Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

Article

Complete Labelling of Pneumococcal DNA-Binding Proteins with Seleno-L-methionine

Fabián Lorenzo-Díaz[1], Inmaculada Moreno-Córdoba[2#], and Manuel Espinosa[2*]

[1]Departamento de Bioquímica, Microbiología, Biología Celular y Genética, Universidad de La Laguna, Santa Cruz de Tenerife, Spain; email: florenzo@ull.edu.es

[2]Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu, 9, 28040 Madrid, Spain; emails: inmamoreno395@gmail.com; mespinosa@cib.csic.es

*Correspondence: mespinosa@cib.csic.es

#Present address: Department of Global Health, GlaxoSmithKline, Severo Ochoa, 2, 28760-Tres Cantos, Madrid, Spain.

Received: date; Accepted: date; Published: date

Abstract: *Streptococcus pneumoniae* is an pathogenic and opportunistic Gram-positive bacteria that is the leading cause of community acquired respiratory diseases, varying from mild- to deathly-infections. Appearance of antibiotic resistant isolates has prompted the search for novel targets. One of the most promising approaches is the structure-based knowledge of possible targets in conjunction to rational design and docking of inhibitors of the chosen targets. A useful technique to help solving protein structures is to label them with a heavy atom, like selenium, that facilitates tracing of the some of the amino acid residues. We have chosen two pneumococcal DNA-binding proteins, namely the relaxase domain of MobM protein from plasmid pMV158, and the RelB-RelE antitoxin-toxin protein complex. Through the update of a previous protocol [1] that uses seleno-L-methionine, we could achieve 100% labelling of the proteins. Furthermore, the labelled proteins retained full activity as judged from relaxation of supercoiled plasmid DNA and from gel-retardation assays.

(*153 words*)

Keywords: *Streptococcus pneumoniae*; protein purification; protein labelling; seleno-methionine; DNA-protein interactions

Abbreviations: COINs (Conjugation inhibitors), DTT (dithiothreitol), EMM (enriched minimal medium), EMSA (electrophoretic mobility shift assays), HGT (horizontal gene transfer), ICEs (integrative and conjugative elements), IPTG (isopropyl β-D-1-thiogalactopyranoside), MALDI Matrix-assisted laser desorption/ionization, OD (optical density), PAA (polyacrylamide), PAGE (polyacrylamide gel electrophoresis), RNAP (RNA polymerase), SeMet (seleno-methionine), SDS (sodium dodecyl sulphate), TAS (toxin-antitoxin systems), TOF (time of flying)

1. Introduction

© 2019 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

Streptococcus pneumoniae (the pneumococcus) is a Gram-positive pathogenic bacterium responsible for the death of nearly about 1.5 million humans per year. The bacterium is especially lethal for children below 5 years, elderly people, and immuno-compromised patients [2]. In addition to the deadly community-acquired pneumococcal pneumonia, the pneumococcus is the causal agent of a number of milder diseases, like meningitis, otitis media and sepsis, all of them requiring medium- to short-sick leaves with the concomitant loss of working hours and economic burden. Management of pneumococcal infections takes into account the vaccination programs, although these have led to selection of the serotypes for which vaccination is not yet available [3]. Antibiotic treatments for pneumococcal and other lethal bacterial pathogens have resulted in the selection of bacteria with high levels of resistance (the so-called 'superbugs') to which there is an urgency to finding treatments alternative to the classical use (and sometimes abuse) of antibiotics [4]. Thus, it has been proposed that tackling pneumococcal diseases should consider approaches different than antibiotic treatments [5]. Among the several possible strategies, two of them seem to be relevant because there are some examples that can be taken as proofs of principle amenable to further advances: i) inhibitors of bacterial conjugation (COINs), and ii) bacterial Toxin-Antitoxin systems (TAS) [5,6].

The first approach, employment of COINs, contemplates the use of molecules that are able to inhibit horizontal gene transfer (HGT) of genetic elements conferring antibiotic resistance, like plasmids, Integrative and Conjugative Elements (ICEs) and other mobile elements. Among inhibitors of HGT, it has been shown that unsaturated fatty acids act as effective COINs molecules [7,8]. Inhibition seems to be due to interference with the ATPase activity of the VirB11-type proteins, which are integrants of the Type IV Secretion System [9].

The second approach contemplates strategies like the exploitation of the toxin proteins from the bacterial TAS as targets for drug discovery, especially of type II (proteic) family. Pioneer work in Gerdes' laboratory showed that TAS are present in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

the majority of prokaryotes, but they are absent in eukaryotes [10], and because of this, they are considered as attractive molecules to be used as antibacterials [11-13]. In general, type II TAS are organized as operons of two genes, the antitoxin preceding the toxin genes; whereas the toxin is a stable protein, the antitoxin is a labile protein mostly because of its unfolded nature [14,15]. During bacterial growth under steady-state conditions, the toxin-antitoxin proteins generate a harmless self-regulated complex. However, under stressful circumstances, the TAS are triggered and the antitoxin, exposed to the environment, would be prone to degradation by proteases, releasing the toxin to act as a poison to halt cell growth, a circumstance that has been proposed to exploit as an antibacterial strategy [16]. The most promising approach to the use of TAS as antibacterials is based on the finding that short peptides can disrupt the rather strong interactions between the toxin and the antitoxin, making it possible to use these peptides as true antibacterials [17,18]. However, this approach requires that the three-dimensional structure of the TA protein-protein complex is known in order to design the proper peptides.

Discovery of novel antibacterials nowadays is less focussed in testing large amounts of compounds or libraries than it used to be, and now drug discovery is helped by the use of known structures of the target. Thus, molecular docking and structure-based designs are the favourite methods of choice, when feasible, in modern drug discovery [19,20]. In the case of the bacterium *S. pneumoniae* and its mobile elements, not very many protein structures with a known function have been determined so far and they include plasmid-encoded proteins [21-24] as well as chromosomally-encoded ones [25-27]. In most, if not all, cases solution of crystal structures has been aided by the employment of protein derivatives, especially those labelled with the amino acid derivative Seleno-L-methionine (SeMet) because it provides a heavy metal that facilitates the tracing of the amino acids in the crystal. In the present work, we present a modified protocol from the early seminal work from Huber's laboratory [1], the use of which has allowed us to attain complete (100%) labelling of pneumococcal DNA-binding proteins with SeMet, namely the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

98 untagged MobMN199 from plasmid pMV158 [28] and the His-tagged RelBE protein
99 complex from the pneumococcal TAs [29]. Further, we show that the SeMet-labelled
100 proteins retained full activity as determined by plasmid DNA relaxation and
101 band-shift experiments.

