Inês Manuel de Sousa Martins de Castro

Halocins and lanthipeptides from *Haloferax mediterranei*

Halocinas e lantipéptidos de *Haloferax mediterranei*

DECLARAÇÃO

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

Inês Manuel de Sousa Martins de Castro

Halocins and lanthipeptides from *Haloferax mediterranei*

Halocinas e lantipéptidos de *Haloferax mediterranei*

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Microbiologia, realizada sob a orientação científica da Doutora Tânia Isabel Sousa Caetano, Investigadora Auxiliar do Departamento de Biologia da Universidade de Aveiro e Professora Doutora Sónia Alexandra Leite Velho Mendo Barroso, Professora Auxiliar com Agregação do Departamento de Biologia da Universidade de Aveiro.

Dedico este trabalho à minha família e aos meus queridos amigos.

o júri

presidente   Prof. Doutora Maria Paula Polónia Gonçalves
professor associado do Departamento de Biologia da Universidade de Aveiro

Prof. Doutora Ana Sofia Duarte
professor auxiliar da Faculdade de Medicina Dentária da Universidade Católica Portuguesa

Doutora Tânia Isabel Sousa Caetano
investigadora auxiliar do Departamento de Biologia e CESAM da Universidade de Aveiro

agradecimentos

Gostaria de agradecer a todas as pessoas que me ajudaram, a minha família, os meus amigos, a minhas orientadoras, especialmente à Doutora Tânia pela dedicação e amizade e à Professora Doutora Sónia pela amabilidade e oportunidade de trabalhar no seu laboratório. Também queria agradecer a toda a família LBM, à Diana por todo o seu apoio (és a maior), às Joanas pela ajuda dada, à Vanessa e à Sofia.

resumo

O estudo dos metabolitos secundários produzidos por Arquea, incluindo as arqueocinas, péptidos com atividade antimicrobiana, é ainda muito limitado, especialmente quando comparado com os estudos existentes relacionados com a produção destes compostos por outros microrganismos. Apenas dois tipos de arqueocinas são conhecidos: i) halocinas produzidas por arqueas halofílicas e ii) sulfolobicinas, produzidas por um extremófilo do género *Sulfolobus* spp.. Também foram reportadas, arqueocinas promissoras com possíveis caraterísticas anticancerígenas. Arquea halofílicas foram recentemente encontradas no intestino humano, mostrando que a sua presença não se restringe apenas a ambientes hipersalinos. A descoberta das halocinas é recente e grande parte da sua caracterização baseia-se em ensaios feitos com sobrenadantes de culturas. Apenas algumas halocinas foram purificadas e sequenciadas com sucesso, e só para um grupo mais restrito é que existe uma proposta de modelo biossintético. *H. mediterranei* ATCC 33500 tem atividade anti-arquea. Estudos determinaram que esta estirpe produz a halocina HalH4. No entanto, ao longo dos últimos anos, foi demonstrado que mesmo na ausência do gene *halH4* esta estirpe manteve a sua capacidade anti-arquea. Assim, a(s) molécula(s) responsável(eis) por tal atividade ainda é desconhecida. *H. mediterranei* codifica no seu genoma três enzimas modificadoras de lantipéptidos de classe II (MedM1, MedM2 e MedM3) e alguns péptidos precursores. Uma elevada percentagem dos lantipéptidos produzidos por bactérias tem actividade antimicrobiana. Este estudo teve como objectivo resumir a informação disponível até agora sobre as haloarqueocinas (as halocinas produzidas por Archaea) a dois níveis: bibliográfico e através da análise dos *clusters* biossintéticos conhecidos até agora utilizando genómica comparativa. Para a haloarqueocina HalC8, foi possível determinar os possíveis genes biossintéticos envolvidos na sua produção. Estes genes codificam uma proteína de função desconhecida (*halU*), dois péptidos localizados na membrana (*halP1* e *halP2*) e um regulador transcricional (*halR*). Outro objetivo foi determinar se os lantipéptidos de *H. mediterranei* ATCC 33500 eram haloarqueocinas. Para tal, mutantes *knock-out* sem os genes *medM1*, *medM2* e *medM3* foram obtidos utilizando a estratégia *pop-in* e *pop-out*. Verificou-se que são necessários aproximadamente 20 dias e 6 meses para obter uma única ou tripla estirpe *knock-out*, respetivamente. A bioatividade do *knock-out* triplo (*ΔM1M2M3*) foi testada contra outras haloarqueas. No entanto, não foram observadas diferenças nos halos produzidos pela estirpe *ΔM1M2M3* e pela sua estirpe parental (WR510). Estes resultados provam que os supostos lantipéptidos de classe II de *H. mediterranei* não estão envolvidos no seu perfil anti-arquea. Assim, a sua função em haloarquea continua desconhecida.

keywords

abstract

The study of archaea's secondary metabolites, including archaeocins, is still limited. These antimicrobial peptides are poorly studied, especially when compared to the numerous studies on antibiotic production by other microorganisms. Only two types of archaeocins are known: i) halocins, produced by halophilic archaea and ii) sulfolobicins, produced by the extremely thermophilic *Sulfolobus* spp. There are also promising reports of archaeocins endowed with anticancer properties. Halophilic archaea have recently been found to be present in the human gut, thus showing that they are not confined to high salt environments alone. Halocins were firstly discovered in the 80's and most of their characterization was solely based on supernatant-based assays. In fact, only a few halocins were successfully purified and sequenced, and even fewer have a proposed biosynthetic mechanism. Also, their mode of action, ecological role and biotechnological potential are still little explored.
H. mediterranei ATCC 33500 has antiarchaeal activity. Studies determined that these strains produced the HalH4 halocin. However, over the last years, it was shown that strains lacking the *halH4* gene retained their antiarchaeal ability. So, the molecule(s) responsible for its microbial activity is still unknown. This strain encodes in its genome three class II lanthipeptide synthetases (MedM1, MedM2 and MedM3) and some putative lanthipeptide precursor peptides. A high percentage of the lanthipeptides produced by Bacteria has antimicrobial activity. This study aimed to summarize the information available so far on haloarcheocins (the halocins produced by Archaea) at two levels: bibliographical and by analysing the gene clusters known so far using comparative genomics. For the haloarcheocin, HalC8, it was possible to determine the putative biosynthetic clusters involved in the production of HalC8 and HalC8-related peptides by Haloarchaea, which includes a protein of unknown function (HalU), two membrane-located peptides (HalP1 and HalP2) and a transcriptional regulator (HalR). Other aim of this study was to determine if the lanthipeptides of *H. mediterranei* ATCC 33500 were haloarcheocins contributing to its antimicrobial profile. To achieve this, knock-out mutants without *medM1*, *medM2* and *medM3* genes were obtained by employing the pop-in and pop-out strategy. It was found that approximately 20 days and 6 months are required to obtain a single or a triple knock-out strains, respectively. The bioactivity of the triple knock-out (Δ*M1M2M3*) was tested against other halobacteria. However, no differences were observed in the halos produced by the Δ*M1M2M3* strain and its parental strain (WR510). These results prove that the putative class II lanthipeptides of *H. mediterranei* are not involved in its antiarchaeal profile. Thus, their function in haloarchaea is still to be unravelled.

Table of contents

List of Figures

List of Tables

List of Abbreviations

5-FOA- 5-fluorotic acid

BRE- B-recognition element

CA- casamino acids

DCCD- dicyclohexylcarbodiimide

Dha- 2,3-didehydroalanine

Dhb- 2,3-didehydrobutyrine

EPS- exopolysaccharide

HalR1- Halocin R1

HLH- helix-loop-helix

LB- Luria Bertani

MCS- multiple cloning site

NHE- Na+/H+ exchangers

OCS- one-component system

PTM- post-translational modification

qRT-PCR- quantitative real-time PCR

RiPPs- ribosomally synthesized peptides that undergo post-translational modifications

YPC- yeast-peptone-casamino acids

YPC-SS- YPC Super-salt

1. INTRODUCTION

Part of this chapter is included in the manuscript that is under revision:

Castro I., Mendo S., Caetano T. Antibiotics from haloarchaea: what can we learn from comparative genomics? Marine Biotechnology (under review).

1.1 Halocins

1.1.1. Introduction

Salt-dependent halophilic archaea belong to the order *Halobacteriales*, phylum Euryarcheaota, and require high content of NaCl to survive (Oren 2014). Thus, their typical habitats include natural brines, salt lakes, marine solar salterns and they were even isolated from a salt sediment dated to the Permian-Triassic period (195–250 million years ago). However, halophilic archaea were also identified in the human microbiome. The genera *Haloferax, Halorubrum, Halococcus, Halosimplex and Natrorubrum* are the most abundantly found in the human intestinal tract and oral cavity (Lagier et al. 2016; Nkamga, Henrissat, and Drancourt 2017). Some strains test positive in humans due to ingestion of salty foods, being merely passengers of the human digestive tract. Yet, others can thrive in the gut, representing a major fraction of the diversity in obese individuals with microbiota alteration (Seck et al. 2019).

In 1982, it was discovered that haloarchaea produce antimicrobial compounds of protein origin that were termed halocins (Rodriguez-Valera *et al.*, 1982). At that time, extreme halophilic archaea still belonged to the domain Bacteria. Thus, halocin was loosely used to describe any proteinaceous antimicrobial compound produced by halophilic archaea and/or bacteria. Nowadays, the term includes any ribosomally synthesized antimicrobial peptide or protein produced by haloarchaeal strains (O'Connor and Shand 2002). Still, the term halocin is sometimes applied to any substance, with antimicrobial activity produced by prokaryotic halophilic strains (archaea or bacteria; (Atanasova, Pietilä, and Oksanen 2013). Occasionally, it is useful to have classification schemes based on the producers. In this context, the use of haloarcheocins and halobacteriocins seems a suitable alternative to distinguish between archaeal and bacterial halocins, respectively. According to this, haloarcheocins, along with sulfolobicins, are archaeocins, since they are antimicrobial peptides/proteins produced by the domain Archaea (O'Connor and Shand 2002). Some archaeocins have also be endowed with anticancer properties. For instance, a "cryptic" peptide derived from the transcription factor Stf76 of *Sulfolobus islandicus* (Notomista et al. 2015) was effective in decreasing cell viability in both human and murein malignant cells (Gaglione et al. 2017).

The halocins described so far have great diversity and the knowledge on their biosynthesis is still scarce, making it difficult to characterize and group them. Due to its relatively recent discovery as well as the establishment of Archaea as a domain, their study is still far behind in comparison to the antimicrobials produced by the other domains.

Table 1: Halocins described to date with the indication of their producers and the conditions of production. ND: Not described

Halocin	Producer	NaCl	Temperature (ºC)	pH	Refs
A4	Haloarchaeon TuA4	4 M	41 °C	7,4	(Haseltine et al. 2001)
C8	*Natrinema* sp. AS7092	4,3 M	37 °C	7,0-7,2	(Li et al. 2003; Sun et al. 2005)
G1	*Halobacterium* strain GRB	ND*	ND*	ND*	(Soppa and Oesterhelt 1989)
H1	*Haloferax mediterranei* M2a (Xai3)	4 M	37-42 °C	7	(Platas, Meseguer, and Amils 2002)
H2	Haloarchaeon GLA22	ND*	ND*	ND*	(Rodriguez-Valera, Juez, and Kushner 1982)
H3	Haloarchaeon GAA12	ND*	ND*	ND*	(Rodriguez-Valera, Juez, and Kushner 1982)
H4	*Haloferax mediterranei* R4	4 M	37 °C	7,2	(I. Meseguer and Rodriguez-Valera 1986)
H5	Haloarchaeon MAA220	ND*	ND*	ND*	(Rodriguez-Valera, Juez, and Kushner 1982)
H6	*Haloferax* gibbonsii Ma2.39	4 M	37 °C	7,2	(M. Torreblanca, Meseguer, and Rodriguez-Valera 1989)
HA1	*Haloferax larsenii* HA1	2,6 M	35/37 °C	7,2	(Kumar and Tiwari 2017b)
HA3	*Haloferax larsenii* HA3	2,6 M	42 °C	7,2	(Kumar and Tiwari 2017a)
KPS1	*Haloferax volcanii* KPS1	4,3 M	40 °C	7	(Kavitha et al. 2011)
R1	*Halobacterium salinarum* GN101	4 M	41 °C	7,4	(Haseltine et al. 2001)
S8	Haloarchaeon S8a	4,3 M	41 °C	7	(Price and Shand 2000)
Sech7a	*Haloferax mediterranei* Sech7a	3,4 M	45 °C	7,5	(Pašić, Velikonja, and Ulrih 2008)
SH10	*Natrinema* sp. BTSH10	3 M	42 °C	8	(Karthikeyan, Bhat, and Chandrasekaran 2013)

*ND: Not described

1.1.2. Halocins described

Halocins were firstly described in 1982, when it was found that supernatants of halophilic isolates from a solar alter in Alicante (Spain) displayed antagonistic activity against closely related strains. Its proteinaceous nature was demonstrated by the loss of activity after protease treatment. Of all strains, 13 were further studied, all of them belonging to the only genus "available" for halophilic bacteria, *Halobacterium* (Rodriguez-Valera, Juez, and Kushner 1982). However, their cell extracts revealed the presence of ether-linked lipids (not ester-linked as in bacteria) and the absence of glucosamine, a key component of the bacterial peptidoglycan. Thus, this was a clue that these strains were not bacteria. At that time, the halocins were classified according to the inhibitory spectrum of the producers against the bioindicator strains (H1, H2, H3, H4, H5 and H6).

