. $\beta$-Galactosidase activity (left *ordinates*) was determined in strains CHA0 (open circles, wild-type context), CHA19 (open triangles, *gacS* mutant), CHA1144 (open diamonds, *rsmXYZ* mutant) and CHA1009 (open squares, *rsmAE* mutant), following procedures reported elsewhere (Valverde *et al*., 2004). The corresponding solid symbols represent the growth curves of each strain (right *ordinates* on an exponential scale). Each value represents the average of the data from three separate cultures ± standard deviation. The scheme above each panel illustrates the genomic context of *orfR1* and *orfR2* in the CHA0 genome. The panels to the right depict the Mfold secondary-structure prediction (set to default parameters at 30°C) of the RsmA/E-binding motif identified for the *orfR1* and *orfR2* 5′-UTRs along with the corresponding standard free energy (Zuker, 2003) and the evolutionary conservation of the 5′-UTR RNA sequence and its secondary structure as inferred by locARN. The block pattern in grey beneath the sequences depicts the degree of nucleotide conservation at each position of the alignment. The string of dots and parentheses on top of the alignment represents the consensus secondary structure of the alignment. Base pairs in the structure are indicated by corresponding opening and closing parentheses. Compatible base pairs are colored in the alignment, where the hue shows the number of different types C-G, G-C, A-U, U-A, G-U or U-G of compatible base pairs in the corresponding columns. Thus, the hue shows sequence conservation of the base pair. The saturation decreases with the number of incompatible base pairs, and indicates the structural conservation of the base pair (Will *et al*., 2012). *Pen*, *P. entomophila*; *Pfl*, *P. fluorescens*; *Ppo*, *P. poae*; *Ppr*, *P.protegens*; *Psy*, *P. syringae* pv. *tomato*; C) Gac-Rsm control strength over *orfR1* and *orfR2*. The local structure and folding energy (below) were predicted with Mfold set to default parameters at 30°C (Zuber *et al*., 2003). The base positions are relative to the first coding nucleotide. The characteristic single-stranded GGA triplets regarded as essential for RsmA/E binding (Valverde *et al*., 2004) are highlighted with grey circles. The RsmA/E binding sites of *hcnA* and *aprA* (Blumer *et al*., 1999) are included for comparison. For each translational fusion, the GacS activation factor is the CHA0/CHA19 expression ratio, the RsmA/E repression factor is the CHA1009/CHA0 expression ratio, and the Gac-Rsm control factor represents the product of the GacS activation factor × the RsmA/E repression factor. The factors were calculated from the $\beta$-galactosidase activities shown in (A) and (B), and correspond to cultures sampled at early stationary phase (i.e., 7 h for *orfR1*'-'*lacZ* (a) and 9 h after inoculation for *orfR2*'-'*lacZ* (B). The *hcnA*'-'*lacZ* and *aprA*'-'*lacZ* factors were calculated from published data (Reimmann *et al*., 2005).

exponential to stationary phase in wild-type CHA0 (Fig. 1B) and a strong downregulation in the absence of functional GacS-sensor or RsmX/Y/Z sRNAs (Fig. 1B); in contrast, the upregulation in an *rsmAE* mutant was only evident in the exponential phase and the expression was indistinguishable from that of the wild-type CHA0 in stationary phase (Fig. 1B). Nevertheless, the conclusion is that both *orfR1* and *orfR2* seem to be direct targets of the Gac-Rsm riboregulatory network in *P. protegens* CHA0. The comparison of the quantitative impact that the Gac and Rsm mutations have on the expression of each translational fusion suggested that the Gac-Rsm pathway exerts a more stringent control over *orfR1* than over *orfR2* (Fig. 1C).

### *Activation of orfamide production in* P. protegens *CHA0 by the Gac-Rsm cascade*

In view of the finding that both putative transcriptional regulators flanking the orfamide-biosynthetic cluster exhibit a typical Gac-Rsm-dependent expression (Fig. 1), we decided to investigate whether or not the production of this CLP is subordinated to the Gac-Rsm cascade in strain CHA0. When grown in King's B medium, *P. protegens* CHA0 produces a set of CLP isoforms that correspond to orfamide derivatives differing in the fatty-acid moiety and in the fourth amino acid residue (Ma *et al*., 2016). Indeed, the drop-collapse activity in the supernatant of wild-type CHA0 was activated in the early stationary phase (Fig. 2B), strongly resembling the expression pattern of HCN production and exoprotease activity (Blumer *et al*., 1999). Moreover, the drop-collapse activity was abolished in *gacS* and *rsmXYZ* mutants, but was stimulated in the *rsmAE* mutant, thus revealing a positive regulation by the Gac-Rsm pathway (Fig. 2A and B). The reverse-phase HPLC profile of purified CLPs from the wild-type CHA0 and its isogenic Gac-Rsm mutants confirmed that the production of orfamide variants is under strict Gac-Rsm control (Fig. 2C) and that this phenotype is consistent with the pattern of drop-collapse activity in the culture supernatants (Fig. 2A). The identity of the major peak eluted at 60 min, both from the wild-type CHA0 and from the *rsmAE* mutant CHA1009, was confirmed as orfamide A by MALDI-TOF-MS fragmentation (Fig. 2C). All together, these results demonstrate that in *P. protegens* CHA0 the Gac-Rsm pathway orchestrates orfamide production, possibly (at least partly) through a direct control over the expression of *orfR1* and *orfR2*, those alleles most likely encoding the Lux-type transcriptional regulators of the CLP biosynthetic cluster.

### *OrfR1 and OrfR2 as functional homologues of the* P. fluorescens *SBW25 viscosin-transcription regulators ViscAR and ViscBCR*

A phylogenetic analysis suggested that the products of the *orfR1* and *orfR2* genes had coevolved independently from each other and from other members of the family that are involved in quorum sensing (Supporting Information Fig. 4). For those CLP gene clusters encoded within one chromosomal locus, one LuxR-type transcriptional regulator is typically encoded upstream from the biosynthetic genes (e.g., *orfR2* in CHA0), whereas a second is encoded downstream from that gene cluster (e.g., *orfR1* in CHA0) (Supporting Information Fig. 3). In *P. fluorescens* strain SBW25, *viscAR* and *visBCR* are the two nonredundant LuxR-type transcriptional regulators required for full production of