102

103 ## 2. Materials and Methods

104 *2.1. Bacterial Strains, Plasmids, and DNA Manipulations*

105 Strains and plasmids used are listed in Table 1.

106

107 **Table 1:** Bacterial strains and plasmids used

Bacterial strain	Genotype	Plasmid or plasmid construct	Source of	References
Escherichia coli B834(DE3)	λ DE3 (*lac*I *lac*UV5-T7 gene 1, *ind*1, *sam*7, *nin*5) F^- ,*dcm*, *lon*, *omp*T, *hsd*S($r_B^-m_B^+$) *gal*, *met*	pMobMN199 (pET24b::*mobMN199*)	MobMN199	[24,28,30]
E. coli B834(DE3)	λ DE3 (*lac*I *lac*UV5-T7 gene 1, *ind*1, *sam*7, *nin*5) F^- ,*dcm*, *lon*, *omp*T, *hsd*S($r_B^-m_B^+$) *gal*, *met*	pET28relBE (pET28a::*relBE*)	$RelBEHis_6$	[30,31]
S. pneumoniae R6	Wild type	none	Genomic DNA	[32]
S. pneumoniae R6	Wild type	pMV158	Plasmid DNA	[33,34]

108

109 *Escherichia coli* bacterial strains were based in the DE3 lambda lysogen
110 developed in Studier's laboratory [30,35] in which the RNA polymerase (RNAP) of
111 bacteriophage T7 is cloned under the control of the *lacUV5* promoter, inducible by
112 isopropyl β-D-1-thiogalactopyranoside (IPTG), whereas the desired gene(s) is
113 cloned in a plasmid under the control of the Φ10 promoter of phage T7. Induction
114 with IPTG triggers the phage RNAP synthesis and, in turn, synthesis of the mRNA
115 of the desired gene is also triggered. T7 RNAP only recognizes its own promoters
116 and is insensitive to rifampicin; thus, addition of the drug will stop transcription
117 from any of the *E. coli* promoters, and only the T7 promoters will be active [30].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

The *S. pneumoniae* strain R6 is a noncapsulated derivative from the virulent strain D39 [32], and was used either for: i) isolation of total chromosomal DNA and PCR-amplification of a DNA region encompassing the *relBE* operon to clone it in the expression vector [29], or ii) amplification of a 256-bp DNA fragment that includes the promoter region, the transcription initiation site, and the Shine-Dalgarno sequences of the *relBE* operon, when functional assays were performed. The same pneumococcal strain harbouring plasmid pMV158 was used to prepare purified plasmid DNA by two consecutive CsCl gradients as reported [34]. Design and cloning the truncated gene *mobMN199* encoding the relaxase domain of the pMV158-encoded *mobM* gene have been detailed elsewhere [28]. Similarly, cloning of the His-tagged pneumococcal *relBE* operon (formerly termed *relBE2Spn*) has been reported previously [31].

2.2. Bacterial Growth Conditions

Pneumococcal cells were grown in the semi-defined medium AGCH [36,37] supplemented with sucrose, until middle exponential phase (2-3 x 10^8 cells/ml). Cells were collected by centrifugation and treated as reported to prepare either total chromosomal DNA [38] or purified plasmid DNA. *E. coli* cells were grown in M9 medium that was prepared 10x [39] and that was enriched with vitamin (riboflavin, biotin, thiamine and pyridoxine at 10 μg/ml, final concentration) and amino acids mixtures (except methionine) at 40 μg/ml, final concentration. Glucose (20 mM, final concentration) was the carbon source used. When needed, methionine (50 μg/ml, final concentration) was added to this enriched minimal medium (EMM+Met). To prepare uniform pre-inocula, cells harbouring plasmids were inoculated into 25 ml of EMM-Met, incubated at 37ºC until mid exponential phase (~5x10^7 cells/ml) and centrifuged, washed with the same medium and concentrated ten times. Glycerol (10%) was added and, after further 10 min incubation at 37ºC, 20 μl aliquots were prepared and stored at -80ºC until further use. Growth was followed by determination of the optical density at 600 nm (OD_{600}) of the cultures.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

147

148 *2.3. Labelling Proteins with SeMet and Protein Purification*

149 Pre-inocula (10 µl) from the frozen samples (above) were inoculated into 5 ml fresh
150 EMM+Met and incubated at 37ºC until OD_{600} reached ~0.9. Four ml of this culture
151 was then diluted into 400 ml of the same pre-warmed medium and incubation
152 resumed until these cultures reached the same OD_{600} (around 4.5 h). Then, cells
153 were collected by centrifugation, washed twice with EMM lacking methionine and
154 resuspended into the original amount (400 ml) of the same medium. To deplete the
155 pool of internal methionine, cells were further incubated 30 min, 37ºC. Then, the
156 culture was added to 3.6 l of pre-warmed EMM+SeMet (50 µg/ml), separated into
157 750-ml portions and incubation resumed at 37ºC. When the OD_{600} reached ~0.8,
158 IPTG (final concentration 1 mM) was added and incubation resumed 2 h more.
159 Then, rifampicin (dissolved in dimethylsulfoxide), final concentration 100 µg/ml
160 was added, and incubation continued 1.5 h more. Cells were harvested by
161 centrifugation and washed with buffer A (MobMN199) or buffer C (RelB-RelE)
162 (Table 2). Purification protocols for the untagged MobMN199 [28] and of the
163 His-tagged RelB-RelE$(His)_6$ complex [31] proteins are described in the Results
164 Section.

165 **Table 2:** Buffer compositions

Buffer	Composition	Use
A	20 mM Tris–HCl pH 7.6, 1mM EDTA, 1mM dithiothreitol, 5% glycerol, 300 mM NaCl	MobMN199 purification (Affinity chromatography)
C	20 mM Tris pH 8.0, 5% ethylene glycol, 1 mM β-mercaptoethanol, 10 mM imidazole, 500 mM NaCl	RelB-RelE purification (IMAC)
S	20 mM Tris pH 7.6, 1 mM EDTA, 5% ethylene glycol, 1 mM DTT, 500 mM NaCl	Gel-filtration chromatography
M	20 mM Tris pH 8.0, 1 mM EDTA, 1 mM DTT, 10% glycerol, 10 mg/ml heparin, 50 mM NaCl	EMSA

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

166

167 *2.4. Mass Spectra Analyses*

168 Matrix-assisted laser desorption/ionization (MALDI) experiments were performed
169 on an Autoflex III MALDI-TOF-TOF instrument (Bruker Daltonics, Bremen,
170 Germany) equipped with a smart-beam laser to measure the time of flying (TOF).
171 The spectra were acquired using a laser power just above the ionization threshold.
172 Samples were analyzed in the positive ion detection and delayed extraction linear
173 mode. Typically, 1,000 laser shots were summed into a single mass spectrum.
174 External calibration was performed using standard proteins (Sigma) namely insulin
175 (5.8 kDa), cytochrome C (12.4 kDa), trypsinogen (23.9 kDa), carbonic anhydrase (29
176 kDa), and protein A (44.6 kDa). The spectra covered the range from 2,000 to 30,000
177 Da.