Nowadays, halocins are often divided into two categories: peptide halocins, also known as microhalocins and protein halocins. Their size boundary is not clear, although it is established to be around 10 kDa. It seems that there is an inverse relationship between halocins molecular weight and their heat stability (Table 2): over 30 kDa they are thermolabile, between 10-15 kDa they are heat-stable up to 80 °C, whereas microhalocins (<15 kDa) are extremely stable to heat, withstanding boiling temperatures (Table 2). Additionally, microhalocins can withstand desalting, prolonged storage and exposure to organic solvents, being more robust and resistant to harsh conditions when compared with protein halocins (Atanasova, Pietilä, and Oksanen 2013; O'Connor and Shand 2002). Up to now, 16 halocins were described so far (Table 1). They are mainly produced by *Haloferax* spp. (32%), followed by *Natrinema* spp. (18%) and *Haloterrigena* spp. (18%). Many of the described halocins were detected in a single haloarchaeon organism (Table 1). Haloarchaeal genera have similar optimal growth requirements, making the conditions of halocins production very similar. They are produced in high-salinity media containing from 2,6 to 4M of NaCl, at a pH between 7 and 8, and at incubation temperatures ranging from 35 °C to 45 °C. Halocins have an acidic nature that is considered an adaptive mechanism to thrive in high salinity environments (Li et al. 2003). Except for halocins Sech7a, HA1, and H1, all the other halocins activity begins to be detected in the early stationary phase (Table 2).

Not all halocins are at the same level of characterization. Up to now, only three of all halocins described have their peptides and their respective gene sequences determined. They are, H4, which was the first halocin described, halocin S8 and halocin C8 (Besse et al. 2015). Therefore, most of the bioactivity studies were performed with supernatants and not with purified peptides.

Table 2: Functional characteristics of halocins (size, thermal stability and salt-dependency) and bioactivity properties (spectrum of activity, detectable activity and mechanism of action).

Halocin	Size (kDa)	Thermal stability	Salt dependent[a]	Spectrum of activity	Detectable activity[b]	Mecanism of action	Refs
A4	< 5	≥ 1 week boiling	N	Broad (cross-phylum)	*ND	*ND	(Haseltine et al. 2001)
C8	7,4	≥ 24 h boiling	N	Broad (in haloarchaeal strains)	TSF	Acts on the cell wall	(Li et al. 2003; Sun et al. 2005)
G1	*ND	*ND	*ND	*H. halobium*	TSF	*ND	(Soppa and Oesterhelt 1989)
H1	31	<50°C	Y	Broad (in haloarchaeal strains)	MEP	Alteration of membrane permeability	(Platas, Meseguer, and Amils 2002)
H2	*ND	*ND	*ND	Broad (in haloarchaeal strains)	*ND	*ND	(Rodriguez-Valera, Juez, and Kushner 1982)
H3	*ND	*ND	*ND	Broad (in haloarchaeal strains)	*ND	*ND	(Rodriguez-Valera, Juez, and Kushner 1982)
H4	34,9	<60°C	Y	Narrow	TSF	Acts on cell membrane: changes in H+ efflux	(Inmaculada Meseguer and Rodriguez-Valera 1985)
H5	*ND	*ND	*ND	Narrow	*ND	*ND	(Rodriguez-Valera, Juez, and Kushner 1982)

			a		b		
H6	31-32	≤ 90°C	N	Narrow, but affects mammalian cells	TSF	Inhibition of Na$^+$/H$^+$ antiporter	(M. Torreblanca, Meseguer, and Rodriguez-Valera 1989)
HA1	14	≤ 80°C	Y	*H. larsenii* HA10, HA3, HA4, and HA9	MEP	Acts on cell membrane	(Kumar and Tiwari 2017b)
HA3	13	≤ 80°C	Y	*H. larsenii* HA10	TSF	Acts on cell membrane	(Kumar and Tiwari 2017a)
KPS1		≤ 80°C	Y	Broad (cross-domain)	TSF	*ND	(Kavitha et al. 2011)
R1	3,8	60°C	N	Broad (cross-phylum)	TSF	Archaeostatic	(Haseltine et al. 2001)
S8	3,6	≥ 24 h boiling	N	Broad (cross-phylum)	TSF	*ND	(Price and Shand 2000)
Sech7a	10.7	≤ 80°C	Y	*H.* salinarum NRC817	EEP	Disruption of cell boundary	(Pašić, Velikonja, and Ulrih 2008)
SH10	*ND	*ND	*ND	*Halorubrum* sp. BTSH03	TSF	Cell lysis	(Karthikeyan, Bhat, and Chandrasekaran 2013)

*ND: Not described
[a] Salt-dependence: Y: yes, N: no
[b] Detectable activity: TSF: transition to stationary-phase; MEP: mid-exponential-phase, EEP: early exponential-phase

1.1.3. Halocin H4: is it really a halocin?

Studies on halocin H4 were performed in *Haloferax mediterranei* H4 (ATCC 33500). Its bioactive supernatants proved to be thermolabile, susceptible to desalting and to have a lytic effect (Rodriguez-Valera, Juez, and Kushner 1982). Halocin H4 was further purified, using filtration processes, in order to partially elucidate its mechanism of action (I.

Meseguer and Rodriguez-Valera 1986; Inmaculada Meseguer and Rodriguez-Valera 1985). The sequence of *halH4* encoded in the *H. mediterranei* megaplasmid pHM300, became available in 1997 (U16389; (Cheung et al. 1997). Its upstream region encodes an ABC transporter and a peptide of unknown function and its downstream region also encodes an ABC transporter and a protein of unknown function (Figure 1).
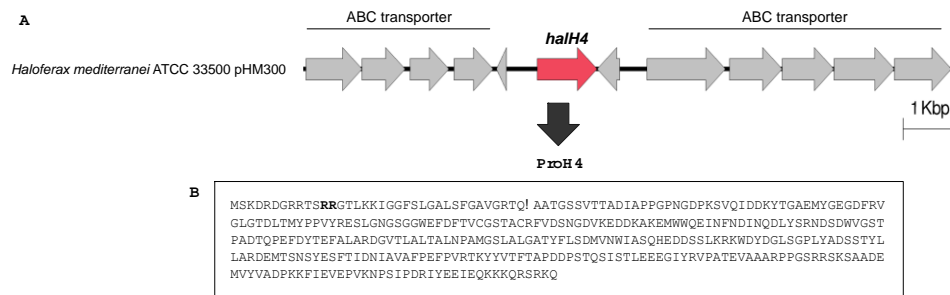
Figure 1: Representation of the genetic environment of halH4 (A) amino acid sequence of the precursor peptide ProH4 (B). The ProH4 contains a TAT-pathway signal peptide and its limit is represented with a down arrow.

Apart from *H. mediterranei*, the halocin H4 is encoded in *Haloterrigena thermotolerans* SS1R12. However, the genome of this strain is not available and, therefore, it was not possible to perform a comparative analysis of the halH4 genetic environment. More recently, a knockout of halH4 was generated and its inhibitory effect was equal to or even greater than that of the wild-type strain (Naor, Yair, & Gophna, 2013). In this study, the knockout of the halocin H4 gene, *halH4* in *Haloferax mediterranei* WR510 was performed using the pop-in & pop-out method (Bitan-banin, Ortenberg, and Mevarech 2003). The resulting strain *H. mediterranei* WR510 Δ halH4 retained its antimicrobial activity against other haloarchaeal strains, which led to the conclusion that halocin H4 is not an antimicrobial peptide (Fig 2). Thus, H4 is neither the sole antagonist nor the most potent halocin produced by *H. mediterranei* H4. The results also raise questions if H4 is indeed an halocin. A more recent study regarding halocin H4, from Chen *et al.*, 2019 also reported that *halH4* deletion mutant strains maintained their antiarchaeal profile against the *Halobacterium salinarum* NRC1 strain, proving once again that halocin H4 is not the sole effector and certainly not the main responsible for the antimicrobial character of *H. mediterranei*

Halocin H1 produced by *Haloferax mediterranei* M2a is a protein halocin of 31 kDa, very similar to HalH4 (Table 2). It is also thermolabile and salt-dependent, but so far

there is no report of a biosynthetic model or genetic environment studies for this halocin (O'Connor & Shand, 2002
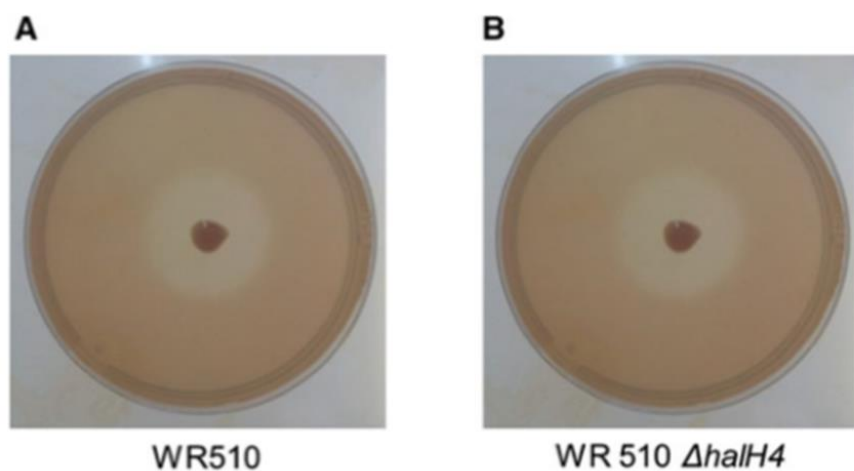
Figure 2: Antimicrobial activity displayed two different strains of *Haloferax mediterranei* against *Halobacterium salinarum*. A: *H.mediterranei* WR510 against *H. salinarum*. B: *H.mediterranei* WR510 *ΔhalH4* against *H. salinarum* (Naor, Yair, and Gophna 2013)

As above mentioned, peptide H4 is not an archaeocin and new studies shed light on its putative role in the haloarchaeal cell. The H4 peptide may be involved in the uptake of exogenous DNA by *H. mediterranei*. In the experiment conducted by Chen et al., an exopolysaccharide (EPS) deficient *H. mediterranei* strain, DF50-ΔEPS was created to facilitate DNA uptake, because EPS are known inhibitors of natural DNA uptake. From DF50-ΔEPS strain was created the DF50-ΔEPSΔhalH4 strain lacking both EPS and the halocin H4 gene. Strain DF50-ΔEPSΔhalH4 showed a decrease in DNA uptake in comparison to the EPS deficient strain, and once the gene halH4 was inserted again, the uptake increased. Scanning electron microscopy also demonstrated that the H4 peptide can cause the cell envelope to form pinholes on its surface, and once the halH4 gene is removed, the cell surface becomes smooth (S. Chen et al. 2019).