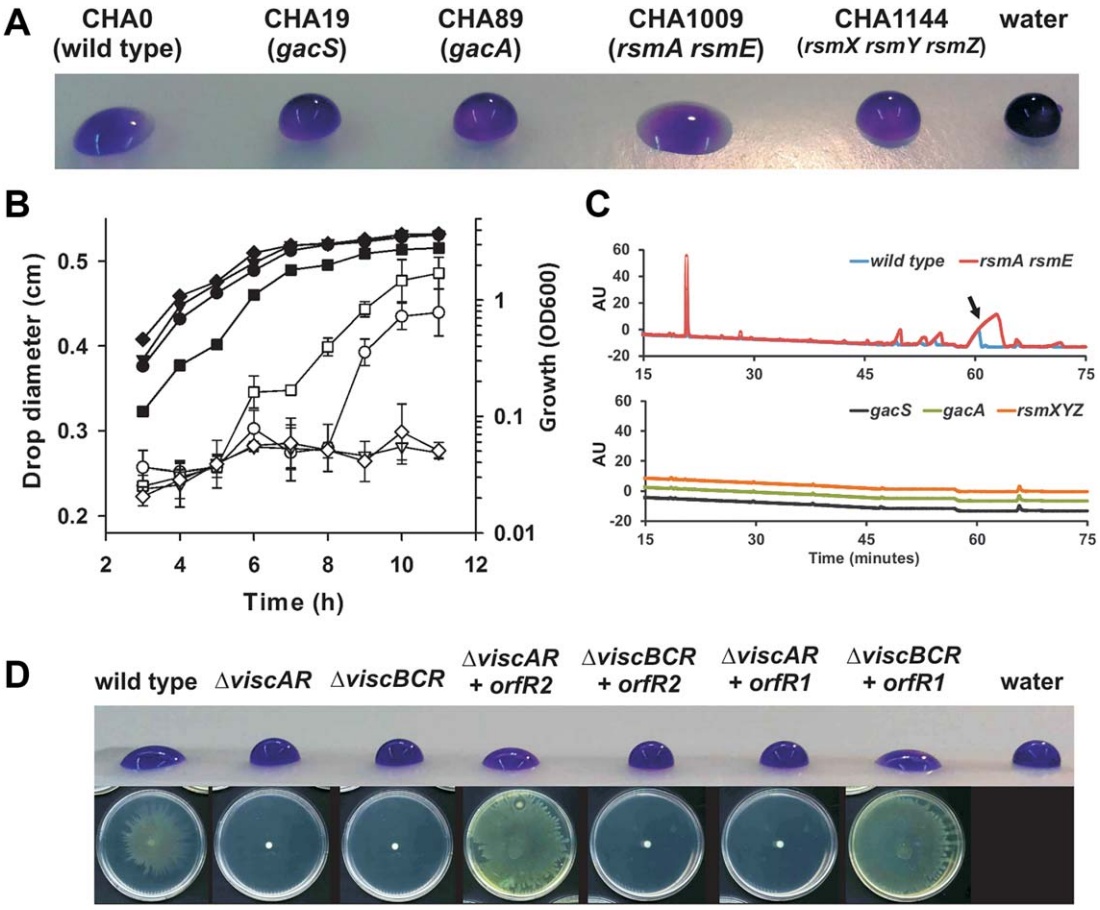



**Fig. 2.** Activation of orfamide production in *P. protegens* CHA0 by the Gac-Rsm cascade.
A. Drop-collapse activity of stationary-phase cell-free-culture supernatants of *P. protegens* CHA0 wild-type and mutant strains lacking different genetic elements of the Gac-Rsm cascade. The standard drop-collapse assay was conducted as described previously (de Bruijn and Raaijmakers, 2009).
B. Time course (*abscissa* in h) for the onset of the drop-collapse activity (left *ordinate*) of wild-type and mutant strains along the corresponding growth curve in nutrient broth. CHA0 (open circles, wild-type context), CHA19 (open triangles, *gacS* mutant), CHA1144 (open diamonds, *rsmXYZ* mutant) and CHA1009 (open squares, *rsmAE* mutant). The corresponding solid symbols represent the growth curves of each strain. Each value is the average of the data from three separate cultures ± standard deviation.
C. CLP HPLC profile of the wild-type and Gac-Rsm mutants (*ordinates* in arbitrary units, AU). The arrow points to the major peak present in the wild-type and in the *rsmAE* mutant strain CHA1009, whose identity was confirmed as orfamide A. The culture lipopeptides were isolated by acid precipitation and purified by RP-HPLC essentially as described (Alvarez *et al*., 2012). The *P. protegens* strains were grown in PSFM medium without $FeCl_3$ (De La Fuente *et al*., 2004) for 24 h at 28°C with agitation (150 rpm). Precipitated lipopeptides were suspended in 100% methanol and injected into an HPLC system (Infinity LC Grad, Agilent) equipped with an Analytical Zorbax C18 reverse-phase column (4.6 mm × 150 mm, Agilent) and an ultraviolet detector (Agilent). The compounds were eluted at a flow rate of 1 ml/min with an increasing gradient of acetonitrile from 10 to 100%. Elution was performed as follows: 10% for 0–10 min, 35–80% for 10–45 min, 80% for 45–55 min and 80–100% for 55–65 min. The identity of the lipopeptides eluted in the fractions corresponding to orfamide peaks were confirmed by MALDI-TOF mass spectrometry analysis at the CEQUIBIEM facility (Center of Chemical and Biochemical Studies by Mass Spectrometry – Buenos Aires University).
D. Functional complementation of *P. fluorescens* SBW25 mutants lacking either of the two transcriptional regulators for viscosin production, ViscAR and ViscBCR, by the *P. protegens* CHA0 transcriptional regulators OrfR2 and OrfR1, respectively. The upper row illustrates the drop-collapse activity of stationary-phase cell-free culture supernatants of the corresponding strains indicated above the panel as compared to water as a negative control. The lower row depicts the swarming phenotype of the same strains. The complemented Δ*viscAR* and Δ*viscBCR* mutant strains contained a low-copy plasmid-borne allele of *P. protegens orfR1* or *orfR2* genes. The swarming motility of the *P. fluorescens* SBW25 strains was assessed on soft (0.6% [w/v] agar) standard succinate agar medium (de Bruijn and Raaijmakers, 2009).

viscosin (de Bruijn and Raaijmakers, 2009). Thus, in order to test if *orfR2* and *orfR1* are respective functional homologs of *viscAR* and *viscBCR*, we complemented SBW25 Δ*viscAR* or Δ*viscBCR* mutants with low-copy plasmids carrying the respective orthologous CHA0 genes. Viscosin production (evidenced by the drop-collapse activity of culture supernatants and the swarming motility of the strains) was fully restored when each cognate ortholog transcriptional-regulator gene was present in a plasmid (Fig. 2D). Thus, *orfR2*