178

179 *2.5. Functional Analyses of SeMet Proteins*

180 Relaxation assays with unlabelled and SeMet-labelled MobMN199 were performed
181 by incubation of supercoiled purified pMV158 plasmid DNA (300 ng, 8 nM) with
182 480 nM protein in buffer A (Table 2), to which 15 mM $MnCl_2$ was added. Incubation
183 was at 30ºC, 20 min, and the reaction products were separated by electrophoresis on
184 1% agarose gels as reported [28,40]. Cleavage of supercoiled (forms FI) by the
185 proteins generated relaxed forms (FII), and the percent reaction was calculated by
186 subtracting the amount of already nicked molecules (faint FII band in the untreated
187 samples most likely generated by mechanical shearing) from the FII-forms
188 generated by protein treatment. The values of relaxed molecules were ~65% for the
189 two unlabelled and native proteins. In the case of the RelB-RelE protein-protein
190 complex, functional assays were done by electrophoretic mobility shift assays
191 (EMSA). To this end, 10 nM of the 256-bp DNA fragment (that includes the
192 promoter, transcription initiation site, and the Shine-Dalgarno sequences of the
193 *relBE* pneumococcal operon), was incubated with different amounts of unlabelled
194 or SeMet-labelled RelB-RelEHis_6 proteins (0.2, 0.4 and 0.9 nM) in buffer M (Table 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

After 20 min at room temperature, samples were separated by electrophoresis on native 5% polyacrylamide (PAA) gels. In all cases, gels were stained with ethidium bromide and the DNA bands were visualized with the aid of a Gel-Doc documentation system (Bio-Rad Laboratories).

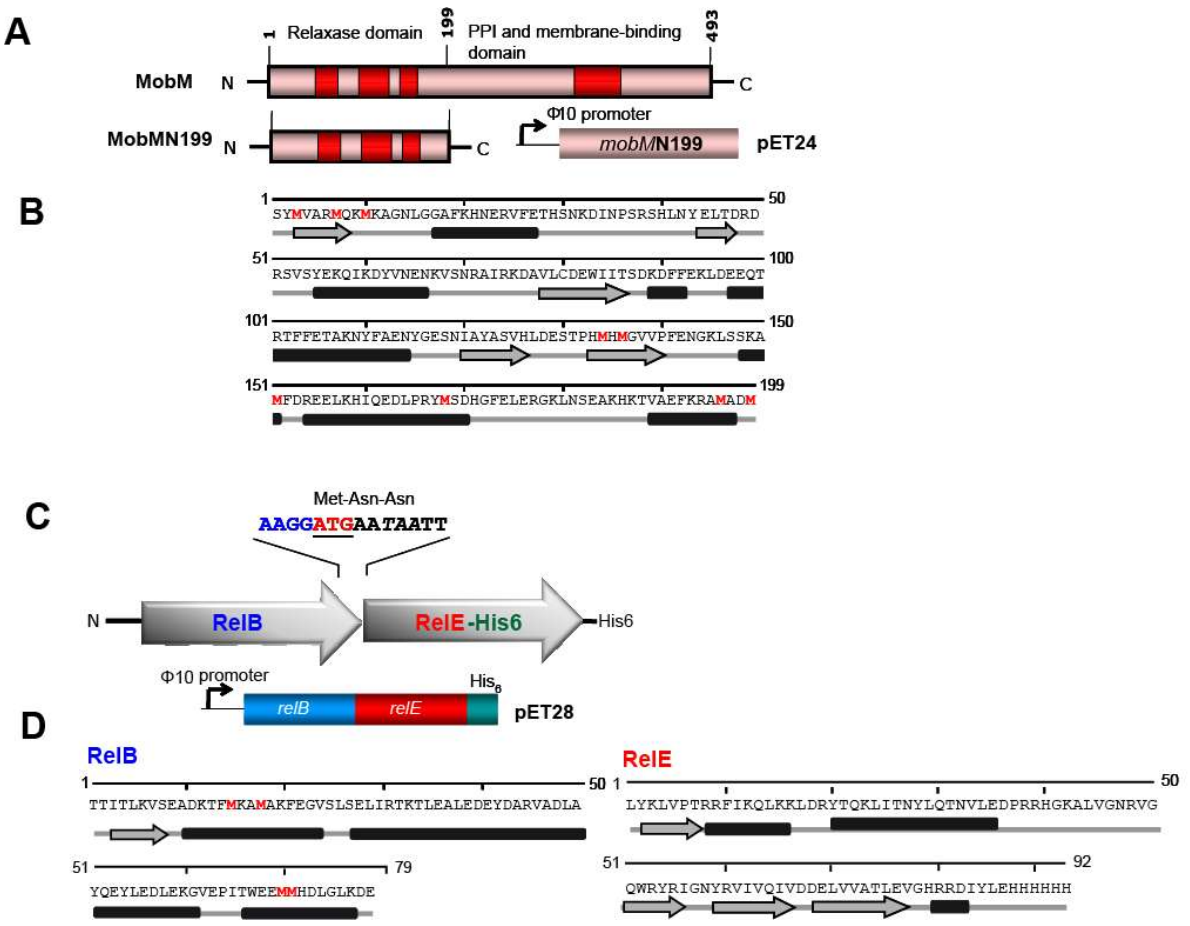
3. Results and Discussion

3.1. Organization of the Pneumococcal Proteins

MobM is the protein that initiates conjugal transfer of plasmid pMV158 by cleavage of the phosphodiesther bond at a specific di-nucleotide (5′-GpT-3′) within the plasmid origin of transfer, *oriT*. The endonuclease activity of MobM is exerted on supercoiled plasmid DNA molecules (forms FI) that are converted into relaxed forms (forms FII) [41]. The native MobM is a prolate-ellipsoid dimer composed by two identical subunits of 494 residues per protomer [42], of which the first Met1 residue is removed after production [28]. The protein has two distinct domains connected by a flexible region (Figure 1A). The N-terminal moiety (around 200 amino acids) encompasses the DNA-binding and relaxase domain [28]. The C-terminal moiety is mostly α-helical and contains: i) the dimerization domain which includes a putative Leu-zipper [43]; ii) the membrane-interaction region [42], and iii) probably the domain involved in interaction with the coupling protein involved in conjugal transfer [43]. Attempts at obtaining crystals using the native full-lenght MobM failed, probably due to the association of MobM with the host membrane [42]. Thus, a truncated version of the protein that includes the first 199 residues harbouring the endonuclease-relaxase domain (MobMN199; Figure 1A) was designed, and the DNA region encoding this truncated protein was cloned into an expression vector. Protein MobMN199 was, thus, the best candidate to achieve the three-dimensional structure of the MobM-relaxase domain in complex with its target DNA [28]. Prediction of the secondary structure of MobMN199 (Figure 1B) showed an alternant distribution of α-helices with β-strands, that is the α/β-fold,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

224 which is found in many of the HUH endonucleases [44]. This prediction was
225 confirmed by circular dichroism analyses [28] and, later on, when we could solve
226 the structure of MobMN199 [24].
227
228



229
230

231 **Figure 1:** Organization (**A**, **C**), and predicted secondary structures (**B, D**) of the proteins analysed in
232 this work. **A.** Protein MobM is encoded by plasmid pMV158 and contains two distinct regions: the
233 N-terminal domain harbours the relaxase domain (MobMN199), whereas the C-terminal domain is
234 involved in protein-protein interactions and membrane association [28,43]. The genetic region
235 encoding MobMN199 was cloned in the expression vector pET24 and the native domain was
236 overexpressed, labelled with SeMet and purified (see Figure 2). **B.** The predicted secondary structure
237 of MobMN199 indicated a distribution of α-helices alternating with β-strands, the so-called α/β-fold
238 typical of many of the HUH endonucleases [44]. The structure of MobMN199 was later solved and
239 corresponded to the prediction [24]. **C.** Schematic representation of the pneumococcal RelB
240 (antitoxin, blue)-RelE (toxin, red) proteins. The termination codon of *relB* (TAA, italics) and the
241 initiation codon of *relE* (ATG, in red) are indicated. The genes encoding these proteins were cloned
242 into the expression vector pET28 which adds a His_6 tag to the C-terminal region of the toxin RelE. **D.**
243 The predicted secondary structure of RelB (left) suggested the existence of an N-terminal β-strand

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

that could be incorporated into a β-sheet by interaction with another RelB molecule to fold as a dimer with a ribbon-helix-helix structure [23]. All Met residues are marked in red; note that due to processing the M1 residue was always removed.