1.1.4. Halocin S8

Halocin S8 (HalS8; Figure 3C) produced by the Haloarchaeon S8a was the first microhalocin to be described. It presents the typical characteristics of these halocins, namely the fact that it can be boiled, desalted, stored at 4 °C and exposed to organic solvents without loss of activity. It is sensitive to proteinase K, but resistant to trypsin (Price and Shand 2000). The *halS8* gene was found on a 200-kbp megaplasmid and it

encodes the prepropeptide ProS8, which is composed by 311 amino acids (Figure 3A). ProS8 has a N-terminus signal peptide, typical of the Tat translocation pathway, followed by a peptide that is proteolytically processed twice to produce the mature HalS8. Thus, HalS8 is a 36 amino acid halocin preceded by a 230 amino acids (LPS8) peptide and followed by other peptide 45 amino acids long (FPS8; Figure 3A). The function of LPS8 and FPS8 peptides has yet to be elucidated, but it was hypothesized that they can be involved in regulation, export or immunity of HalS8 (O'Connor and Shand 2002; Price and Shand 2000).
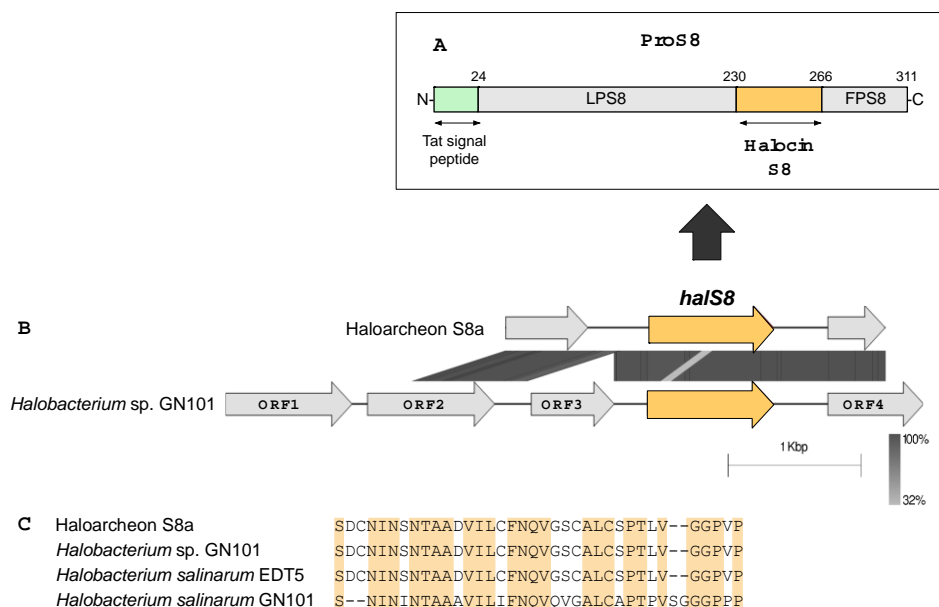
Figure 3: Halocin S8 processing, genetic environment and sequence. (A) ProS8 peptide processing: a double proteolysis originates halocin S8 and two additional peptides, LPS8 and FPS8. (B) Comparison of the genetic environment surrounding *halS8* in Haloarchaeon S8a and in *Halobacterium* sp. GN101. (C) Multiple alignment of the halocin S8 peptide sequence from haloarcheon S8a, *Halobacterium* sp. GN101, *Halobacterium salinarum* EDT5 and *Halobacterium salinarum* GN101.

The biosynthetic gene cluster of this halocin in the Haloarchaeon S8a has not yet been characterized. However, the *halS8* gene is also found in *Halobacterium salinarum* ETD5 (accession: KR611165) and ETD8 strains (accession: KR611166) as well as in a megaplasmid of (accession: from nc 129975 to 130928 of EU080936). The *Halobacterium* sp. GN101 megaplasmid 2 is completely sequenced and it allowed us to analyse the genetic environment of *halS8* (Figure 3B). In *Halobacterium* sp. GN101, *halS8* is preceded by a putative transposase of the IS6 family (ORF3) and two proteins of unknown function (ORF2 and ORF1). According to Psortb, the protein encoded by ORF1 should be cytoplasmic and the subcellular location of the one encoded by ORF2 is unknown. However, ORF2 protein analysis in TMHMM server showed that its N-terminus has a

cytoplasmic domain, followed by two transmembrane helices and an extracellular domain of 252 aa. Downstream of *halS8* there is the ORF4 that encodes a protein with unidentified molecular function that should be localized in the cytoplasmic membrane. This protein has 2 transmembrane helices, followed by a cytoplasmic region of 100 amino acids and then other 2 transmembrane helices.

The Halocin R1 (HalR1) produced by *Halobacterium salinarum* GN101 is a HalS8-like halocin since it is 63% identical and 71% similar to HalS8 (Figure 3C;O'Connor and Shand 2002). However, these two halocins have different physical properties since HalR1 is not as stable as HalS8 (it only withstands temperatures up to 60 °C; Table 2) and has a different spectrum (it is archeostatic rather than archeocidal). The key factor of these differences can be the different number of Cys residues: HalS8 has four Cys that can form disulfide bridges, thus increasing halocin stability, whereas HalR1 has only one (O'Connor and Shand 2002)

1.1.5. C8 and C8-like halocins

Halocin C8 (HalC8) is a 76 amino acids long microhalocin (Figure 4), with a molecular weight of 7.427 kDa (calculated 7440.7 Da). Compared with other halocins, HalC8 has a broad spectrum of activity (Table 2). It can be desalted, boiled up to 100°C for 1 hour or stored at 4 °C for a period longer than 1 year, without losing its bioactivity (Li et al. 2003). HalC8 is rich in cysteine residues (Figure 5), which form disulphide bridges that surely contribute for its stability under harsh conditions (Besse et al. 2017). The gene *halC8* encodes a 283 amino acid preproprotein (ProC8) of 29.38 kDa. ProC8 undergoes proteolysis by an unknown mechanism and originates: i) a N-terminus immunity protein HalI, a hydrophilic polypeptide (207 amino acids) with a short hydrophobic region at its N-terminus region and ii) a C-terminus acidic and hydrophobic microhalocin HalC8 (Figure 4). The HalC8 biosynthetic cluster was first described by Li *et al.* (2003) in *Natrinema* sp. AS7092 (formerly known as *Halobacterium* sp. AS7092). In addition to *halC8*, it also includes (Figure 4): i) *halU* that encodes a protein with 12 transmembrane helices with unknown function, ii) *halR* that encodes a DNA-binding transcriptional regulator, iii) *halT1* that encodes a transmembrane protein, iv) *halT2* that encodes a protein with a secretion signal peptide and v) *halT3* that encodes and ABC transporter. The *halC8* gene cluster was detected in other haloarchaeal genomes: in 6 *Natrinema* spp. and in 1 *Haloterrigena* spp. (Table S1 and Figure 4B). These strains also have the other *hal* genes, except the cluster of *N. gari* JCM 14663 that misses *halR* (Figure 4B). It is also worth to mention that *N. gari* JCM 14663 HalC8 is shorter (Figure 4 and Figure 5). Besse *et al.* (2017) were able to amplify the *halC8* gene in other *Natrinema* spp, including *N. ejinorense*, *N. salaciae* and

N. versiforme. We did not identify *halC8* in the draft genomes available for these three species (Figure 4C). However, *halT1*, *halT2* and *halT3* were detected (Figure 4C). This finding casts doubt on the requirement of *halT1*, *halT2* and *halT3* for the biosynthesis of HalC8, which is also supported by the fact that *halT2* and *halT3* are not present in *H. salina* JCM 13891 cluster (Figure 4B).
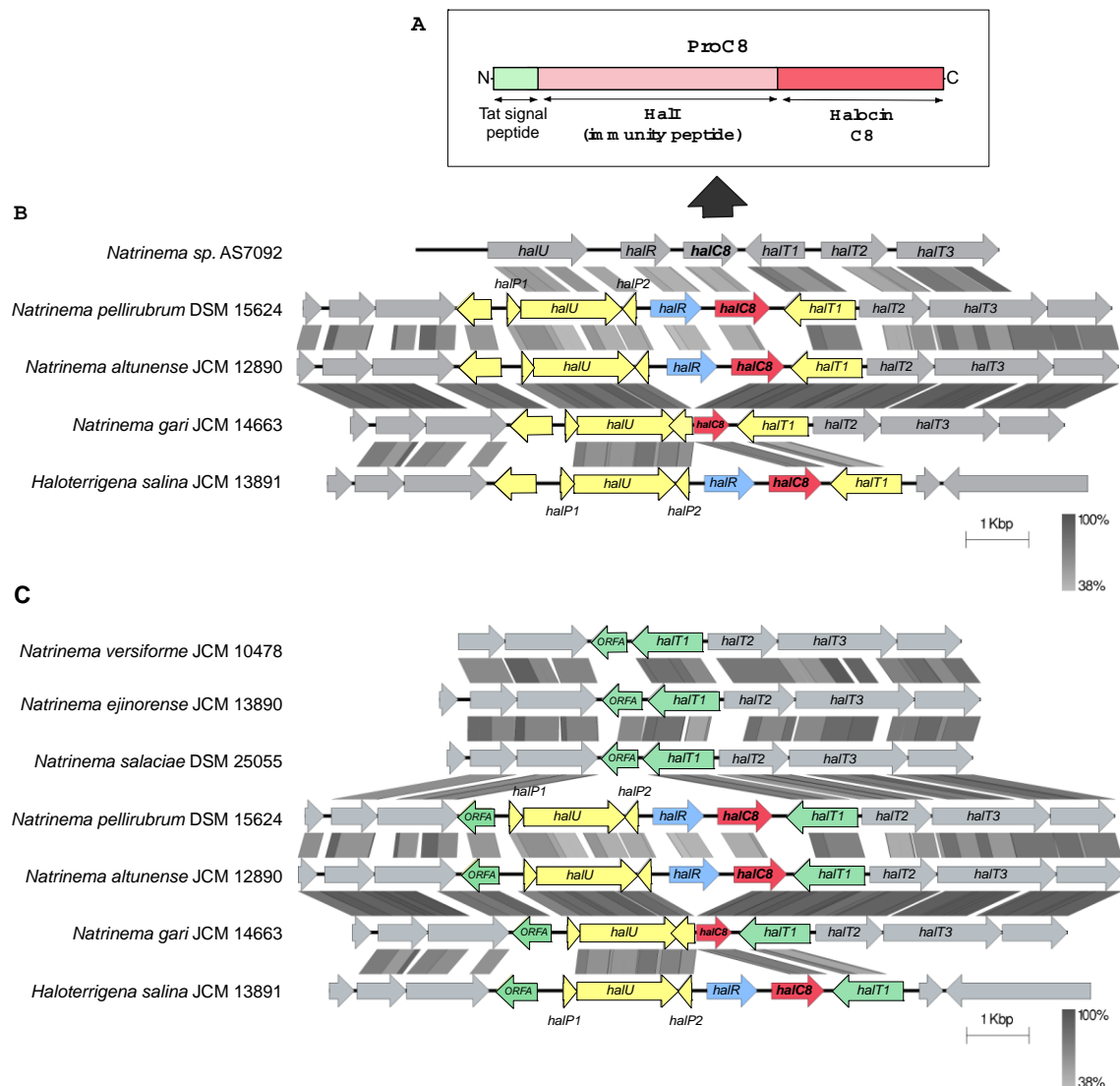
Figure 4: (A) *halC8* gene encodes the ProC8 peptide possessing a Tat signal peptide, the immunity peptide HalI and the halocin HalC8 (B). Comparative genomics of *halC8* gene of diverse haloarchaea and the original HalC8 producer (*Natrinema* sp. AS7092; (C). Comparison of the upstream and downstream regions of *halC8* biosynthetic cluster among different *Natrinema* spp. and *Haloterrigena salina* JCM 1389 showing that *halT1*, *halT2* and *halT3* are not preceded by *halC8*, *halR* and *halU* in *N. versiforme*, *N. ejinorense* and *N. salaciae* (C). *halC8* gene was amplified from these three strains by Besse *et al.* (2017) but it is not found in the same genetic environment of the other *halC8*-containing *Natrinema* spp. Some of the strains included in Table S 1 were not included in this analysis either because they were analogous to some of these gene clusters or because their genomes were not annotated (Inês Castro, Mendo, and Caetano n.d.).
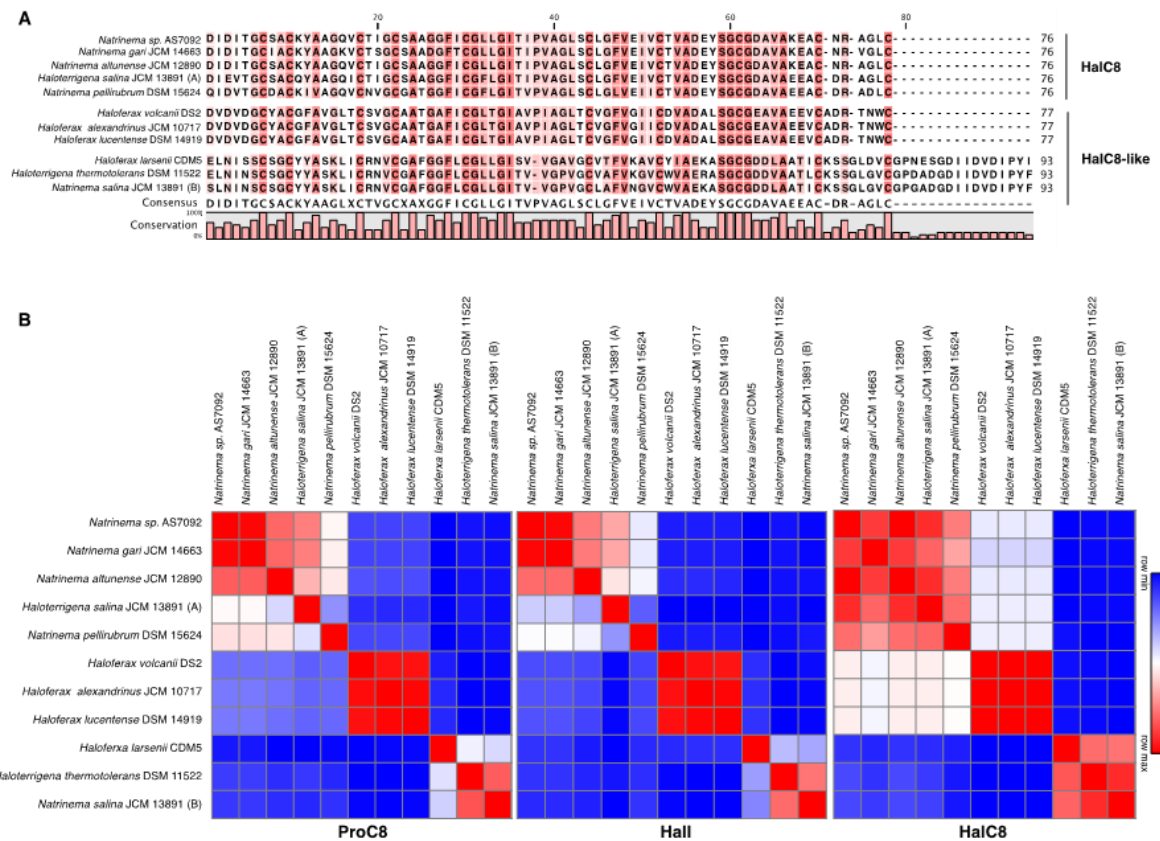
Figure 5: Alignment of HalC8 and HalC8-like peptides of Halobacteria with conserved residues colored in red (A) and heat map of ProC8, HalI and HalC8 global similarity matrix (B) generated with SIAS and visualized with Morpheus (https://software.broadinstitute.org/morpheus).