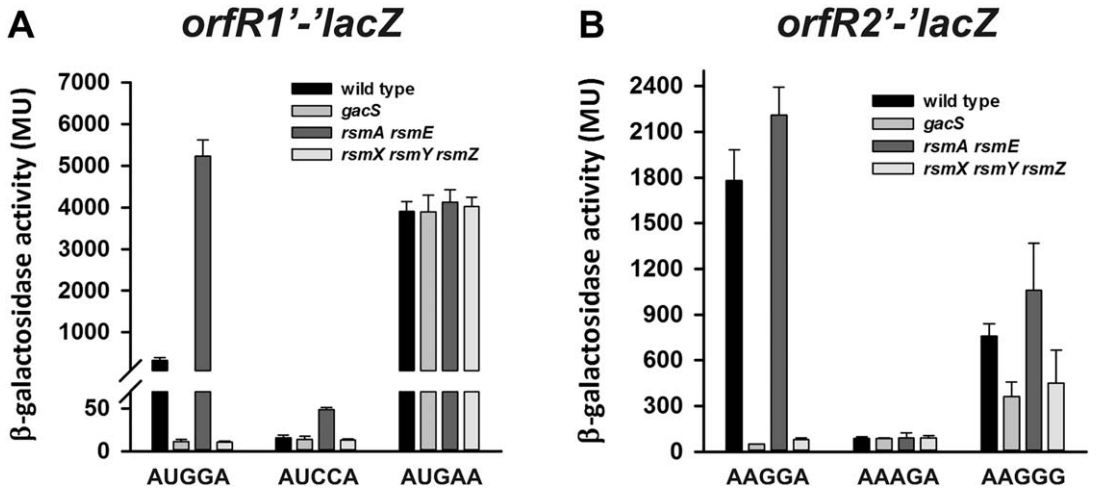



**Fig. 3.** Impact of point mutations within the putative RsmA/E-binding site on the Gac-Rsm control over *orfR1*'-'*lacZ* and *orfR2*'-'*lacZ* translational fusions. Comparison of the translational activity of wild-type and mutant *orfR1*'-'*lacZ* (A), or wild-type and mutant *orfR2*'-'*lacZ* (B), at the transition from exponential to early stationary phase (9 h after inoculation) in strains lacking different elements of the CHA0 Gac-Rsm cascade. In the figures, the ß-galactosidase activity is plotted on the *ordinate* for each of the genotypes indicated on the *abscissa*. Key to the genotypes denoted by the bar shades: black, wild-type; dark grey, *rsmA/E* mutant; medium grey, *gacS* mutant; light grey, *rsmX/Y/Z* mutant.

and *orfR1* proved to be functional LuxR-type transcriptional regulators of the *viscAR* and *viscBCR* families (Fig. 2D; Supporting Information Figs 3 and 4).

### *Evidence of direct control by RsmA/E over the 5′-UTR region of the* orfR1 *and* ofrR2 *genes*

To confirm that the sequence motifs identified within the 5′-UTR of *orfR1* and *orfR2* were responsible for the observed Gac-Rsm-dependent regulation (Fig. 1), we introduced point mutations in both translational fusions. For *orfR1*, the GG→CC replacement within the UGGA tetraloop completely abolished the expression of the fusion, most probably resulting from an ablation of the Shine–Dalgarno sequence (Fig. 3A). In contrast, a G→A replacement disrupting the original GGA motif within the tetraloop resulted in an active translational fusion with a full loss of Gac-Rsm regulation (Fig. 3A). For the second LuxR-transcriptional regulator gene *orfR2*, we obtained essentially similar results: a G→A replacement producing an AAGA tetraloop resulted in a translationally inactive fusion (Fig. 3B); the alternative A→G replacement to generate an AGGG tetraloop strongly reduced Gac-Rsm-regulatory expression (Fig. 3B). As with *orfR1*, the predicted binding site is essential for full Gac-Rsm control over *orfR2* expression; but, in contrast to *orfR1*, a residual dependence on Gac-Rsm in the mutant fusion bearing the AGGG tetraloop was observed, which relationship may be attributable to the presence of a second RsmA/E binding site in the vicinity of the start codon (i.e., AAGGA, at 6 nucleotides downstream from the identified motif). The relevance of this second neighbouring GGA for RsmA/E binding remains to be elucidated.

Finally, to garner independent evidence for a direct interaction of RsmA/E and the detected motifs within the *orfR1* and *orfR2* 5′-UTRs, we performed RNA-retardation assays using purified functional histidine-tagged versions of the repressor proteins RsmA and RsmE (Supporting Information Fig. 5), and synthetic RNA 20-mers spanning the putative binding sites. Both RsmA-$His_6$ and RsmE-$His_6$ shifted the wild type 20-mer of the *orfR1* 5′-UTR at protein:RNA molar ratios of at least 0.5 (RsmA-$His_6$) and 1.5 (RsmE-$His_6$; Fig. 4A and B). The single base replacement within the UGGA tetraloop that abolished *in vivo* Gac-Rsm control of the corresponding *orfR1*'-'*lacZ* fusion (Fig. 3A), impeded the binding of RsmA-$His_6$ or RsmE-$His_6$ (Fig. 4A and B). The wild-type 20-mer representing the RsmA/E target site within the *orfR2* 5′-UTR, also formed discrete complexes with RsmA-$His_6$ and RsmE-$His_6$ at protein:RNA molar ratios of at least 0.3 (RsmA-$His_6$) and 2.4 (RsmE-$His_6$; Fig. 4C).