In the case of the pneumococcal RelB-RelE proteins, the genetic structure of the intergenic region (Figure 1C) showed that the ATG start codon of *relE* (encoding the toxin) is placed three codons upstream of the termination (TAA) codon of *relB* (encoding the antitoxin), a situation indicative of translational coupling of both proteins. The operon was cloned into an expression vector that tags the C-terminus of *relE* with six His residues. Concerning the secondary structure predictions of the proteins (Figure 1D), RelB would start with a β-strand followed by three α-helices, a situation that resembles a protein with a ribbon-helix-helix structure [23], whereas RelE would appear to have a distribution of α-helices and β-strands indicative of a more complex structure. Circular dichroism studied supported these predictions [31]. Analytical ultracentrifugacion and native mass spectrometry experiments indicated that the RelB-RelE complex appeared to be a heterohexamer composed by four antitoxin and two toxin protomers [31]. However, this hypothesis must wait until the three-dimensional structures of the RelB-RelE proteins are solved.

3.2. Purification of the SeMet-labelled proteins

Purification of the SeMet MobMN199 and SeMet RelB-RelE complex was done by induction of *E. coli* strain B834(DE3) harbouring plasmid pMobMN199 or plasmid pET28relBE, respectively. Cultures from the frozen inocula were diluted in EMM medium and grown at 37ºC. At OD_{600} = 0.5, cells were induced with 1 mM IPTG, followed by addition of rifampicin and further incubation (see Methods). At the end of the incubation period, cells were centrifuged, washed twice, and suspended (20 x concentrated) in buffer A (MobMN199) or C (RelB-RelE) to which two tablets of a protease inhibitor cocktail (Complete EDTA-free; Roche) were added. The cell paste was passed twice through a French pressure cell.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

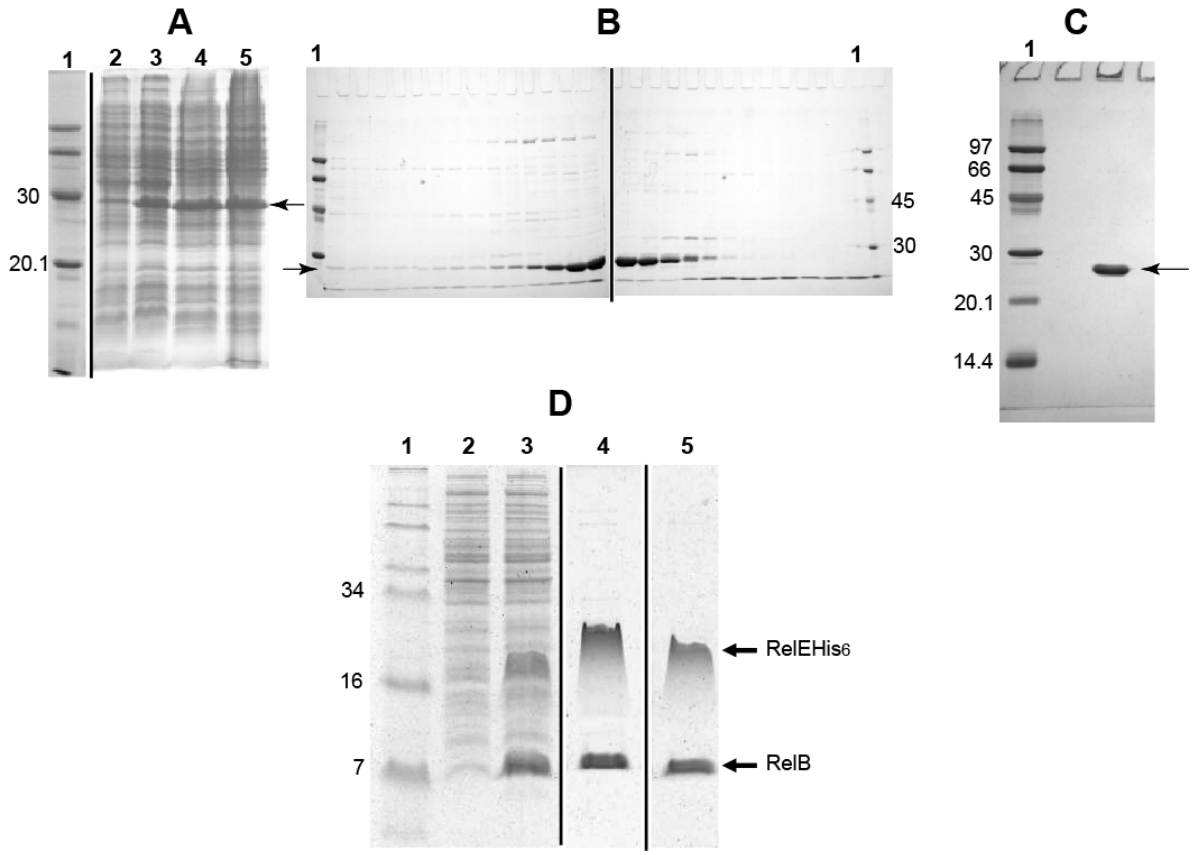


Figure 2: Purification stages of SeMet-labelled proteins. **A-C**: Cells harbouring the plasmid encoding *mobM*N199 were grown in SeMet-containing medium and the different purification stages were analyzed by electrophoresis on 15%-SDS-Tris-glycine PAA gels. Samples shown were from: uninduced cultures (lane 2); cultures induced with IPTG and rifampicin (lane 3); supernatant after PEI precipitation (lane 4), and supernatant of the ammonium sulphate precipitation step after dialysis against buffer A (lane 5). This latter fraction was loaded onto a heparin-agarose column and the retained proteins were eluted by a 0.3-0.8 M NaCl gradient (**B**) as reported [28]. Panel **C** shows the final purified MobMN199 protein. The yield of the purified protein was about 2 mg/ml. Lanes 1 are the molecular weight standards with their mass (kDa) are indicated. **D**: Purification of RelB-RelEHis$_6$ proteins. Cells harbouring the plasmid encoding the *relBE* operon were grown in SeMet-containing medium, and the different purification stages were analyzed by electrophoresis on 16% SDS-Tricine-PAA gels. Proteins were detected by staining with Coomassie Brilliant blue R-250 (Bio-Rad). Samples in the gels were: lane 1, molecular weight standard (SeeBlue Plus2, Invitrogen); lanes 2 and 3, total cell extracts from uninduced and induced cultures, respectively; lane 4, fractions from the eluted samples from the nickel column, and lane 5, purified proteins after gel filtration. Relevant protein positions are indicated. Migration of the two proteins of the complex was anomalous and due to their isoelectric points [31].