Other peptides with less than 51% similarity with ProC8 were detected in other Haloarchaea including *Haloferax* spp. (Besse et al. 2017). Their HalI and HalC8 peptides share less than 43% and 72% similarity with those of *Natrinema* sp. AS7092, respectively (Figure 6 and Figure S1). Global similarity of HalC8 analysis reveals that these peptides can be divided in three groups: i) the highly similar to *Natrinema* sp. AS7092 HalC8, ii) the *Haloferax* spp. HalC8 and iii) the HalC8 of *H. larsenii*, *H. thermotolerans* and *N. salina* (Figure 5B). Based on amino acid sequence, we propose that group ii) and iii) as HalC8-like peptides whose halocin activity has yet to be proven. There is no evident conservation amongst the *halC8* genomic environments of the three groups, indicating that HalC8 biosynthesis may be distinct from that of *Natrinema* sp. AS7092 (Figure 6). Halocin A4 is a microhalocin produced by the euryarchaeal halophilic strain TuA4. Its N-terminal sequence shares 75% identity with HalC8 (Besse et al. 2015). Therefore, it is likely that A4 is also a HalC8-like halocin. It is heat-stable and retains activity when desalted (Table 1). A4 is interesting since it has a cross-phylum spectrum of activity (Table 2).
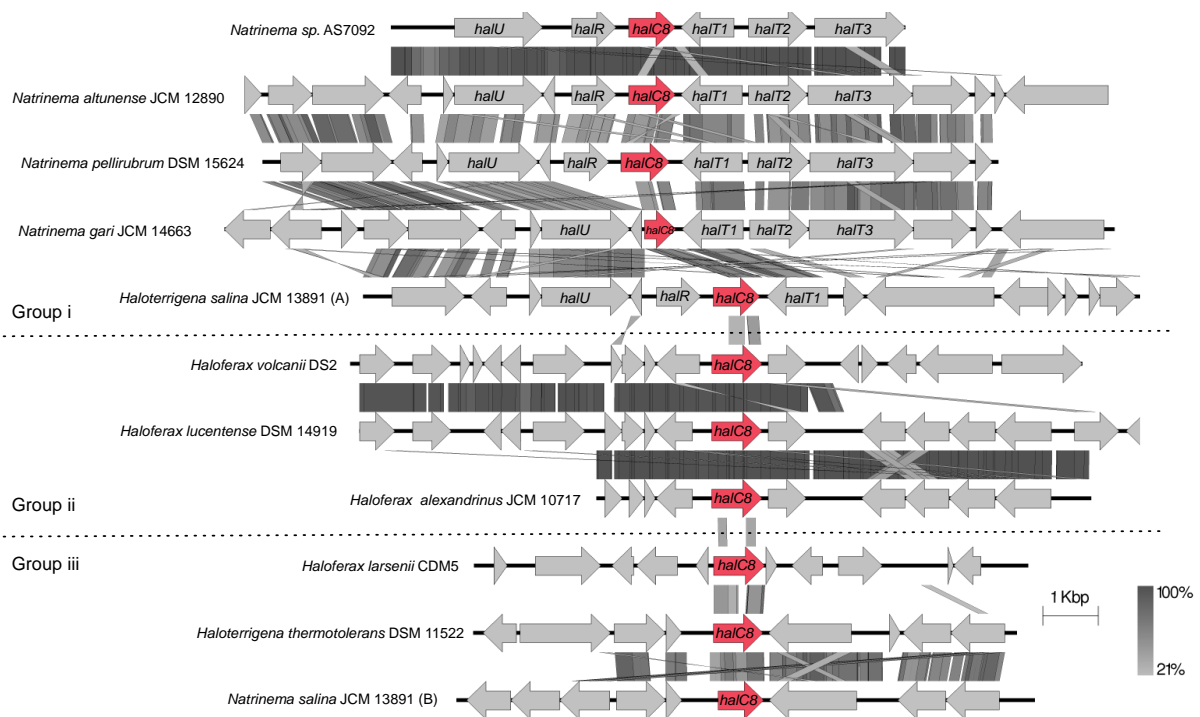
Figure 6: Comparative genomics of the genetic environment of *halC8* gene of the original HalC8 producer (*Natrinema* sp. AS7092) and other haloarchaea were *halC8* or *halC8*-like genes were identified. The image was produced with Easyfig 2.1 applying a blast minimal length of 50 (Inês Castro, Mendo, and Caetano n.d.).

1.1.6. Other halocins

Halocin H6 is a 31-32 kDa protein halocin produced by the *Haloferax gibbonsii* Ma2.39 strain. Unlike the other protein halocins, H4 and H1, it is heat resistant, retaining its activity after exposure to 100 °C and can be desalted. The robust character of this halocin is typical of microhalocins but its size is not and no additional studies have been done to clarify its biosynthesis (Price and Shand 2000).

There is also some confusion about this halocin, which is named H6 or H7 when it is produced by *Haloferax gibbonsii* Ma2.39 or its mutant, *Haloferax gibbonsii* Alicante SPH7, respectively. *Haloferax gibbonsii* Alicante SPH7 is a mutant that over-expresses H6 (O'Connor and Shand 2002). The halocin is the same, but it is produced in higher amounts and therefore should not take other designation. Apart from the H4, S8 and C8 halocins, the peptide sequence or the genetic context of the other halocins is unknown, and as such, it is difficult to include them in one category.

1.1.7. Model of biosynthesis

The biosynthetic gene cluster of the majority of halocins is unknown. Furthermore, halocins seem to be a very heterogeneous group of compounds, which will make it difficult to establish a common biosynthetic mechanism. So far, the HalC8 is the most well-known and characterized model. As abovementioned, HalC8 is encoded by a precursor peptide ProC8 that also encodes the immunity peptide HalI (Figure 4A). ProC8 contains a N-terminus Tat signal peptide with the classical tripartite structure and the twin-arginine motif (Figure S1). Therefore, ProC8 is a substrate of the Tat secretion pathway in which the cleavage site corresponds to Ala39 (Sun et al. 2005; Figure S1). Based on this, the most probable biosynthetic model includes the export of ProC8 peptide by the Tat translocation pathway, originating the HalI:HalC8 peptide (Figure 7).
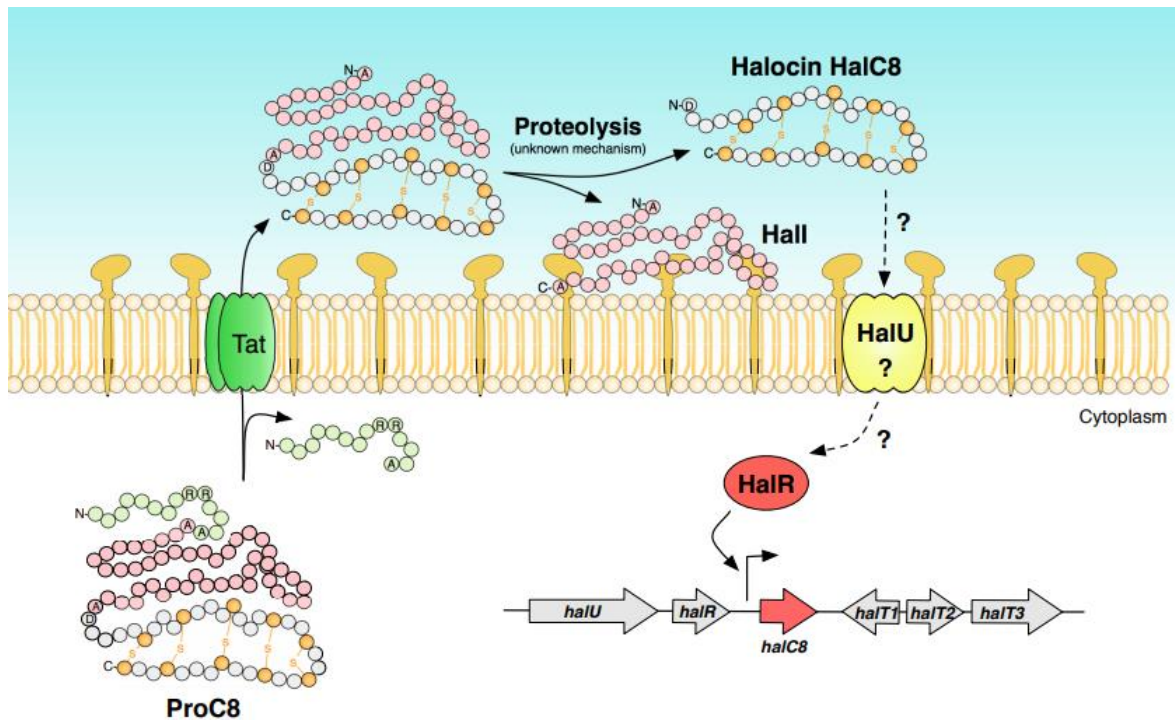
Figure 7: Halocin C8 biosynthesis model. The ProC8 peptide is exported via Tat-pathway. The processed peptide then suffers proteolysis via an unknown mechanism originating both halocin C8 and the immunity peptide HalI; then, HalC8 is released to the environment, and the immunity peptide HalI binds to the cell membrane.HalU is a membrane bound protein of unknown function and that could be involved in *halC8* genetic regulation (Inês Castro, Mendo, and Caetano n.d.).

The Tat pathway is capable of translocating prefolded proteins without disruption of the membrane and in *Halobacterium* spp. this is the main secretion pathway. The export of proteins mainly by this route is considered an adaptive response to the high

intracellular (~4 M KCl) and extracellular (~4 M NaCl) salt conditions (Dilks, Giménez, and Pohlschröder 2005; Rose et al. 2002). Once in the extracellular environment, the HalI:HalC8 peptide will undergo proteolysis by an unknown protease to give rise to the mature HalC8 and the immunity peptide HalI (Figure 7). The putative *halC8* cluster also encodes the HalR protein (269 amino acids) that has a winged helix-like DNA-binding domain at its N-terminus (approximately the first 100 amino acids). No transmembrane helices or signal-transducing and sensing domains were detected in HalR. Thus, HalR is not a one-component system (OCS), the predominant class of regulatory systems in both Bacteria and Archaea (Ulrich and Koonin 2005), or, alternatively it is an OCS with novel and uncharacterized signal-transducing and sensing domains.

The archaeal transcription machinery (including the general transcription factors) resembles that of eukaryotes. However, with few exceptions, their transcription regulators are like those of bacteria. Almost all archaeal transcription regulators known are repressors and some of them act by binding to inverted repeats of the promoter (Karr 2014). The *halC8* promoter has one of such sequences upstream of its B-recognition element (BRE) and of the TATA box (Sun et al. 2005), which can be the target of HalR. In fact, the transcription of *halC8* gene was found to be upregulated when cells enter the stationary phase (Sun et al. 2005). The involvement of the other proteins encoded in the *hal* gene cluster in the biosynthesis of HalC8 is yet to be discovered.