All together, the results of the impact of the point mutations under investigation here on the Gac-Rsm control over translational fusions (Fig. 3) and the findings on the RNA-mobility-shift assays (Fig. 4), strongly support the notion that the identified sequence and/or structural motifs within the 5′-UTRs of *orfR1* and *orfR2* constitute the binding sites for the translational-repressor proteins RsmA and RsmE and are instrumental in activation of the translation of the mRNAs at the post-transcriptional level upon GacS/GacA signal transduction followed by the decoying of the molecular mimic sRNAs RsmX/Y/Z (Fig. 5). The involvement of the Gac-Rsm cascade in the control of CLP production in plant probiotic pseudomonads has been reported more than a decade ago in

*Pseudomonas* sp. strain DSS73 in relation to the production of amphisin (Koch *et al*., 2002) and more recently with respect to the synthesis of, at once, sclerosin in *Pseudomonas* sp. DF41 (Berry *et al*., 2012), the CLP white-line-inducing principle (WLIP) in *P. putida* RW10S2 (Rokni-Zadeh *et al*., 2012) and the CLP massetolide in *P. fluorescens* SS101 (Song *et al*., 2015). A direct post-transcriptional control was inferred upon inspection of the 5′-UTR of *massAR* (encoding the CHA0 *orfR2* ortholog) and detection of a putative RsmA/E-binding motif overlapping the RBS, but the second massetolide-regulator gene *massBCR* was not detected

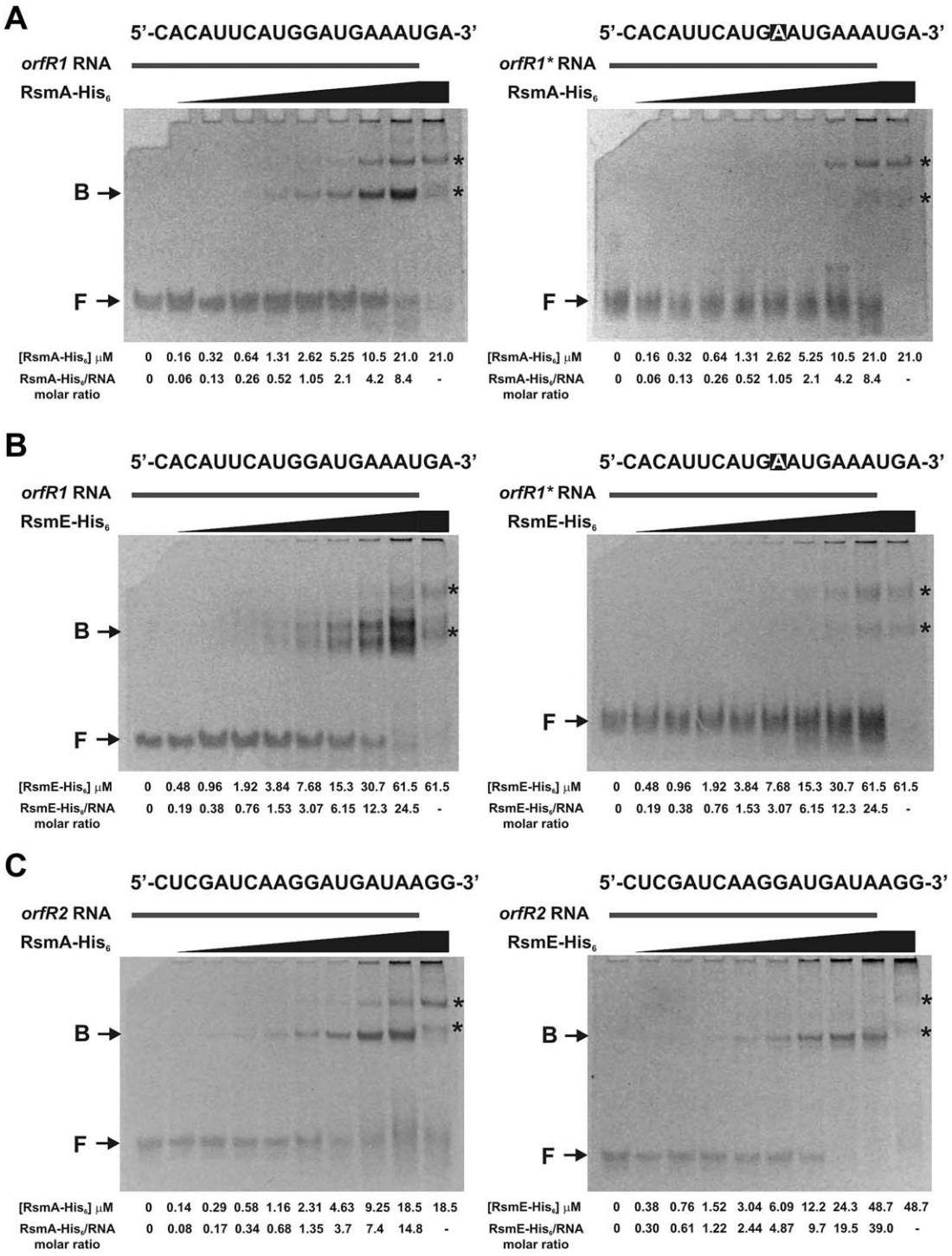


**Fig. 4.** Binding of histidine-tagged RsmA and RsmE proteins to RNA oligonucleotides encompassing the predicted target sites of the *orfR1* and *orfR2* 5′-UTRs. Different concentrations of RsmA-$His_6$ or RsmE-$His_6$ were incubated with each 20-mer RNA oligonucleotide shown above the panel before fractionation on nondenaturing gels, staining with GelRed, and imaging.
A. Wild-type (left panel) or mutant (right panel) *orfR1* motif versus RsmA-$His_6$.
B. Wild type (left panel) or mutant (right panel) *orfR1* motif versus RsmE-$His_6$. The replaced base is shaded in the mutant 20-mer.
C. Wild-type *orfR2* motif versus RsmA-$His_6$ (left panel) or versus RsmE-$His_6$ (right panel). F and B indicate the positions of the respective free- and bound-RNA species. The asterisks on the right side of the gel denote the migrating position of RNA species co-purified with the recombinant proteins loaded in the wells not containing the 20-mers. *P. protegens* CHA0 RsmA-$His_6$ or RsmE-$His_6$ were expressed in *E. coli*, and purified by Ni-NTA chromatography as described (Reimmann *et al*., 2005). The eluted proteins were dialyzed against 10 mM Tris.acetate pH 8.0, quantified by Bradford Coomassie-blue colorimetry with BSA as a standard, and stored at −20°C. The 20-nucleotide RNA fragments used in this study were purchased from Invitrogen Argentina SRL. The oligonucleotides were resuspended in diethyl-pyrocarbonate-treated water at 100 $\mu$M. The RNA fragments were preheated at 70°C for 5 min and incubated for a further 5 min at room temperature. Binding reactions (10 $\mu$l) contained 10 mM Tris (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl, 50 mM KCl and 5% (w/v) glycerol, the RNA oligonucleotides at 2.5 $\mu$M, and different concentrations of the RsmA or RsmE proteins. The mixtures were incubated for 20 min at 37°C for complex formation. Samples were loaded onto a 15% (w/v) PAGE. Electrophoresis was performed in Mini-PROTEAN III cells (Bio-Rad Laboratories) in 0.5× TBE buffer under nondenaturing conditions at 20 mA for 2 h at room temperature. After electrophoresis the gel was stained with the fluorescent dye GelRed (1:10 000 dilution in diethyl-pyrocarbonate-treated water) for 30 min and the bands were visualized in a conventional transilluminator.