The cell lysate containing SeMet-labelled MobMN199 protein was cleared by centrifugation for 30 min, 9500 rpm, 4ºC, and the supernatant was treated with

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

0.2% (v/v) polyethyleneimine (Sigma) to precipitate nucleic acids. Proteins in the supernatant were precipitated at 70% (w/v) ammonium sulphate saturation. The proteins in the precipitate were collected by centrifugation dissolved in Buffer A (Figure 2A). After dialysis against the same buffer, the sample was loaded onto a 100-ml heparin-agarose (BioRad) column (flow rate of 50 ml/h). After washing with 5-column volumes of the same buffer, a 400-ml 0.3-0.8 M NaCl gradient was applied to elute the proteins retained. Fractions were analyzed by 15% SDS-Tris-glycine polyacrylamide gel electrophoresis (SDS-PAGE) followed by staining with Bio-safe Coomassie (BioRad Laboratories) (Figure 2B). Fractions containing the peak of MobMN199 were pooled, dialysed against Buffer A containing 500 mM NaCl, and concentrated by filtering through 3 kDa cut-off membranes (Pall) until the sample volume reached 1 ml. The protein sample was next automatically injected at 0.5 ml/min onto a HiLoad Superdex 200 gel-filtration column (Amersham) and subjected to fast-pressure liquid chromatography (FPLC; Biologic DuoFlow from BioRad). Fractions containing pure MobMN199 protein (>98%) were pooled and concentrated until the final concentration was 5 mg/ml protein (Figure 2C).

In the case of SeMet RelB-RelE, the cell lysate was cleared by centrifugation, 30 min, 9500 rpm, 4ºC, and the supernatant was loaded onto a nickel column (His-select Nickel Affinity Gel, Sigma). After washing with buffer C, proteins were eluted in the same buffer supplemented with 250 mM imidazol. Fractions were analyzed by gel SDS-PAGE on 16% SDS-Tricine, and proteins were detected by staining with Coomassie Brilliant blue R-250 (Bio-Rad Laboratories). Fractions containing the peaks of the desired proteins were pooled, dialyzed against buffer S and applied to a gel filtration column (Superdex 200 XK16/60 column, Amersham Pharmacia Biotech). Fractions were analyzed by gel electrophoresis (Figure 2D), and we found that the RelE(His)$_6$ protein exhibited an anomalous migration (M_r~18000) higher than the theoretical value (M_r~11500). This is probably due to the high isoelectric point of RelE (pI = 10.27) that makes the protein not to achieve a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

uniform negative charge. Fractions containing the desired proteins were pooled and concentrated by filtration through 3 kDa cut-off filters (Pall). The purified proteins were stored at -80ºC where they remained active for at least one year.

3.3. MALDI-TOF Molecular Weight Determinations

To assess the degree of SeMet labelling of the pneumococcal proteins, we determined the molecular weight of the unlabelled and labelled proteins by employment of mass spectra experiments. The results are depicted in Figure 3 and they are summarized in Table 3. MALDI-TOF spectra of MobMN199 unlabelled (Figure 3A) or SeMet-labelled (Figure 3B) showed the existence of a major peak exhibiting molecular weights of 23,128.1 Da and 23,551.3 Da, respectively. The difference between both values was of 423.2 Da, which is consistent with the calculated theoretical value of MobMN199 protomers in which all the 9 Met residues (Met1 was processed) were substituted by SeMet ones. Thus, the efficiency of labelling was of 100%.

In the case of RelB proteins, unlabelled or labelled, two major peaks were detected, the main one corresponding to RelB lacking M1, and a minor one pertaining to the full RelB (Figure 3C). A similar finding was observed for RelE, but in this case the main peak corresponded to the protein lacking the first three residues (M1, N2, and N3), whereas the second minor peak could be assigned to the entire RelE (Figure 3D). Again, the efficiency of labelling was of 100%.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

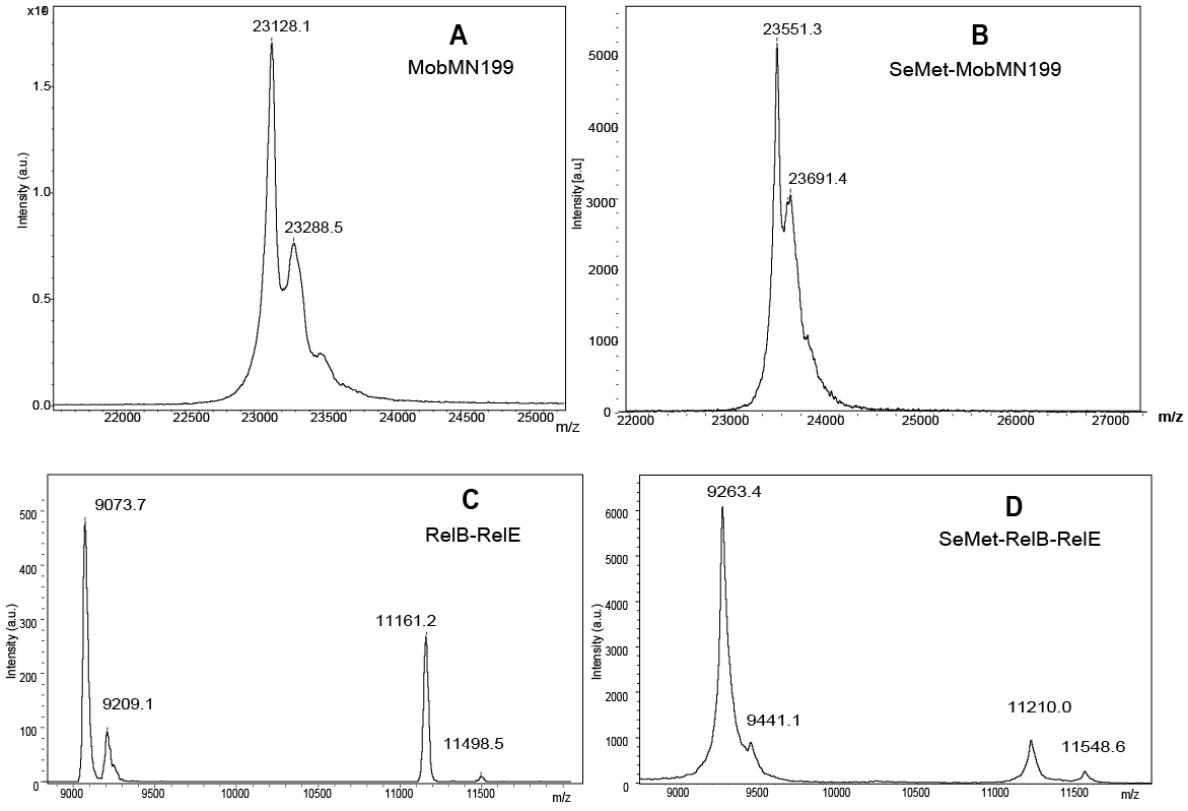


Figure 3: MALDI-TOF mass spectrometry of the pneumococcal DNA-binding proteins. Proteins analyzed were: MobMN198 (**A**, **B**), and RelB-RelE toxin-antitoxin proteins (**C**, **D**), unlabelled (**A**, **C**) and SeMet-labelled (**B**, **D**).The spectra and masses of the different molecules identified are indicated. The first Met residue of MobMN199 was processed, thereby leaving a 198-amino acid protein (MobMN198) containing 9 Met residues. Similarly, the first Met of RelB was removed and the resulting protein has 4 Met residues. In the case of RelE, most of the synthesized protein lacks the first three residues (Met-Asn-Asn), and the resulting protein has no Met residues [1].