1.1.8. Self-protection/Immunity mechanisms

Just like other producers, for example, of bacteriocins and eucariocins, halocin producers must survive to the action of their own halocin by encoding mechanisms of self-protection/immunity. To the best of our knowledge, this mechanism was investigated only for HalC8 and it seems unique because one protein (ProC8) will originate both the immunity polypeptide (HalI) and the halocin (HalC8) (Figure 4A; Figure 7). The co-expression of these elements can be cost saving to the producer strain.

Concerning the mechanism, it was found that HalI is associated with the cytoplasmic membrane and binds directly to HalC8, inhibiting its action (Sun et al. 2005). *In vitro*, HalI is able to capture several molecules of HalC8 due to the ability of HalC8 to form oligomers, forming a large heterocomplex (Mei et al. 2008). *In vitro* and *in vivo* studies demonstrated that the helix-loop-helix (HLH) motif located at the N-terminal region of HalI is the key feature for the self-protection mechanism. Site-directed mutagenesis targeting one or both helices of the HLH motif resulted in an almost complete loss of HalI-HalC8 binding ability and immunity. In this context, it was established that the Leu64 and Leu73 residues of HalI (Figure S1) are fundamental for its immunological activity (Mei et

al. 2008). These two residues are located in the hydrophobic region of the HLH motif. Thus, it is proposed that HalI interacts with HalC8 via this HLH motif due to the molecular force of the hydrophobic interaction, resembling the interaction of colicin E1 and its immunity protein. HalI also binds specifically and intercepts HalC8 at the surface of the cytoplasmic membrane, a mechanism highly similar to that of the immunity of the lantibiotic nisin (NisI:nisin complex) (Mei et al. 2008)

1.1.9. Bioactivity of halocins - targets and mechanism of action

Most of the cross-inhibition testing among haloarchaeal strains was done using the double-layer agar method, where clear inhibition halos correspond to halocin production (Marina Torreblanca, Meseguer, and Ventosal 1994). Most halocin are archaeocidal (inhibit the growth by killing), but halocins with archaeostatic effect were also reported (Table 2).

Some halocins have a broad range of activity, and even cross-phylum inhibition. Halocins A4, S8 and R1, display antiarchaeal activity against strains belonging to the genus *Sulfolobus*, a hyperthermophilic Crenarchaeota.

Halocin A4 has a cytocidal effect and halocins S8 and R1 have a cytostatic effect (Haseltine et al. 2001). The halocin KPS1, produced by *H. volcanii* KPS1 has a cross-domain activity. It inhibits not only the growth of the haloarchaeal *Halobacterium sodomens*, but also Gram-negative and Gram-positive bacteria such as *Bacillus subtilis*, *Streptococcus mutans, Staphylococcus aureus, Escherichia coli* and *Pseudomonas aeruginosa* (Kavitha et al. 2011). As soon as halocins were discovered, experiments with the goal of describing halocins' specific mode of action against other microorganisms quickly began.

To my knowledge, halocin H6 is the only halocin whose mechanism has been described in more detail. In a preliminary experiment, it was observed that its target is the archaeal cell membrane. Specifically, it was shown that the effect was due to the ion gradient across the membrane because, after exposure to HalH6, the sensitive cells lost their cytoplasmic content, becoming "ghosts"(M. Torreblanca, Meseguer, and Rodriquez-Valera 1990). After exposure to HalH6, sensitive strains showed changes in their intracellular volume and pH, in membrane potential, proton motive force and sodium and proton flux on membrane (O'Connor and Shand 2002). HalH6 had the same effect as dicyclohexylcarbodiimide (DCCD), a substance known to inhibit the archaeal Na+/H+ antiporters (Inmaculada Meseguer, Torreblanca, and Konishi 1995). Na+/H+ antiporters or exchangers (NHE) are also found in mammalian cells and are responsible for the regulation of intracellular pH levels. They play an important role on ischemia-reperfusion

injury, a condition known to cause heart failure. The ischemia-reperfusion injury results from ischemia and causes the activation of NHE, which leads to intracellular acidification and, consequently, increases the intracellular levels of sodium. The high levels of intracellular Na^+ cause a reversion in the Na^+/Ca^{2+} exchangers that normally remove calcium from the cells and in this condition accumulate it, resulting in an increase in intracellular Ca^{2+}. After ischemia, during reperfusion, the NHE cannot be inactivated resulting in critical levels of intracellular Na^+. One solution for this disease is the inactivation of the NHE, an activity that is promoted by HalH6 in haloarchaea. And, in fact, some experiments showed that HalH6 is also able to inhibit the NHE of human cells, *in vitro*. Additionally, HalH6 displayed an *in vivo* protective effect on damaged myocardium tissue of dogs caused by ischemia-reperfusion injury (Lequerica et al. 2006).

Halocin HalC8 seems to have the archaeal cell wall as primary target. Archaeal sensitive cells treated with this halocin begin to swell and release their cytoplasmic content due to cell wall disruption. After prolonged exposure most of the cells burst (Li et al. 2003). This hypothesis is also supported due to HalC8 oligomerization, a characteristic often observed in pore-forming antimicrobial peptides (Mei et al. 2008). However, investigation of halocin's activity has its challenges. It is difficult to determine whether the damage observed is a direct or a secondary effect and large amounts are required to cause significant changes in cell's morphology (M. Torreblanca, Meseguer, and Rodriquez-Valera 1990).

Moreover, interactions between halophilic archaea and halophilic bacteria were evaluated using culture supernatants from different saline environments from around the world in a cross-study where all strains were tested against each other (Atanasova, Pietilä, and Oksanen 2013). In this study, archaea inhibited much more bacterial strains than the other way around. In fact, only one strain of *Halorubrum* sp., a haloarchaea, was sensitive to halophilic bacteria supernatants. *Haloferax* sp. strains were the ones with the broader inhibitory spectrum and showed higher antimicrobial activity, demonstrated by the larger inhibition halos produced. Moreover, along with an *Haloplanus* strain, *Haloferax* sp. strains were resistant to all the supernatants tested. Interestingly, *Haloferax* sp. s5a-1 supernatants had a different spectrum activity depending on the previous protease treatment, indicating that this strain may produce different halocins (Atanasova, Pietilä, and Oksanen 2013). In the future, an effort should be made to use analytical and spectrometry techniques to characterize the peptides responsible for inhibition phenotypes since they are currently more readily available and user friendly.

1.1.10. Ecological role of halocins

One of the main challenges for the detection of their activity is the fact that the secreted halocins can be easily incorporated by the sensitive cell or cellular debris, making it difficult to detect (Price and Shand 2000). However, experiments using strains collected from several high-salinity environments such as salterns and saline lakes revealed that halocin production is ubiquitous among haloarchaeal and, therefore, they should play an important role in their ecology (Marina Torreblanca, Meseguer, and Ventosal 1994). In line with bacteriocins, it has been hypothesized that halocins are involved in competition amongst haloarchaeal strains (Kis-Papo and Oren 2000). To test the latter, *in vivo* studies were conducted. The halocin activity of processed saltern brine samples was analyzed, but no strain was inhibited by the environmental samples. Thus, it was not possible to draw clear conclusions because: i) ecologically, the production of halocins may not be involved in haloarchaeal competition or ii) the levels of production in their natural habitat can be very low, making their detection in the laboratory difficult or iii) activity may be lost due to downstream processing such as filtration or proteolysis (e.g. halophilic proteases excreted by some strains) (Kis-Papo and Oren 2000).

Studies regarding the production-sensitivity interactions of halophilic strains of bacteria and archaea from different saline environments around the world demonstrated that the majority of the interactions were observed between strains of geographical distant habitats, meaning that habitat conditions dictate the diversity available in a specific environment (Atanasova, Pietilä, and Oksanen 2013). Furthermore, the sensitivity that some producer strains display to their own halocin could be a type of cell density control mechanism (Atanasova, Pietilä, and Oksanen 2013).

1.2. Lanthipeptides

1.2.1. Introduction

Lanthipeptides are natural products part of the RiPPs family, which are ribosomally synthesized peptides that undergo post-translational modifications (PTMs). These peptides can often be referred as lantibiotics when they display antimicrobial activity (Repka et al. 2017). Lanthipeptides main characteristic is their lanthionine (Lan) and/or methyl-lanthionine (MeLan) amino acids that are formed through thioethercross-links, hence their name. Biosynthetic enzymes are responsible for the formation of these thioethercross-links, that are induced post-translationally by essentially dehydration and cyclization of residues. Dehydration of Ser residues to 2,3-didehydroalanine (Dha) or Thr to -2,3-didehydrobutyrine (Dhb) allows a thioether bridge formation by the means of a Michael-type addition of a cysteinyl thiol (Knerr & van der Donk, 2012). The thioether cross-links are the main contributors to the activity, overall stability and bioactivity of the lanthipeptide (Knerr and van der Donk 2012).

Common to all lanthipeptides is the biosynthesis of the precursor peptide, designated LanA. It is synthesised by the ribosome and is composed of two parts, the leader peptide that directs the installation of PTMs and the core peptide where dehydration and cyclization occur to form the mature peptide, able to be exported. (Fig. 8; Repka et al., 2017)
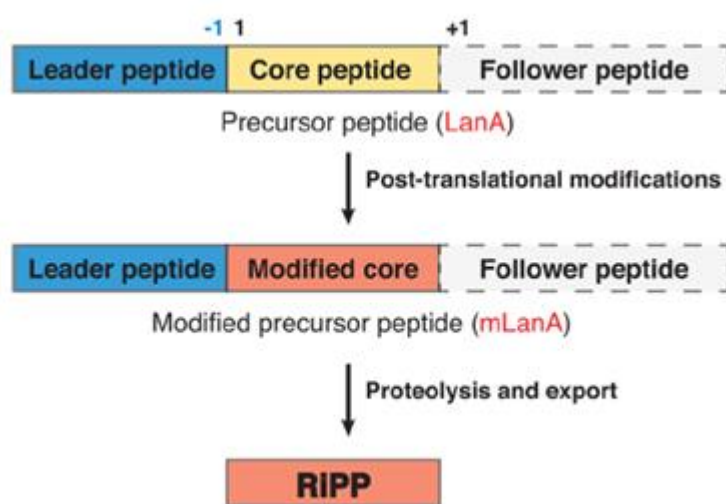
Figure 8: Representation of the RiPPs cascade of events leading to the final mature peptide. Adapted from Repka et al., 2017

1.2.2. History

The biotechnological potential of lanthipeptides was first introduced in the 1950s, with the discovery of nisin's potential as a food preservative, preventing microbial growth with no apparent microbial resistance occurrence (Knerr and van der Donk 2012). At the time, lanthipeptides were thought to be strictly produced by gram-positive bacteria. Nowadays, genome-mining technologies coupled with facilitated genome sequencing techniques allowed the emergence of potentially new lanthipeptide producers as well as the discovery of new pathways, a better understanding of biosynthetic machinery and their mode of action (Knerr and van der Donk 2012). Lanthipeptide's biosynthetic enzymes are found in all branches of life, ranging from archaea to higher eukaryotes. Nevertheless, actual protein activity was only detectable and characterized in bacteria (Repka et al. 2017). Lantibiotics also pose as attractive alternatives to classic antibiotics, with demonstrated activity against drug-resistant gram-positive bacterial strains like *Staphylococcus*, *Streptococcus*, *Enterococcus*, and *Clostridium* and also some gram-negative pathogens like *Neisseria* and *Helicobacter* (Knerr and van der Donk 2012). Although clinically interesting, lanthipeptide antimicrobial activity is not strictly antibacterial, some having also antifungal, morphogenetic, antiviral, antinociceptive and antiallodynic properties (Repka et al. 2017). Their mode of action is still to be unravelled but preliminary studies reveal that these antibacterial peptides can either inhibit cell wall biosynthesis or disrupt the cytoplasmic membrane via pore formation (Knerr and van der Donk 2012).