as a putative RsmA/E target. Nevertheless, no experimental support for the direct control of RsmA/E over *massAR* has been reported (Song *et al*., 2015). In the *P. protegens* group, the requirement of a functional Gac-Rsm system for orfamide production has been recently documented in *P. protegens* Pf-5. For instance, a directed *gacA* inactivation resulted in a 64-fold reduction in orfamide-A production (Hassan *et al*., 2010), but also spontaneous mutations in *gacA* or *gacS* produced a decreased orfamide-A synthesis (Song *et al*., 2016). Both studies also reported a downregulation of the *orfR1* ortholog Pfl_2150 (Hassan *et al*., 2010; Song *et al*., 2016) and of the *orfR2* ortholog Pfl_2143 (Song *et al*., 2016) in the absence of a functional GacA protein, which control could be a consequence of an enhanced mRNA turnover following the translational arrest. Our results (Figs 3 and 4) provide the first evidence of a direct Gac-Rsm control over the genes encoding the two transcriptional regulators of CLP biosynthesis in *Pseudomonas*.

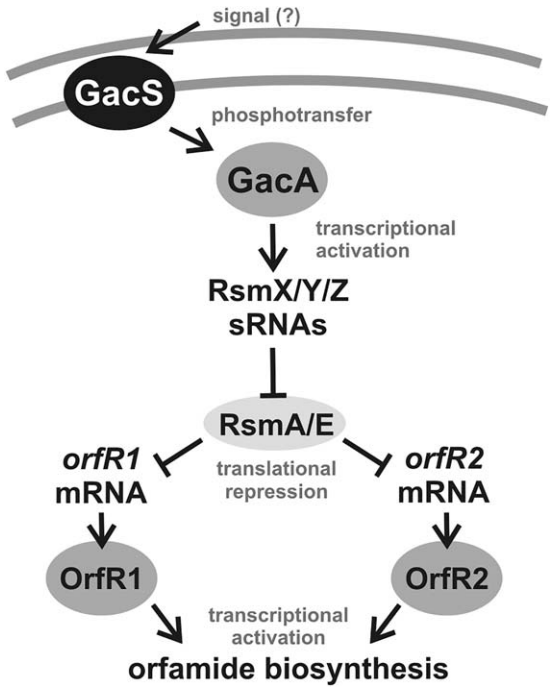



**Fig. 5.** Model of the hierarchical control by the Gac-Rsm cascade of the transcriptional regulators of the orfamide biosynthetic gene cluster in *P. protegens* CHA0.

## Conclusions and perspectives

In *P. protegens* CHA0, the post-transcriptional control of gene expression through the Gac-Rsm pathway has a major impact on secondary metabolism and, as a consequence, on several phenotypes linked to the ability of this soil bacterium to biocontrol different phytopathogens and regulate its interaction with eukaryotes (Haas and Defago, 2005). In the present work, we carried out an *in silico* search of novel genes and processes that may be subject to control by the Gac-Rsm cascade on the basis of sequence, structural and evolutionary constraints for the RsmA/E binding site (Supporting Information Fig. 1). The approach detected a set of 43 novel candidate targets of the Gac-Rsm pathway (Supporting Information Table 2). Indeed, two of the most prominent ones discovered in our *in silico* survey corresponded to a pair of LuxR-type transcriptional regulators (*orfR1* and *orfR2*) encoded in the immediate vicinity of the gene cluster for orfamide-A production (Fig. 1 and Supporting Information Fig. 3). From these results, we now provide solid evidence of a direct control of Gac-Rsm over the translation of both the *orfR1* and *orfR2* mRNAs (Figs (1 and 3) and 4), and consequently over orfamide production in *P. protegens* CHA0 (Fig. 2). That OrfR1 and OrfR2 were able to functionally replace the viscosin transcriptional regulators ViscAR and ViscBCR of *P. fluorescens* SBW25 (Fig. 2D) suggests the direct involvement of OrfR1 and OrfR2 in activating the expression of the orfamide-A biosynthetic cluster at the transcriptional level in strain CHA0. This possibility is currently under

study in our laboratory. Nevertheless, we have demonstrated thus far that orfamide production is another trait that is part of the Gac-Rsm regulon in *P. protegens* CHA0.

Several questions arise from the identification of these transcriptional regulators flanking the orfamide-biosynthetic gene cluster as Gac-Rsm targets. These queries include the issues regarding which signal(s) control the transcription of *orfR1* and *orfR2* and whether or not *orfR1* and *orfR2* are autoregulated like other Lux-type transcriptional elements (Chatterjee *et al*., 1996). Moreover, since OrfR1 and OrfR2 seem not to be functionally redundant (Fig. 2D), the further question arises as to whether those proteins have specific and independent binding sites in the promoter of the orfamide-biosynthetic operon and whether they form heterodimers. Further experiments are required to test all these hypotheses. Nevertheless, irrespective of the transcriptional-regulatory pattern of *orfR1* and *orfR2*, our data clearly show that expression of these alleles is subject to control by the Gac-Rsm pathway at the translational level. Thus, the activation of orfamide production is coordinated with the induction of several other biosynthetic pathways producing secondary metabolites in *P. protegens* CHA0 following stimulation of the Gac-Rsm cascade (Fig. 5). Finally, the regulation of orfamide production in *P. protegens* CHA0 represents an instance of hierarchical post-transcriptional control over transcriptional regulators, which example is added to a list of reported cases of regulatory networks already documented involving small noncoding RNAs that directly control transcription factors – for example, LuxO, LuxR and AphA in *Vibrio* spp. (Feng *et al*., 2015) and FlhDC and CsgD in Enterobacteriaceae (Mika and Hengge, 2013). To our knowledge, this work represents the first demonstration of a direct post-transcriptional control by the Gac-Rsm cascade over a pair of transcriptional regulators in *Pseudomonas* spp.

## Acknowledgements

This work was supported by grants from CONICET, ANPCyT and Universidad Nacional de Quilmes. AM was supported by a CONICET fellowship. CV and EJ are research members of CONICET. The authors wish to thank Jos Raaijmakers for providing strains SBW25, SBW25 Δ*viscAR* and SBW25 Δ*viscBCR*, and Pia Valacco and Silvia Moreno (CEQUIBIEM, Universidad Nacional de Buenos Aires, Argentina) for their technical assistance with the mass spectrometer. The authors declare no conflicts of interest. Finally, Dr. Donald F. Haggerty, a retired academic career investigator and native English speaker, edited the final version of the manuscript.