Table 3: Expected and MALDI-TOF determined molecular weights of the SeMet-labelled pneumococcal proteins.

Features / Protein	MobMN199	RelB	RelEHis$_6$
Predicted MW of unlabelled protomer (Da)	23129.8	9205[a]	11505.3
Determined MW of unlabelled protomer (Da)	23128.1	9073.7	11161.2[b]
Number of Met residues [a]	9	4	0
Theoretical MW of SeMet-labelled protomer (Da)	23551.9	9262	11552[b]
Determined MW of SeMet-labelled protomer (Da)	23551.3	9263.4	11210[b]
Percent incorporation	**100**	**100**	**---**[c]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

[a] The first Met residue of MobM and of RelB was removed by processing, as determined by N-terminal sequencing of the proteins [31,43].

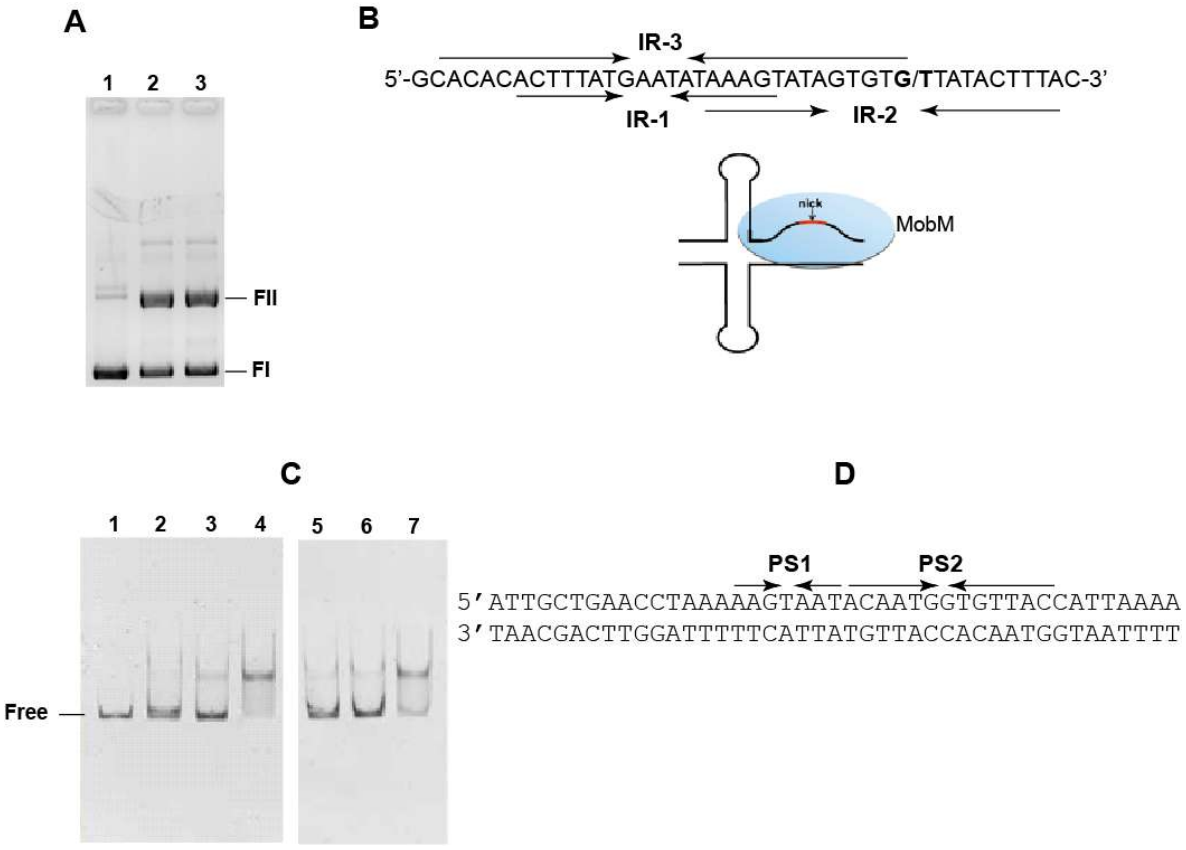
[b] The vast majority of the protein lacked the first three amino acid residues (M1, N2, and N3) as previously determined [31].

[c] The percent incorporation could not be determined due to lack of any Met residue of the majority of the protein.

3.4. Functional Assays

To determine whether the SeMet-labelled proteins retained full activity, their interactions with substrate DNA were assayed. As stated above, MobMN199 harbours the relaxase domain of the parental protein [28] and thus, the truncated version is able to relax supercoiled cognate DNA as efficiently as the full-length MobM. We tested the activity of the SeMet-MobMN199 on CsCl-purified pMV158 DNA under standard conditions (Methods and Figure 4A). Briefly, 8 nM supercoiled pMV158 DNA was incubated (lanes 2 and 3) or not (lane 1) with 480 nM of unlabelled (lane 2) or SeMet-labelled (lane 3) MobMN199 in buffer A (Table 2) containing 15 mM $MnCl_2$ (final concentration). After 20 min incubation at 30ºC, the reaction products were separated by electrophoresis on 1% agarose gels and stained with ethidium bromide. The results showed that the amount of forms FII (relaxed DNA molecules) generated by both proteins was roughly the same, an amount of 65% of relaxed FII forms derived from the supercoiled FI substrate, indicating that the cleavage reaction mediated by MobM (Figure 4B) was independent on whether the protein has its Met residues substituted or not by SeMet.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019



386

387

388 **Figure 4:** Functional assays of the SeMet-labelled proteins. (**A**). MobMN198 has endonucleolytic
389 nicking-closing activity on supercoiled plasmid DNA molecules (Forms FI) that are converted into
390 relaxed molecules (Forms FII). Lanes shown in the gel are: 1, no protein, 2 unlabelled MobMN199,
391 and 3, SetMet-labelled MobM. The weak band above relaxed forms FII has been observed before and
392 might correspond to relaxed DNA dimers and thus not cleaved by the proteins [28,40]. (**B**). The
393 minimal origin of transfer recognized by MobMN198 harbours three inverted repeats (IR-1 to IR-3)
394 and contains the di-nucleotide (GpC, boldface letters) cleaved by the protein; below it is shown a
395 scheme showing the MobMN198-generation of a DNA structure that opens the DNA strands and
396 facilitates cleavage of the protein [40]. (**C**). RelB-RelE protein complex has high affinity to bind linear
397 double-stranded DNA containing its target, and no differences were found between the unlabelled
398 (lanes 2-4) and SeMet-labelled proteins (lanes 5-7); lane 1, no protein was added. The amount of
399 DNA used was 10 nM, whereas the proteins were used at 0.2 (lanes 2 and 5), 0.4 (lanes 3 and 6), and
400 0.9 nM (lanes 4 and 7). (**D**). Nucleotide sequence of a region of DNA spanning the target of the RelBE
401 proteins. This region contains two palindromic sequences (PS1 and PS2) that are postulated to be the
402 target of the RelBE complex [31].