1.2.3 Classification of lanthipeptides

As above mentioned, the core peptide LanA undergoes thioether cross-links formation through essentially dehydrations and cyclizations by biosynthetic enzymes (Knerr and van der Donk 2012). Lanthipeptides are classified according to these biosynthetic enzymes (Fig 9). Class I lanthipeptides encompass in their biosynthetic cluster a LanB dehydratase, responsible for the serine/threonine dehydration and a LanC cyclase in charge of the thioether cyclization (Repka et al. 2017). Class II contains a lanthipeptide synthetase with bifunctional activity, LanM, with both dehydratase and cyclase domains located N-terminally and C-terminally, respectively. In class III, all reactions are carried out by a trifunctional enzyme, LanKC, which has three domains: i) a N-terminal lyase domain, ii) a central kinase domain, and iii) a putative C-terminal cyclase domain lacking the conserved zinc-binding motif (Cys-Cys-His/Cys) found in LanC, LanM and LanL (Zhang et al. 2012). Lastly, in class IV lanthipeptides, a trifunctional enzyme

designated LanL is responsible for dehydration and cyclization reactions. LanLs have the same N-terminal and central domains as LanKC but its cyclase domain is homologous to the one found in LanC and LanM (Repka et al. 2017). Class II lanthipeptides will be discussed in more detail in frame with the objectives of this thesis.
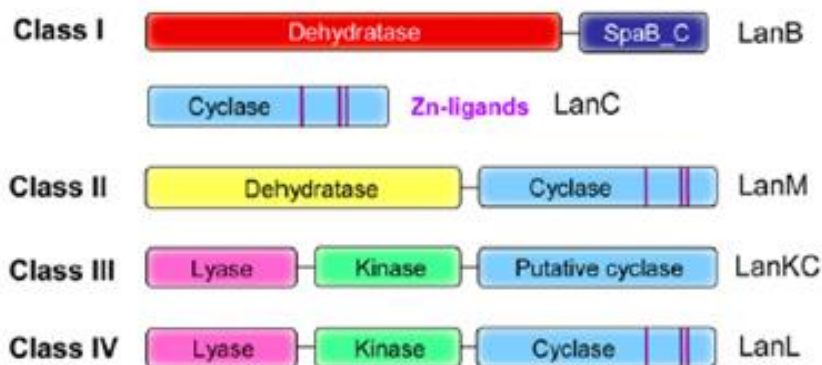
Figure 9: Schematic representation of all four classes of lanthipeptides described to date. Abbreviations: LanB, lantipeptide dehydratase; LanC, lantipeptide cyclase; LanM, class II lantipeptide synthetase; LanKC, class III lanthipeptide synthetase; LanL, class IV lantipeptide synthetase; SapB-modifying synthetase (Zhang et al. 2012)

1.2.4. Class II lanthipeptides

In this class, the bifunctional synthetase, LanM is responsible for executing all dehydration and cyclization reactions. The length of LanMs can vary, ranging from 900 to 1200 residues. These enzymes contain two domains, a dehydratase domain located N-terminally, and a cyclase domain located C-terminally. Although they share the same functional domain, the LanM dehydratase domain is not homologous to the one found in the class I LanB synthetase. Interestingly, the LanM cyclase domain and the LanC cyclase domain share a 25% similarity and they both contain zinc-ligands, thus being homologous (Knerr and van der Donk 2012).The dehydration and cyclization reactions are independent but not sequential, i.e. the cyclization is initiated while the dehydration process is still occurring (Repka et al. 2017).

1.2.4.1. Dehydration & Cyclization in class II

The dehydration events in class II lanthipeptides, are composed of two main steps: phosphorylation and posterior phosphate elimination (Fig 10; Repka et al., 2017). Studies conducted on the class II synthetase, LctM (involved in the biosynthesis of lacticin 481), demonstrated that the dehydration occurs via phosphorylation of serine/threonine followed

by the elimination of a phosphate ester that results in the formation of Dha/Dhb by the means of ATP and Mg^{2+} (Chatterjee et al. 2005).
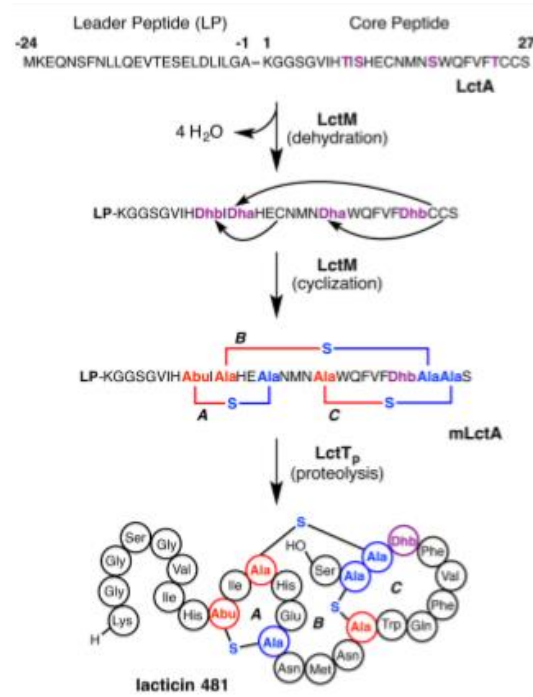
Figure 10: Representation of biosynthesis of the class II lanthipeptide lacticin 481 produced by *L. lactis*. (Repka et al. 2017).

Mutagenesis assays on LctM, revealed the importance of Asp242 and Asp259 in the interaction and regulation of the phosphorylation reaction, Arg399 and Thr405 in elimination of phosphate and Lys159 is involved in both reactions (You and van der Donk 2007). Another well studied lanthipeptide class II synthetase is CylM, the synthetase of cytolysin, which is a two-peptide lantibiotic, meaning that it can only exert its function when coupled to another lanthipeptide. Up to date, the report of two-peptide lantibiotics (also referred as two-component peptides) is restricted to class II lanthipeptides (Repka et al. 2017). In this group, one peptide, the α-peptide, binds to the lipid II component of the bacterial target cell while the other component, the β-peptide, binds to the previously formed complex and can form pores in the in the membrane (Fig 11; Repka et al., 2017).
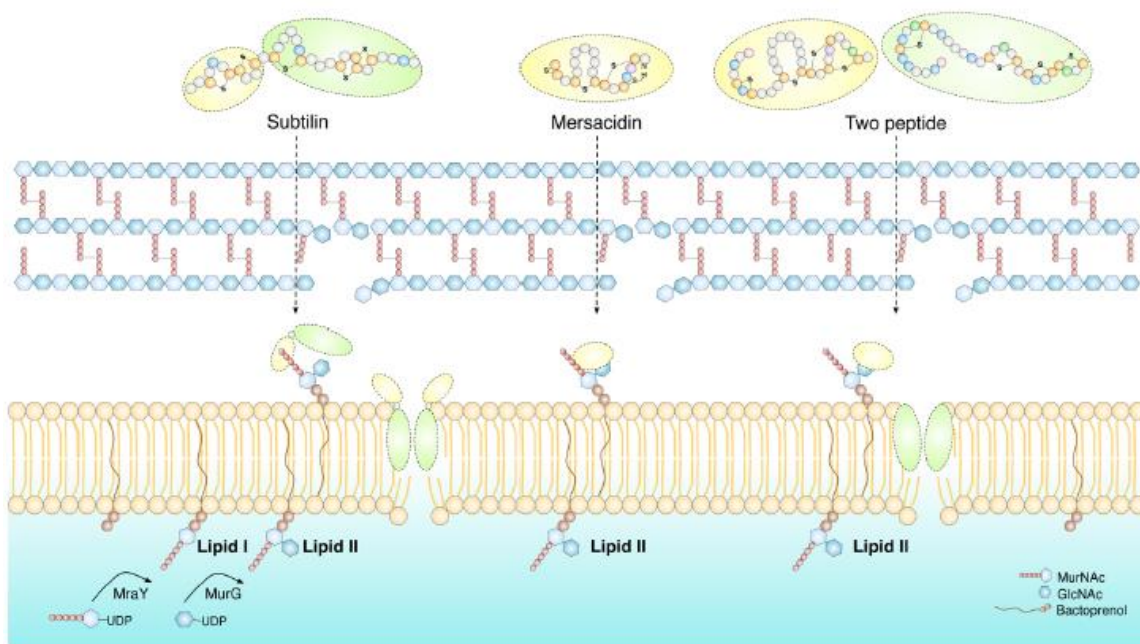
Figure 11: Schematic representation of the mode of action of different lantibiotics (Barbosa, Caetano, and Mendo 2015).

As previously mentioned, the dehydratase domain of LanMs bears no homology to the one found in the dehydratase LanB from class I, but rather resembles the lipid kinases and lipid-kinase like proteins found in eukaryotes (Dong et al. 2015). CylM also has important residues involved in the phosphate elimination step and important subdomains, like the KA (kinase-activation) domain. Arg 506 and Thr512 are located in the KA domain, and responsible for maintaining the stability of the activation loop and supplying catalytic residues to the active site for phosphate elimination. Asp252 and His254 are involved in the substrate phosphorylation step (Dong et al. 2015).

After dehydration, the precursor peptide will have several dehydro amino acids as well as cysteine residues. In theory, if all cysteine residues can form thioether rings, it means that a myriad of isomers can be potentially created. However only one isomer is produced, meaning that the cyclization process is a tightly selective and regulated event. In fact, even the current synthetic chemistry technologies cannot always fully replicate it (Repka et al. 2017). Studies on the synthetase HalM2, involved in the biosynthesis of haloduracin β-peptide revealed that its interaction with the leader peptide is essential during cyclization (Yang and Van Der Donk 2015). The mechanism of cyclization is similar to the one of LanC, and it is hypothesized that the Zn^{2+} ion conducts the activation of cysteine thiols (Paul, Patton, and van der Donk 2007). Although LanM C-terminus and LanC share a similar mechanism and are even homologous, phylogenetic studies on the

evolution of both lanthipeptide synthetases revealed different clades and a distant kinship, meaning that these enzymes most probably evolved independently (Zhang et al. 2012).

1.2.4.2. Cleavage and transport in class II

After the precursor (LanA) undergoes dehydration and cyclization, the modified peptide (mLanA) suffers proteolysis and transport (Fig. 10). In class II lanthipeptides, this final steps are usually carried out by the means of a LanT$_P$ bifunctional enzyme with both protease and transport domains belonging to the ABC-transporter maturation and secretion (AMS) protein family (Repka et al. 2017).

In general, the leader sequence is cleaved in a conserved domain of two glycine residues (GG-motif) (Havarstein, Diep, and Nes 1995). The GG-motif is essential in this step, as studies performed on the gene *mutA*, which encodes the precursor peptide of the lantibiotic mutacin II, demonstrated that mutation of this motif translates in loss of proteolytic processing and export (P. Chen et al. 2001). In class II two-component lantibiotics, as is the case of lichenicidin, a bifunctional enzyme LicT$_P$ removes the majority of the leader peptide from Licβ, leaving an hexapeptide, NDVNPE attached to Licβ. The hexapeptide is then cleaved from the core peptide by a protease, LanP to generate the mature peptide (Fig 12; Caetano et al. 2011).
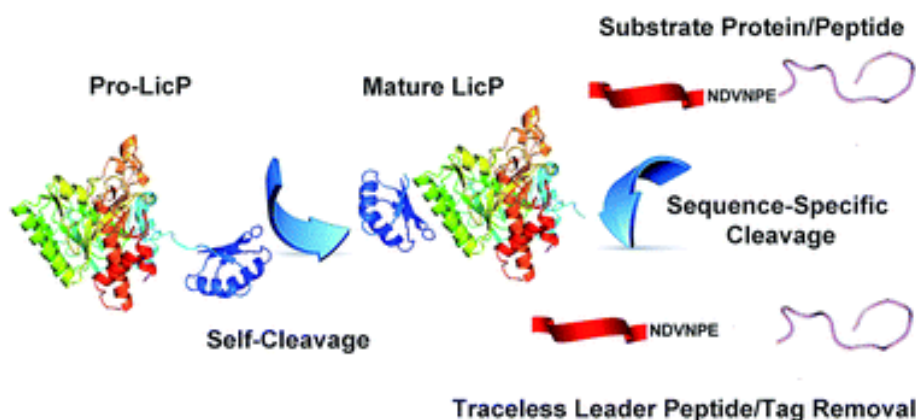
Figure 12: Proteolytic processing of the class II two-component lantibiotic lichenecidin. The unprocessed peptide, Pro-LicP undergoes a first cleavage, generating the mature LicP followed by a another cleavage, removing the hexapeptide NDVNPE (adapted from Tang et al. 2015).

1.2.4.3. Immunity/Self-protection of lantibiotic producers

As previously described, lantibiotics seem to target lipid II in the cytoplasmic membrane (Fig. 11) To prevent the producer from its own "toxicity", lantibiotic producers develop immunity mechanism to protect their cells from its own peptides. In the biosynthetic cluster of class II lanthipeptides are often encoded transporters like LanEFG and LanI-like proteins that are involved in immunity (Fig. 13). LanEFG constitutes an ABC type transport system that is responsible for the efflux of lanthipetides. The LanI proteins excert their immunity behavior by sequestering the extracellularly secreted product (Repka et al. 2017). Sometimes, a third intervenient in lantibiotic immunity can be present, as it's the case of the accessory protein LanH. LanH has been described as an ancillary protein that aids the assembly and proper function of the ABC transporter. Because it's an accessory, LanH proteins are usually not present in lantibiotic associated biosynthetic gene clusters (Hill et al. 2008).
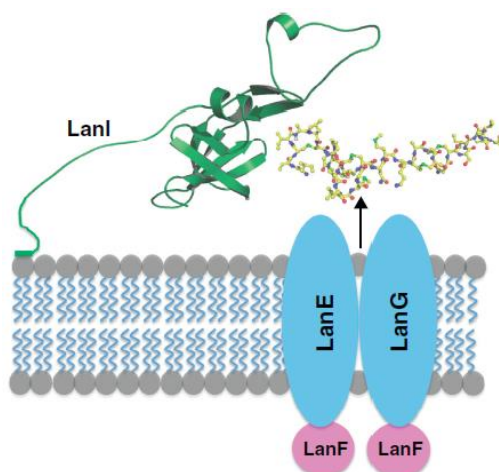
Figure 13: Schematic representation of the enzymes generally involved in the immunity mechanism of lantibiotic producing strains (adapted from van der Donk and Nair 2014).