## References

Alvarez, F., Castro, M., Principe, A., Borioli, G., Fischer, S., Mori, G., and Jofre, E. (2012) The plant-associated *Bacillus amyloliquefaciens* strains MEP2 18 and ARP2 3 capable of producing the cyclic lipopeptides iturin or surfactin and fengycin are effective in biocontrol of sclerotinia stem rot disease. *J Appl Microbiol* **112**: 159–174.

Berry, C.L., Brassinga, A.K., Donald, L.J., Fernando, W.G., Loewen, P.C., and de Kievit, T.R. (2012) Chemical and biological characterization of sclerosin, an antifungal lipopeptide. *Can J Microbiol* **58**: 1027–1034.

Blumer, C., Heeb, S., Pessi, G., and Haas, D. (1999) Global GacA-steered control of cyanide and exoprotease production in *Pseudomonas fluorescens* involves specific ribosome binding sites. *Proc Natl Acad Sci USA* **96**: 14073–14078.

de Bruijn, I., and Raaijmakers, J.M. (2009) Diversity and functional analysis of LuxR-type transcriptional regulators of cyclic lipopeptide biosynthesis in *Pseudomonas fluorescens*. *Appl Environ Microbiol* **75**: 4753–4761.

Chatterjee, J., Miyamoto, C.M., and Meighen, E.A. (1996) Autoregulation of luxR: the *Vibrio harveyi* lux-operon activator functions as a repressor. *Mol Microbiol* **20**: 415–425.

Crespo, M.C., and Valverde, C. (2009) A single mutation in the oprF mRNA leader confers strict translational control by the Gac/Rsm system in *Pseudomonas fluorescens* CHA0. *Curr Microbiol* **58**: 182–188.

D'Aes, J., De Maeyer, K., Pauwelyn, E., and Hofte, M. (2010) Biosurfactants in plant-*Pseudomonas* interactions and their importance to biocontrol. *Environ Microbiol Rep* **2**: 359–372.

De La Fuente, L., Thomashow, L., Weller, D., Bajsa, N., Quagliotto, L., Chernin, L., and Arias, A. (2004) *Pseudomonas fluorescens* UP61 isolated from birdsfoot trefoil rhizosphere produces multiple antibiotics and exerts a broad spectrum of biocontrol activity. *Eur J Plant Pathol* **110**: 671–681.

Feng, L., Rutherford, S.T., Papenfort, K., Bagert, J.D., van Kessel, J.C., Tirrell, D.A., *et al*. (2015) A qrr noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* **160**: 228–240.

Flury, P., Vesga, P., Pechy-Tarr, M., Aellen, N., Dennert, F., Hofer, N., *et al*. (2017) Antimicrobial and insecticidal: cyclic lipopeptides and hydrogen cyanide produced by plant-beneficial *Pseudomonas* strains CHA0, CMR12a, and PCL1391 contribute to insect killing. *Front Microbiol* **8**: 100.

Gross, H., Stockwell, V.O., Henkels, M.D., Nowak-Thompson, B., Loper, J.E., and Gerwick, W.H. (2007) The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. *Chem Biol* **14**: 53–63.

Haas, D., and Defago, G. (2005) Biological control of soilborne pathogens by fluorescent pseudomonads. *Nat Rev Microbiol* **3**: 307–319.

Hassan, K.A., Johnson, A., Shaffer, B.T., Ren, Q., Kidarsa, T.A., Elbourne, L.D., *et al*. (2010) Inactivation of the GacA response regulator in *Pseudomonas fluorescens* Pf-5 has far-reaching transcriptomic consequences. *Environ Microbiol* **12**: 899–915.

Heeb, S., Blumer, C., and Haas, D. (2002) Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0. *J Bacteriol* **184**: 1046–1056.

Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., *et al*. (2016) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *EMBO J* **35**: 991–1011.

Jousset, A., Lara, E., Wall, L.G., and Valverde, C. (2006) Secondary metabolites help biocontrol strain *Pseudomonas fluorescens* CHA0 to escape protozoan grazing. *Appl Environ Microbiol* **72**: 7083–7090.

Jousset, A., Schuldes, J., Keel, C., Maurhofer, M., Daniel, R., Scheu, S., and Thuermer, A. (2014) Full-genome sequence of the plant growth-promoting bacterium *Pseudomonas protegens* CHA0. *Genome Announc* **2**. pii: e00322-14.

Kay, E., Dubuis, C., and Haas, D. (2005) Three small RNAs jointly ensure secondary metabolism and biocontrol in *Pseudomonas fluorescens* CHA0. *Proc Natl Acad Sci USA* **102**: 17136–17141.

Koch, B., Nielsen, T.H., Sorensen, D., Andersen, J.B., Christophersen, C., Molin, S., *et al*. (2002) Lipopeptide production in *Pseudomonas* sp. strain DSS73 is regulated by components of sugar beet seed exudate via the Gac two-component regulatory system. *Appl Environ Microbiol* **68**: 4509–4516.

Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**: 255–258.

Lapouge, K., Sineva, E., Lindell, M., Starke, K., Baker, C.S., Babitzke, P., and Haas, D. (2007) Mechanism of hcnA mRNA recognition in the Gac/Rsm signal transduction pathway of *Pseudomonas fluorescens*. *Mol Microbiol* **66**: 341–356.

Lapouge, K., Schubert, M., Allain, F.H., and Haas, D. (2008) Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol Microbiol* **67**: 241–253.

Lapouge, K., Perozzo, R., Iwaszkiewicz, J., Bertelli, C., Zoete, V., Michielin, O., *et al*. (2013) RNA pentaloop structures as effective targets of regulators belonging to the RsmA/CsrA protein family. *RNA Biol* **10**: 1031–1041.

Ma, Z., Geudens, N., Kieu, N.P., Sinnaeve, D., Ongena, M., Martins, J.C., and Hofte, M. (2016) Biosynthesis, chemical structure, and structure-activity relationship of orfamide lipopeptides produced by *Pseudomonas protegens* and related species. *Front Microbiol* **7**: 382.

Mika, F., and Hengge, R. (2013) Small regulatory RNAs in the control of motility and biofilm formation in *E. coli* and *Salmonella*. *Int J Mol Sci* **14**: 4560–4579.

Mortier-Barriere, I., Velten, M., Dupaigne, P., Mirouze, N., Pietrement, O., McGovern, S., *et al*. (2007) A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* **130**: 824–836.