403

404 In the case of the RelB-RelE pneumococcal protein complex, the antitoxin RelB
405 has a moderate DNA-binding affinity that is augmented by the toxin RelE, which
406 acts as a co-repressor of their own synthesis [29,31]. The DNA-binding ability of the
407 RelB-RelE complex was tested by EMSA assays (Figure 4C), using as substrate a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

double-stranded 256-bp fragment containing the RelB-RelE target (palindromes PS1 and PS2 in Figure 4D; [31]). The results showed that the free DNA (Figure 4C, lane 1) molecules were progressively retarded into protein-DNA complexes that increased with the protein concentrations used; the increase of retarded bands was nearly in the same proportion independently whether the proteins were unlabelled (lanes 2-4) or labelled with SeMet (lanes 5-7).

5. Conclusions

The results presented here represent an optimization of previous methods [1] to achieve full labelling of pneumococcal DNA-binding proteins with SeMet. Two variants were chosen: untagged (MobMN199) and His-tagged (RelB-RelEHis$_6$) proteins, demonstrating the applicability of the method, in which 100% efficiency of labelling was attained in all cases. Further, the SeMet-labelled proteins were shown to retain full activity as judged by DNA relaxation and EMSA experiments through comparison of the unlabelled and labelled proteins. Finally, the structure of MobMN199 has been solved with the aid of SeMet-labelled protein [24], whereas that of the RelB-RelE complex is being worked out at present (unpublished results).

Author Contributions: All authors designed the experiments, which were performed by F.L-D. and I.M-C. M.E. analyzed the results and wrote the first draft, which was corrected by all authors.

Funding: Research supported by Grant BIO2015-69085-REDC from the Spanish Ministry of Economy and Competitiveness

Acknowledgements: Thanks are due to Concha Nieto, Alicia Bravo, and Cris Fernández-López for helpful discussions and suggestions, and to our crystallographer colleagues (Miquel Coll's group, IBM-Barcelona, Spain) for discussions along the crystallization experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Budisa, N.; Steipe, B.; Demange, P.; Eckerskorn, C.; Kellermann, J.; Huber, R. High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eu. J. Biochem.* **1995**, *230*, 788-796.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

2. Mandell, L.A.; Wunderink, R.G.; Anzueto, A.; Bartlett, J.G.; Campbell, G.D.; Dean, N.C.; Dowell, S.F.; File, T.M., Jr; Musher, D.M.; Niederman, M.S., *et al.* Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin. Infect. Dis.* **2007**, *44*, S27-72.

3. Meichtry, J.; Born, R.; Küffer, M.; Zwahlen, M.; Albrich, W.C.; Brugger, S.D.; Mühlemann, K.; Hilty, M. Serotype epidemiology of invasive pneumococcal disease in Swiss adults: A nationwide population-based study. *Vaccine* **2014**, *32*, 5185-5191.

4. Bravo, A.; Ruiz-Cruz, S.; Alkorta, I.; Espinosa, M. When humans met superbugs: Strategies to tackle bacterial resistances to antibiotics. *Biomol. Concepts* **2018**, *9*, 216-226.

5. Chan, W.T.; Balsa, D.; Espinosa, M. One cannot rule them all: Are bacterial toxins-antitoxins druggable? *FEMS Microbiol. Rev.* **2015**, *39*, 522-540.

6. Chan, W.T.; Espinosa, M. The antibacterials that have yet to be found. *Atlas of Science* 2016.

7. Fernandez-Lopez, R.; Machon, C.; Longshaw, C.M.; Martin, S.; Molin, S.; Zechner, E.L.; Espinosa, M.; Lanka, E.; de la, C., F. Unsaturated fatty acids are inhibitors of bacterial conjugation. *Microbiology* **2005**, *151*, 3517-3526.

8. García-Cazorla, Y.; Getino, M.; Sanabria-Ríos, D.J.; Carballeira, N.M.; de la Cruz, F.; Arechaga, I.; Cabezón, E. Conjugation inhibitors compete with palmitic acid for binding to the conjugative traffic ATPase TrwD, providing a mechanism to inhibit bacterial conjugation. *J. Biol. Chem.* **2018**, *293*, 16923-16930.

9. Grohmann, E.; Christie, P.J.; Waksman, G.; Backert, S. Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* **2018**, *107*, 455-471.

10. Pandey, D.P.; Gerdes, K. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res.* **2005**, *33*, 966-976.

11. Mutschler, H.; Gebhardt, M.; Shoeman, R.L.; Meinhart, A. A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biol.* **2011**, *9*, e1001033.

12. Hayes, F.; Kędzierska, B. Regulating toxin-antitoxin expression: Controlled detonation of intracellular molecular timebombs. *Toxins* **2014**, *6*, 337-358.

13. Kędzierska, B.; Hayes, F. Emerging roles of toxin-antitoxin modules in bacterial pathogenesis. *Molecules* **2016**, *21*, 790.

14. Ainelo, A.; Tamman, H.; Leppik, M.; Remme, J.; Hõrak, R. The toxin GraT inhibits ribosome biogenesis. *Molecular Microbiology* **2016**, *100*, 719-734.

15. Kędzierska, B.; Hayes, F. Transcriptional control of toxin–antitoxin expression: Keeping toxins under wraps until the time is right. In *Stress and environmental regulation of gene expression and adaptation in bacteria*, John Wiley & Sons, Inc.: 2016; pp 463-472.

16. Williams, J.J.; Hergenrother, P.J. Artificial activation of toxin-antitoxin systems as an antibacterial strategy. *Trends Microbiol.* **2012**, *20*, 291-298.

17. Fernández-Bachiller, M.; Brzozowska, I.; Odolczyk, N.; Zielenkiewicz, U.; Zielenkiewicz, P.; Rademann, J. Mapping protein–protein interactions of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

resistance-related bacterial Zeta toxin–Epsilon antitoxin complex (ε2ζ2) with high affinity peptide ligands using fluorescence polarization. *Toxins* **2016**, *8*, 222.

18. Lioy, V.S.; Rey, O.; Balsa, D.; Pellicer, T.; Alonso, J.C. A toxin-antitoxin module as a target for antimicrobial development *Plasmid* **2010**, *63*, 31-39.