Most of the early information provided on lantibiotic immunity was obtained from studies on the class I lantibiotic nisin. Nisin immunity is provided by the combination of NisFEG and NisI action (Hill et al. 2008). Nowadays, immunity mechanism have been described for several lantibiotics, like subtilin, epidermin, Pep5, nukacin ISK-1, mersacidin, cytolisin and lacticin 3147 (Hill et al. 2008).

Although some LanEFG and LanI proteins from different lantibiotic producers show some high similarity between them, events of cross inhibition between different producers are rare, which indicates that the immunity mechanism is specific to each producer (Hill et

al. 2008). In class II lantibiotics we can observe different immunity mechanism. In nukacin ISK-1, the proteins involved in immunity are NukFEG and NukH. NukH acts as an accessory protein to NukFEG. In mersacidin the self-protection is provided by an ABC transporter only. Conversely, in cytolisin, the immunity of the producer cell relies on the peptide CylI, that bears no homology to other LanI. The two-component lantibiotic lacticin 3147 also relies on an immunity protein, LctI, to protect its cells from the produced lantibiotic (Hill et al. 2008).

1.2.5. Class II lanthipeptides in Archaea

Although the production and isolation of lanthipeptides is still restricted to Bacteria, lanthipeptide biosynthetic genes are found in all domains. In Archaea, a study performed by Costa in 2017 revealed the presence of the class II modification enzyme LanM in several haloarchaeal genomes (chromosome and plasmid). Thus, the possible production of class II lanthipeptides in Archaea seems restricted to the class Halobacteria, phylum Euryarcheota (Fig. 14). No other lanthipeptide modification enzymes were identified in Archaea to date.
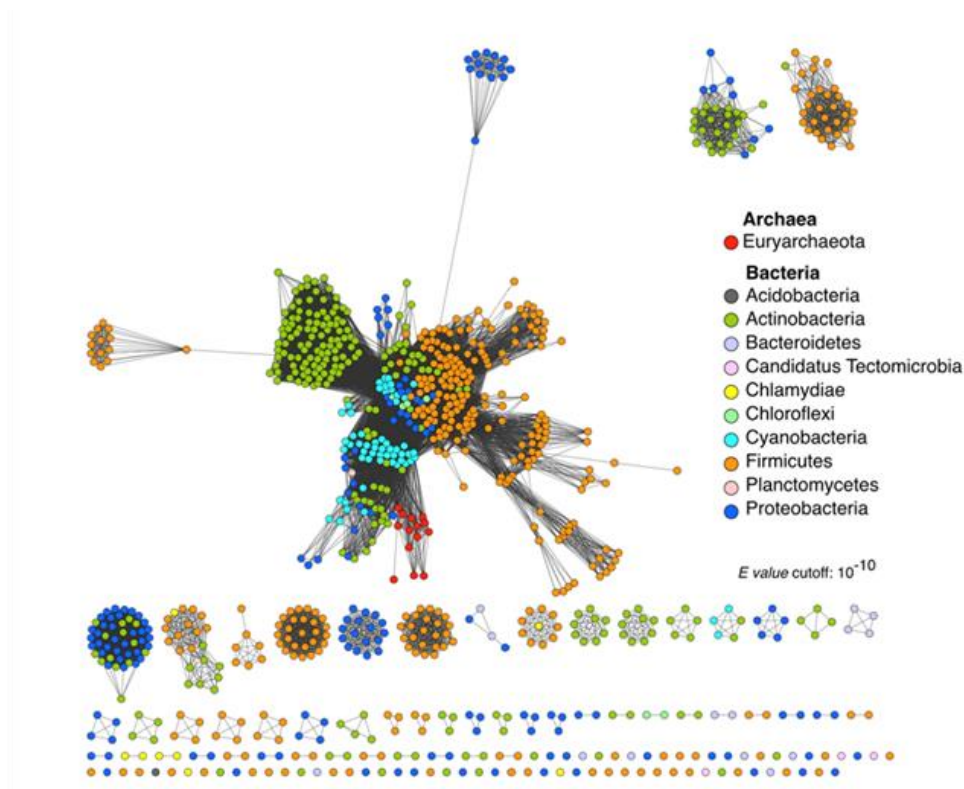
Figure 14: Sequence similarity network of LanMs generated using EFI-EST and visualized in Cytoscape with an alignment score threshold of 1^{10} (Castro *et al.* n.d.).

All the identified LanMs were subjected to phylogenetic analysis and the result is a division into 8 groups (Fig. 15). It was observed that some strains encode more than one LanM: *H. mediterranei* ATCC 33500, *H. cibarius* D43 and the halophilic archaeon J07HX5 (Fig. 15).
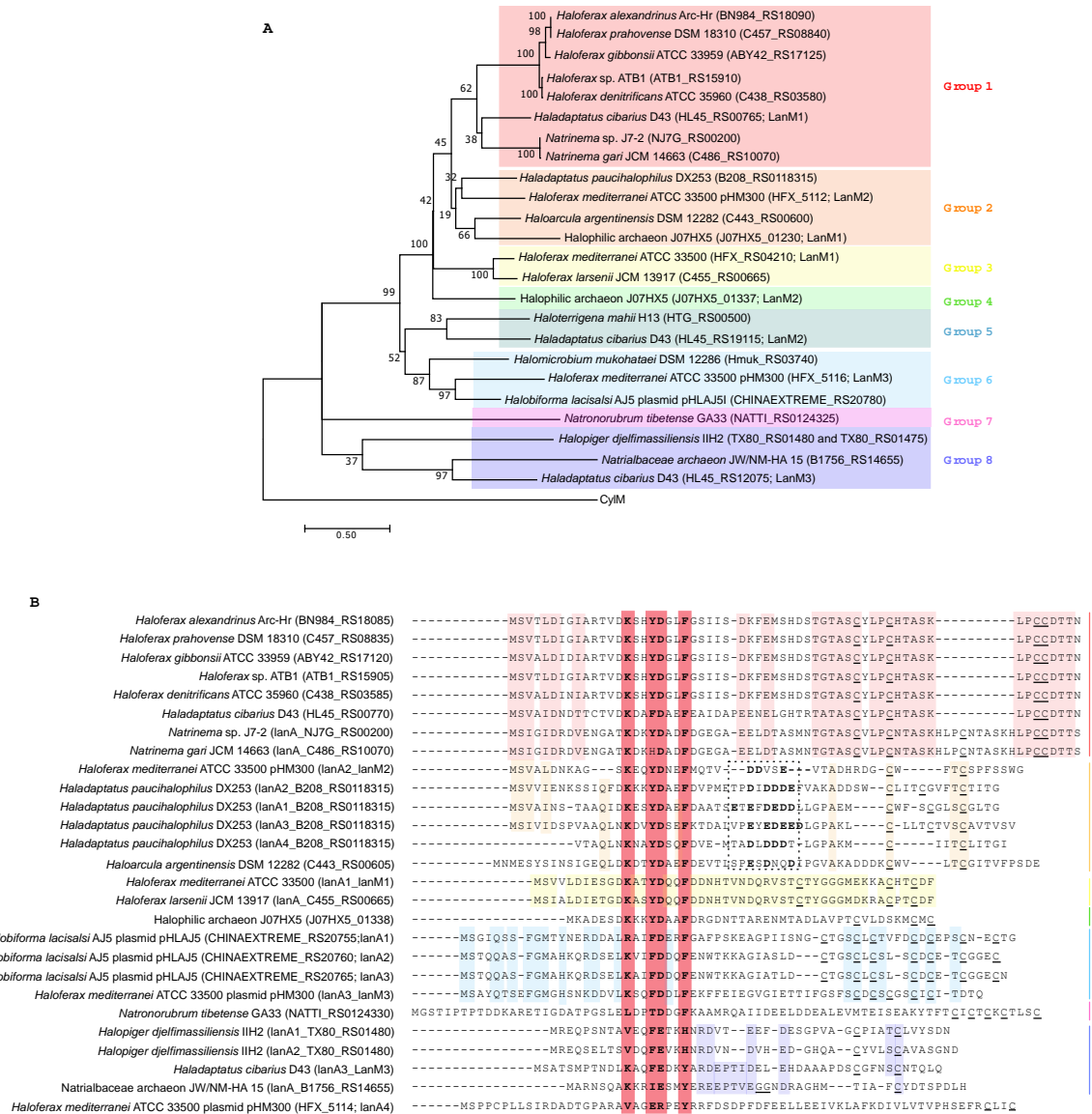
Figure 15: Phylogenetic tree of LanM proteins identified in Archaea domain. Seven groups were proposed based on their phylogenetic relationship. The branch scale determines the number of substitutions per site (I Castro et al. n.d.).

In *H. mediterranei*, three putative LanMs were found (Costa 2017). The enzymes were named MedM1, MedM2 and MedM3, being the first one located on the chromosome and the last two located on the megaplasmid pHM3000 (Fig. 16). Most recently, class II

lantibiotic precursors, LanMs, were also found in 19 haloarchaeal genomes, in which 8 belonged to the same protein family (Makarova et al. 2019).
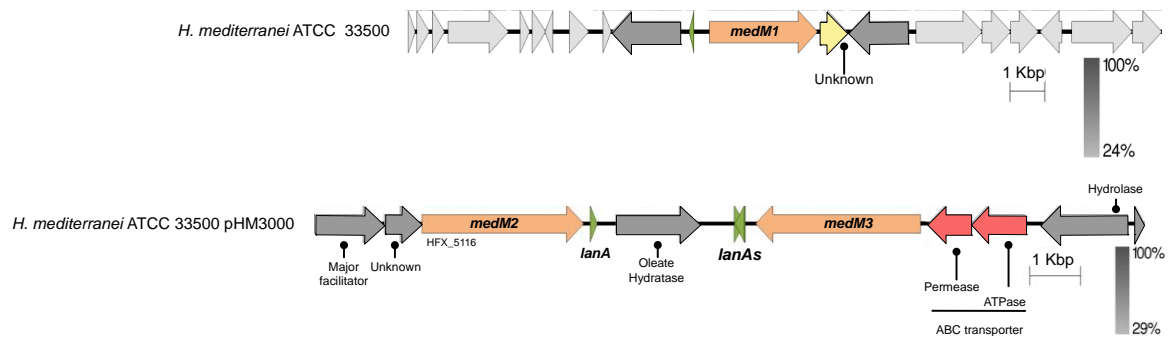
Figure 16: Biosynthetic gene clusters of the putative class II lanthipeptides present in *H. mediterranei*, MedM1, MedM2 and MedM3. *medM1* is located on the chromossome, while *medM2* and *medmM3* are located in the megaplasmid pHM3000 (Costa 2017).

Attempts of heterologous expression of the *H. mediterranei* lanthipeptides (putative LanM plus putative LanA) in both *E. coli* and *Haloferax volcanii* were unsuccessful and these lanthipeptides are yet to be isolated from haloarchaea (Costa 2017).

1.3. Context and objectives of this thesis

Halocin production by haloarchaeal strains have been reported since the 80's (Rodriguez-Valera, Juez, and Kushner 1982). One of the first haloarchaeal strains to have its antiarchaeal character described was *Haloferax mediterranei*. Halocin H4 was described as the main peptide involved in its antimicrobial activity (Inmaculada Meseguer and Rodriguez-Valera 1985). Studies performed with *H. mediterranei* mutants lacking the *halH4* gene revealed that they retained its ability to inhibit the growth of other haloarchaea (S. Chen et al. 2019; Naor, Yair, and Gophna 2013).

More recently, three putative class II biosynthetic gene clusters were found on *H. mediterranei* genome, and the respective modification enzymes were termed MedM1, MedM2 and MedM3. These findings coupled with the unknown origin of the antagonistic activity displayed by *H. mediterranei* leads to the hypothesis of class II lanthipeptides involvement in antiarchaeal activity and the idea of them being haloarcheocins as well.