Murray, H., and Errington, J. (2008) Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell* **135**: 74–84.

Ongena, M., and Jacques, P. (2008) *Bacillus* lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol* **16**: 115–125.

Raaijmakers, J.M., De Bruijn, I., Nybroe, O., and Ongena, M. (2010) Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol Rev* **34**: 1037–1062.

Reimmann, C., Valverde, C., Kay, E., and Haas, D. (2005) Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0. *J Bacteriol* **187**: 276–285.

Rokni-Zadeh, H., Li, W., Sanchez-Rodriguez, A., Sinnaeve, D., Rozenski, J., Martins, J.C., and De Mot, R. (2012) Genetic and functional characterization of cyclic lipopeptide white-line-inducing principle (WLIP) production by rice rhizosphere isolate *Pseudomonas putida* RW10S2. *Appl Environ Microbiol* **78**: 4826–4834.

Romanowski, A., Migliori, M.L., Valverde, C., and Golombek, D.A. (2011) Circadian variation in *Pseudomonas fluorescens* (CHA0)-mediated paralysis of *Caenorhabditis elegans*. *Microb Pathog* **50**: 23–30.

Schubert, M., Lapouge, K., Duss, O., Oberstrass, F.C., Jelesarov, I., Haas, D., and Allain, F.H. (2007) Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nat Struct Mol Biol* **14**: 807–813.

Song, C., van der Voort, M., van de Mortel, J., Hassan, K.A., Elbourne, L.D., Paulsen, I.T., *et al*. (2015) The Rsm regulon of plant growth-promoting *Pseudomonas fluorescens* SS101: role of small RNAs in regulation of lipopeptide biosynthesis. *Microb Biotechnol* **8**: 296–310.

Song, C., Kidarsa, T.A., van de Mortel, J.E., Loper, J.E., and Raaijmakers, J.M. (2016) Living on the edge: emergence of spontaneous gac mutations in *Pseudomonas protegens* during swarming motility. *Environ Microbiol* **18**: 3453–3465.

Valverde, C., and Haas, D. (2008) Small RNAs controlled by two-component systems. *Adv Exp Med Biol* **631**: 54–79.

Valverde, C., Lindell, M., Wagner, E.G., and Haas, D. (2004) A repeated GGA motif is critical for the activity and stability of the riboregulator RsmY of *Pseudomonas fluorescens*. *J Biol Chem* **279**: 25066–25074.

Wei, X., Huang, X., Tang, L., Wu, D., and Xu, Y. (2013) Global control of GacA in secondary metabolism, primary metabolism, secretion systems, and motility in the rhizobacterium *Pseudomonas aeruginosa* M18. *J Bacteriol* **195**: 3387–3400.

Will, S., Joshi, T., Hofacker, I.L., Stadler, P.F., and Backofen, R. (2012) LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA* **18**: 900–914.

Zuber, S., Carruthers, F., Keel, C., Mattart, A., Blumer, C., Pessi, G. *et al*, (2003) GacS sensor domains pertinent to the regulation of exoproduct formation and to the biocontrol potential of *Pseudomonas fluorescens* CHA0. *Mol Plant Microbe Interact* **16**: 634–644.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Supplementary Table 1**. List of 5′-UTRs (-250 to +50 relative to the annotated start codon) from the genome of

*P. protegens* CHA0 that contain sequence motifs matching the query 5′-WCANGGANGW-3′ (Schubert *et al*., 2007).

**Supplementary Table 2**. Predicted RsmA/RsmE binding sites in the 5′-UTR of annotated genes in the chromosome of *P. protegens* CHA0.

**Supplementary Table 3**. Strains, plasmids and oligonucleotides used in this study. *Pseudomonas* spp. and *Escherichia coli* were grown on nutrient agar (NA; blood agar base 40 g/l, yeast extract 5 g/l) and in nutrient yeast broth (NYB; nutrient broth 25 g/l, yeast extract 5 g/l) at 28°C and 37°C, respectively. When required, antibiotics were added to the medium in the following concentrations: 125 μg/ml tetracycline for *Pseudomonas protegens*; 25 μg/ml tetracycline, 100 μg/ml ampicillin for *E. coli*. *P. protegens* strains were grown at 35°C to improve their capacity to accept heterologous DNA in electrotransformation with plasmids (Valverde *et al*., 2004).

**Supp. Fig. 1**. Panel a) Schematic representation of the *in silico* workflow utilized to predict novel mRNAs that are direct targets of the Gac/Rsm cascade in the genome of *P. protegens* CHA0. The procedure stated in brief: the 5′-untranslated regions (UTRs) of all annotated genes were recruited from the CHA0 genome, and the string search WCANGGANGW was carried out with the DNA-pattern tool from the RSAT server (Medina-Rivera *et al*., 2015). The RNA secondary prediction was done with the mFOLD server (Zuker, 2003). Sequence conservation of the candidate motifs was conducted by obtaining the orthologous genomic segments from other *Pseudomonas* spp. genomes (www.pseudomonas.org, (Winsor *et al*., 2016)). Multisequence alignments were carried out through the use of CLUSTAL OMEGA (Sievers *et al*., 2011). Panel b) Distribution frequency of RsmA/E binding motif locations relative to mRNA regions. c) Distribution frequency of pentanucleotides within the central portion of the identified motifs. Panel d) Location of the predicted RsmA/E binding sites (solid circles) and of the corresponding RBSs within the 5′-UTRs of the identified mRNA targets (open circles). The relative position on the *abscissa* is referred to the first nucleotide of the corresponding coding sequence (CDS). The Shine-Dalgarno sequence of every open reading frame (ORF) of *P. protegens* CHA0 was obtained by retrieving 30 nucleotides from −10 to +20 relative to the first nucleotide of the annotated start codon through the RSAT server (Medina-Rivera *et al*., 2015). Then, sequences were queried for the presence of the core consensus sequence string 5′-GGAGG-3′ (Shine and Dalgarno, 1975), with up to one mismatch permitted, through the use of the DNA-pattern tool from the RSAT server (Medina-Rivera *et al*., 2015).