19. Bienstock, R.J. Computational drug design targeting protein–protein interactions. *Curr. Pharm. Des.* **2012**, *18*, 1240-1254.

20. Tsao, D.H.H.; Sutherland, A.G.; Jennings, L.D.; Li, Y.; Rush, T.S.; Alvarez, J.C.; Ding, W.; Dushin, E.G.; Dushin, R.G.; Haney, S.A., *et al.* Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. *Bioorg. Med. Chem.* **2006**, *14*, 7953-7961.

21. Boer, D.R.; Ruíz-Masó, J.A.; López-Blanco, J.R.; Blanco, A.G.; Vives-Llàcer, M.; Chacón, P.; Usón, I.; Gomis-Rüth, F.X.; Espinosa, M.; Llorca, O., *et al.* Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. *EMBO J.* **2009**, *28*, 1666-1678.

22. Espinosa, M. Plasmids as models to study macromolecular interactions: The pMV158 paradigm. *Res. Microbiol.* **2013**, *164*, 199-204.

23. Gomis-Ruth, F.X.; Sola, M.; Acebo, P.; Parraga, A.; Guasch, A.; Eritja, R.; Gonzalez, A.; Espinosa, M.; del Solar, G.; Coll, M. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J.* **1998**, *17*, 7404-7415.

24. Pluta, R.; Boer, D.R.; Lorenzo-Díaz, F.; Russi, S.; Gómez, H.; Fernández-López, C.; Pérez-Luque, R.; Orozco, M.; Espinosa, M.; Coll, M. Structural basis of a histidine-DNA nicking/joining mechanism for gene transfer and promiscuous spread of antibiotic resistance. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6526-E6535.

25. Gutiérrez-Fernández, J.; Saleh, M.; Alcorlo, M.; Gómez-Mejía, A.; Pantoja-Uceda, D.; Treviño, M.A.; Voß, F.; Abdullah, M.R.; Galán-Bartual, S.; Seinen, J., *et al.* Modular architecture and unique teichoic acid recognition features of choline-binding protein L (CbpL) contributing to pneumococcal pathogenesis. *Sci. Reports* **2016**, *6*, 38094.

26. Perez-Dorado, I.; Campillo, N.E.; Monterroso, B.; Hesek, D.; Lee, M.; Paez, J.A.; Garcia, P.; Martinez-Ripoll, M.; Garcia, J.L.; Mobashery, S., *et al.* Elucidation of the molecular recognition of bacterial cell wall by modular pneumococcal phage endolysin Cpl-1. *J. Biol. Chem.* **2007**, *282*, 24990-24999.

27. Rued, B.E.; Alcorlo, M.; Edmonds, K.A.; Martínez-Caballero, S.; Straume, D.; Fu, Y.; Bruce, K.E.; Wu, H.; Håvarstein, L.S.; Hermoso, J.A., *et al.* Structure of the large extracellular loop of FtsX and its interaction with the essential peptidoglycan hydrolase PcsB in *Streptococcus pneumoniae*. *mBio* **2019**, *10*, e02622-02618.

28. Lorenzo-Díaz, F.; Dostál, L.; Coll, M.; Schildbach, J.F.; Menendez, M.; Espinosa, M. The MobM-relaxase domain of plasmid pMV158: Thermal stability and activity upon Mn^{2+}-and DNA specific-binding. *Nucleic Acids Res.* **2011**, *39*, 4315-4329.

29. Nieto, C.; Pellicer, T.; Balsa, D.; Christensen, S.K.; Gerdes, K.; Espinosa, M. The chromosomal *relBE2* toxin-antitoxin locus of *Streptococcus pneumoniae*: Characterization and use of a bioluminescence resonance energy transfer assay to detect toxin-antitoxin interaction. *Mol. Microbiol.* **2006**, *59*, 1280-1296.

30. Studier, F.W.; Moffatt, B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **1986**, *189*, 113-130.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2019

31. Moreno-Córdoba, I.; Diago-Navarro, E.; Barendregt, A.; Heck, A.J.R.; Alfonso, C.; Díaz-Orejas, R.; Nieto, C.; Espinosa, M. The toxin-antitoxin proteins RelBE2*spn* of *Streptococcus pneumoniae*: Characterization and association to their DNA target. *Proteins* **2012**, *80*, 1834-1846.

32. Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidelberg, J.; DeBoy, R.T.; Haft, D.H.; Dodson, R.J., *et al.* Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **2001**, *293*, 498-506.

33. Burdett, V. Identification of tetracycline-resistant R-plasmids in *Streptococcus agalactiae* (group B). *Antimicrob. Agents Chemother.* **1980**, *18*, 753 - 760.

34. del Solar, G.; Díaz, R.; Espinosa, M. Replication of the streptococcal plasmid pMV158 and derivatives in cell-free extracts of *Escherichia coli*. *Mol. Gen. Genet.* **1987**, *206*, 428-435.

35. Studier, F.W.; Rosenberg, A.H.; Dunn, J.J.; Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* **1990**, *185*, 60-89.

36. Lacks, S. Genetic regulation of maltosaccharide utilization in pneumococcus. *Genetics* **1968**, *60*, 685-706.

37. Ruiz-Cruz, S.; Solano-Collado, V.; Espinosa, M.; Bravo, A. Novel plasmid-based genetic tools for the study of promoters and terminators in *Streptococcus pneumoniae* and *Enterococcus faecalis*. *J. Microbiol. Meth.* **2010**, *83*, 156-163.

38. López, P.; Espinosa, M.; Stassi, D.L.; Lacks, S.A. Facilitation of plasmid transfer in *Streptococcus pneumoniae* by chromosomal homology. *J. Bacteriol.* **1982**, *150*, 692-701.

39. Maniatis, T.; Fritsch, E.F.; Sambrook, J. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press: New York, 1982.

40. Lorenzo-Díaz, F.; Fernández-López, C.; Guillén-Guío, B.; Bravo, A.; Espinosa, M. Relaxase MobM induces a molecular switch at its cognate origin of transfer. *Front. Mol. Biosci.* **2018**, *5*.

41. Guzmán, L.M.; Espinosa, M. The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid *oriT*. *J. Mol. Biol.* **1997**, *266*.

42. de Antonio, C.; Farias, M.E.; de Lacoba, M.G.; Espinosa, M. Features of the plasmid pMV158-encoded MobM, a protein involved in its mobilization. *J. Mol. Biol.* **2004**, *335*, 733-743.

43. Fernández-López, C.; Lorenzo-Díaz, F.; Pérez-Luque, R.; Rodríguez-González, L.; Boer, R.; Lurz, R.; Bravo, A.; Coll, M.; Espinosa, M. Nicking activity of the pMV158 MobM relaxase on cognate and heterologous origins of transfer. *Plasmid* **2013**, *70*, 120-130.

44. Chandler, M.; Cruz, F.; Dyda, F.; Hickman, A.B.; Moncalian, G.; Ton-Hoang, B. Breaking and joining single-stranded DNA: The HUH endonuclease superfamily. *Nat Rev Microbiol* **2013**, *11*.