**Supp. Fig. 2**. Experimental validation of Gac/Rsm translational control of selected putative target mRNA genes. The impact of the *gacS*, *rsmA/rsmE*, and *rsmX/Y/Z* mutations on the expression of (**a**) *PFLCHA0_c00250* (*dprA*), (**b**) *PFL CHA0_c02690* (*pksP*), (**c**) *PFL CHA0_c04880* (*soj1*) and (**d**) *PFL CHA0_c21910* (*orfR1*). The candidate target PFLCHA0_c00250 -automatically annotated as *smf*, or *dprA*, for DNA-processing chain A- is a broadly conserved protein family with the ability to bind to single-stranded DNA to protect the ligated stretch from degradation (Mortier-Barriere *et al*., 2007) and it has been implicated in the process of natural DNA transformation in different species; such as *Haemophilus influenzae* (Karudapuram *et al*., 1995), *Streptococcus pneumoniae* (Berge *et al*., 2003) and *Bacillus subtilis* (Tadesse and Graumann, 2007). The candidate target gene *PFLCHA0_c02690* (*pksP*) encodes a small putative acyl carrier protein (*Pseudomonas* Orthologous Group POG019731), and is located within a region containing gene functions related to fatty acid modifications; *pksP* homologues are apparently restricted to a few pseudomonad genomes including the *P. protegens* cluster (Winsor *et al*., 2016). The candidate target *PFLCHA0_c04880* (*soj1*) is involved in the cell-division process (Murray and Errington, 2008). β-Galactosidase expression of plasmid-borne translational '*lacZ* fusions was determined in strains CHA0 (in MU on the left *ordinates*; open circles, wild-type context), CHA19 (open triangles, *gacS* mutant), CHA1144 (open diamonds, *rsmXYZ* mutant) and CHA1009 (open squares, *rsmAE* mutant). The corresponding solid symbols represent the growth curves of each strain as plotted on the right *ordinates* in an exponential scale and expressed as the optical density at 600 nm ($OD_{600}$). Each value is the average of the data from three different cultures ± standard deviation. The images above each panel show, to the left, the genomic context of each target gene in the CHA0 genome, and to the right, the evolutionary conservation of the 5′-UTR-RNA sequence and its secondary structure as inferred by locARN (Will *et al*., 2012). See the legend to Fig. 1 for the explanation of the secondary structure conservation analysis. The Mfold prediction of the identified RsmA/E binding motif with the corresponding standard free energy indicated below the loop structures (Zuker, 2003) is shown to the right of each panel.

**Suppl. Fig. 3.** Synteny of the CLP biosynthetic locus in related pseudomonads. Genes involved in the biosynthesis of the correspondingly CLPs are indicated with solid black arrows, whereas loci encoding predicted or experimentally verified transcriptional regulators of the biosynthetic genes are indicated with open white arrows. The upper grey box encloses CLP gene clusters encoded within a single locus, whereas the lower box encloses genes that are encoded by two unlinked loci.

**Suppl. Fig. 4**. Evolutionary relationships of the correspondingly dedicated LuxR-type transcriptional regulators. The evolutionary history was inferred through the use of the Maximum Likelihood method based on the JTT matrix-based model (Jones *et al*., 1992). The tree with the highest log likelihood (-5752.9238) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying neighbour-joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The analysis involved 30 amino acid sequences from the following *Pseudomonas* strains: *P. fluorescens* Pf-5 (Pfl_2143 and Pfl_2150, GenBank accessions YP_259250 and YP_259257), *P. syringae* pv. *syringae* B64 (SalA, GenBank accession ELS42861.1), *P. fluorescens* SBW25 (ViscAR and ViscBCR, GenBank accessions CAY50418 and CAY48792), *P. syringae* pv *tomato* DC3000 (Pspto2833 and SyfR, GenBank accessions NP_792637.1 and NP_792632.1), *P. fluorescens* Pf-01 (Pf01_2216 and Pf01_2210, GenBank accessions ABA73959 and ABA73953), *P. fluorescens* SS101 (MassAR and MassBCR,

GenBank accessions Q0PH96 and EIK62142), *P. putida* PCL1445 (PsoR and LuxR-like protein PCL1445, GenBank accession DQ151887), *P. sp*. MIS38 (Orf1, GenBank accession AB107223), *P. cichorii* JBC1 (CichofactinR1 and CichofactinR2, GenBank accessions BH81_RS12990 and BH81_RS13015), *P. fluorescens* BW11P2 (BanG and BanH, GenBank accession KX437753), *P. poae* RE*1-1-14 (PoaAR and PoaBCR, GenBank accessions H045_10955 and H045_07055), *P. entomophila* L48 (EtlR, GenBank accession PSEEN3335), *P. putida* BW11M1 (XtlR, GenBank accession KC297505), *P. putida* RW10S2 (WlpR, GenBank accession KC297505). Other LuxR-type regulators included in the analysis are *Vibrio fischeri* LuxR (GenBank accession AAQ90196), *P. aeruginosa* PAO1 LasR and RhlR (GenBank accessions BAA06489 and NP_252167), *P. putida* PpuR (GenBank accession AAZ80478) and the full-length ATP/dATP-binding protein MalT of *E. coli* (GenBank accession AAA83888). All positions containing gaps and missing data were eliminated. There were a total of 155 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar *et al*., 2016).

**Suppl. Fig. 5**. The C-terminal histidine-tagged RsmA and RsmE derivatives are functional repressors of Gac-dependent genes in *P. protegens* CHA0. The functionality of the His-tagged version of both repressor proteins was tested in different phonotypical assays: (**a**) qualitative exoprotease activity was detected on skim-milk nutrient agar (Valverde *et al*., 2003); (**b**) HCN production (Voisard *et al*., 1989); β-galactosidase expression of chromosomal *hcnA*'-'*lacZ* (**c**) and *aprA*'-'*lacZ* (**d**) fusions (Reimmann *et al*., 2005); (**e**) stabilization of the sRNA RsmY upon induction of each RNA binding protein, as revealed by Northern blotting (Valverde *et al*., 2004). (**f**) co-purification of the sRNA RsmY with the His-tagged repressor proteins during Ni-affinity purification from the Δ*rsmAE* strain CHA1009, as revealed by Northern blotting of phenol extracts (Reimmann *et al*., 2005). In Panel a, isopropyl ß-D-thiogalactoside (IPTG) was added to the plate when appropriate. In Panel b, the concentration of HCN generated (μM) per unit of $OD_{600}$ is plotted on the *ordinate* for each of the strains grown on nutrient yeast broth (NYB) indicated on the *abscissa* either in the presence (grey bars) or absence (white bars) of IPTG. In panels c and d, ß-galactosidase activity in MU is plotted on the *ordinates* for each of the strains indicated on the *abscissas*. Panel e depicts the cellular level of RsmY sRNA at the time denoted after His-tagged RsmA/E induction as visualized by the Northern-blot density. Panel f shows the amount of RsmY sRNA co-purified with the His-tagged RsmA/E for the strains indicated above each lane.