Thus, the main objective of this thesis was to test whether the putative class II lanthipeptide biosynthetic clusters identified were involved in the halocin production by *H. mediterranei*. More specifically, this work intended to:

i) Review available information available on halocins produced by Archaea;

ii) Perform novel analysis of halocins clusters, including comparative genomics and amino acid sequence analysis based on advances in haloarchaeal genome sequencing and the annotation made in recent years;

iii) Implement the pop-in pop-out strategy to generate knockout mutants in the laboratory;

iv) Generate *H. mediterranei* single mutants of the genes encoding each modification enzyme, *medM1*, *medM2* and *medM3*, and evaluate its impact on halocin production phenotype;

v) Generate a triple mutant lacking all the genes encoding the MedM modification enzymes and evaluate its impact on halocin production phenotype;

The specific points i) and ii) did not involve laboratory procedures and were used to prepare a review manuscript. As such, these points were addressed in the Introduction of this thesis and were not included in the Results section.

2. MATERIALS AND METHODS

2.1. Strains and culture conditions

All haloarchaeal strains (Table III), except *Halobacterium salinarum*, were routinely grown in YPC (yeast-peptone-casamino acids) medium containing, per litre, 144 g of NaCl (2.5 M NaCl), 18 g of $MgCl_2.6H_2O$, 21 g of $MgSO_4.7H_2O$, 4.2 g of KCl, and 0.44 g of $CaCl_2$ (3 mM $CaCl_2$).

The strain *Halobacterium salinarum*, was routinely grown on a variant of YPC medium containing a higher NaCl concentration, denominated YPC Super-salt (YPC-SS). YPC-SS contains 240 g of NaCl per litre (4.1 M NaCl) and the same amount of the other reagents used in YPC medium.

For the YPC broth medium preparation, 600 mL salt-water 30% (SW 30%), 300 ml of distilled water (dH_2O), 100 ml of YPC 10X and 6 mL of 0.5 M $CaCl_2$ solution, per litre were autoclaved separately, and then mixed together in a sterile environment. For the YPC-SS broth medium preparation, 900 ml of salt-water 30% (SW 30%), 100 ml of YPC 10X and 6 mL of 0.5 M $CaCl_2$ solution, per litre, were autoclaved separately, and then mixed together in a sterile environment. The salt-water 30%, contains, per litre, 240 g of NaCl, 30 g of $MgCl_2.6H_2O$, 35g of $MgSO_4.7H_2O$, 7g of KCl and 20 mL of a 1 M Tris-HCl solution (pH 7.5). The YPC 10x contains, per litre, 50 g of yeast extract (Liofilchem), 10 g of meat peptone (Merck), 10 g of casamino acids (Difco) and 17,6 mL of a 1M KOH solution.

The CA (casamino acids) broth was prepared as YPC broth, substituting the YPC 10X solution by the CA 10X solution that contains 70 g of casamino acids per litre.

To prepare YPC or CA agar plates, agar was added to the previously warmed salt-water and distilled water to the final concentration 15g per litre.

For the counterselection of the pop out mutants, 5-FOA was added to the final concentration of 250 µg/mL. 5-FOA was initially diluted in DMSO at the final concentration of 10 mg/mL.

Table 3: Strains and plasmids used in this study.

Strains	Properties	Reference
H. mediterranei ATCC 33500	Wild-type	-
H. mediterranei WR510	Δ*pyrE*	(Naor, Yair, and Gophna 2013)
H. volcanii H53	Δ*pyrE*, Δ*trpA*	(Allers et al. 2004)
H. denitrificans H1456	Wild-type	DSM 4425
H. salinarum NRC-1	Wild-type	(Naor, Yair, and Gophna 2013)

Plasmids	Properties	Reference
pTA131	Contains the *pyrE2* as selectable gene marker	(Naor, Yair, and Gophna 2013)
pKOM1	pTA131 + flanking regions of the LanM1 gene	This study
pKOM2	pTA131 + flanking regions of the LanM2 gene	This study
pKOM3	pTA131 + flanking regions of the LanM3 gene	This study

2.2. Generation of knock-out mutants with the pop-in and pop-out strategy

2.2.1. Basis of pop-in and pop-out strategy.

In order to inactivate the biosynthesis of lanthipeptides in *H. mediterranei* ATCC 33500, we aimed to obtain a triple mutant without LanM1, LanM2 and LanM3 genes (*H. mediterranei* ΔM1ΔM2ΔM3) using the pop-in and pop-out strategy (Figure 17; Bitan-banin, Ortenberg, and Mevarech 2003).
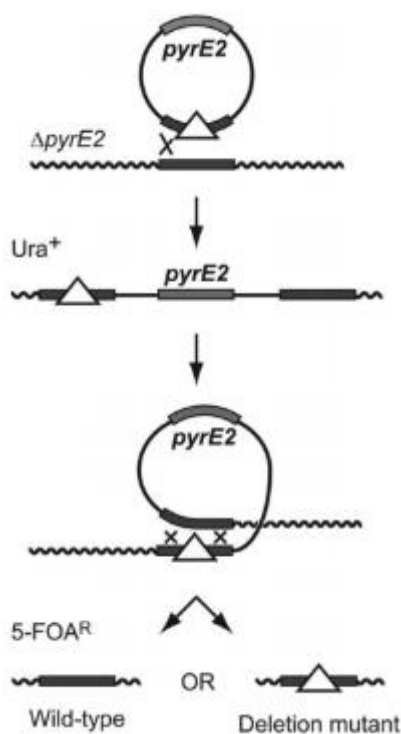
Figure 17: Schematic representation of the pop-in and pop-out technique that involves knock-out and auxotrophic strains (Allers *et al.* 2004).

This method uses "suicide plasmids", also known as knock-out plasmids that are initially incorporated into the chromosome of an uracil auxotrophic *H. mediterranei* strain (without the gene *pyrE*; Δ*pyrE*). The *pyrE2* gene encodes an orotate phosphoribosyltransferase that is involved in uracil biosynthesis. *H. mediterranei* wild-type strains are naturally susceptible to the orotate analogue compound, the 5-fluorotic acid (5-FOA) (Fig. 18). Δ*pyrE* strains cannot grow on media without an uracil source and are resistant to 5-FOA because they cannot metabolize this compound to F-dUTP, which is toxic (Fig.18).
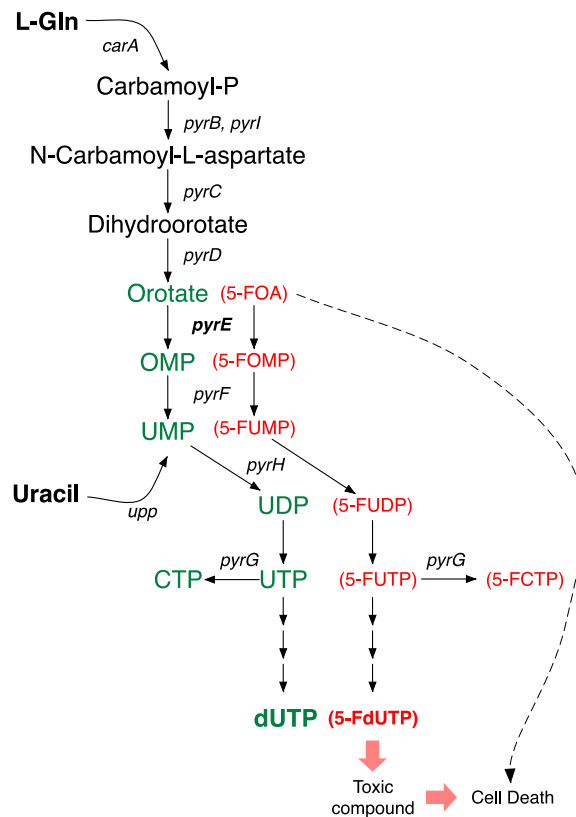
Figure 18: Schematic representation of pyrimidine metabolism. 5-FOA is an analogue of orotate. Thus, 5-FOA is metabolized by the same enzymes. However, the final product 5-FdUTP is toxic for the cells and causes cell death. Adapted from (Sakaguchi et al. 2013)

Knock-out plasmids have the *pyrE* gene that will complement the uracil auxotrophy of Δ*pyrE* strain and will allow selection of transformants in media lacking uracil. They also contain the flanking regions of the target gene (to be deleted) in order to induce recombination with it. After the transformation of Δ*pyrE* strains with these plasmids, transformants are selected based on their growth in media without uracil. This phenotype results from the incorporation of the plasmid in the chromosome (pop-in), since it does not

have an origin of replication for *Haloferax* spp. These pop-in transformants will be susceptible to 5-FOA, once *pyrE* gene is again present in the genome (Fig 17 and Fig 18). Afterwards, the pop-in transformants are then pressured by using 5-FOA in the culture media to induce an intra-chromosomal recombination with the target gene. This will induce the excision of the plasmid from the chromosome and pop-out transformants are selected. Pop-out transformants are again resistant to 5-FOA since they lost their *pyrE* gene (Fig 17 and Fig 18). From this excision two outcomes can be expected: i) restoration of the wild-type target or ii) knock-out of the target gene (Fig 17).

In this study the knock-out plasmids used to delete *medM* genes were derived from the pTA131 plasmid (Fig 18; Table 3). The plasmid pTA131 (Fig. 19), has an origin of replication for *E. coli* but not for *H. mediterranei*, and contains the gene *pyrE2* as a selectable marker for archaeal transformation. The plasmid has also some features like the ampicillin resistance gene, *AmpR*, the reporter gene *lacZ* and an origin of replication for *E. coli*.

For the purpose of this study the *H. mediterranei* WR510 strain (Table 3) lacking the *pyrE2* gene was used as a recipient of knock-out plasmids. Thus, the triple *medM* mutants were obtained from *H. mediterranei* WR510 strain (Table 3).
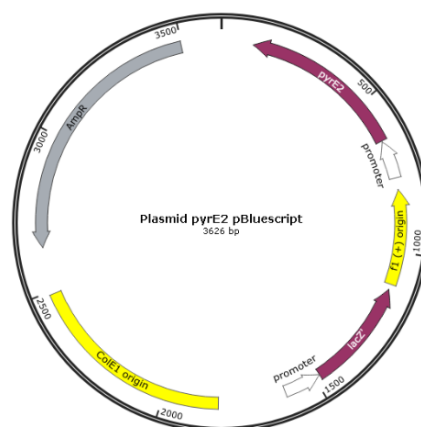
Figure 19: Plasmid pTA131. Plasmid pTA131 was used as a backbone to construct the knock-out plasmids in this study. This plasmid contains an haloarchaeal selective marker, the *pyrE2* gene.

2.2.2. First step: extraction of pTA131 plasmid

E. coli with the pTA131 was routinely cultivated in LB agar plates containing ampicillin (100 μg/mL) and grown at 37°C overnight. For plasmid extraction, a single colony was picked and grown in LB broth supplemented with ampicillin (100 μg/mL) at

37°C, overnight with aeration (180 rpm). 1,5 mL of this culture was used to the plasmid with the Nzytech miniprep kit, according to manufacturer's instructions. The concentration of the plasmid was measured with Qubit (Thermo) and integrity was evaluated by agarose gel electrophoresis (agarose percentage 0.8 %).

2.2.3. Second step: construction of the "suicide plasmids": pKOM1, pKOM2 and pKOM3

For the construction of the knock-out plasmids to allow the deletion of class II synthetase genes in *H. mediterranei*, it was necessary to clone two fragments into the MCS of pTA131, for each gene to be deleted. Thus, pKOM1, pKOM2 and pKOM3 (Table 3) were constructed with the aim of *medM1*, *medM2* and *medM3* deletion in *H. mediterranei*, respectively. Primers were constructed for each gene in order to amplify (Fig 20): i) the upstream region of the target gene and the 5' region of the target gene and ii) the 3' region of the target gene and its downstream region. Primer design was based on the procedure described to obtain the *halH4* knock-out in *H. mediterranei* (Naor, Yair, and Gophna 2013) and using the software primer3 (Untergasser et al. 2012). Restriction sites for 3 restriction enzymes were added to the primers as explained in Fig 20.
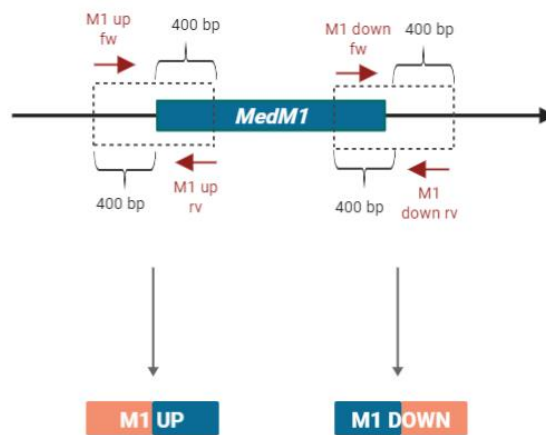
Figure 20: Schematic representation of the upstream and downstream inserts of the knock-out plasmidsusing *medM1* gene as example. Upstream fragment: M1up (≈ 800 bp), and downstream fragment: M1down (≈ 800 bp).

Each fragment was amplified in a PCR reaction containing 12.5 µL of Platinum SuperFi PCR Master Mix (2X), 1.5 µL of primer forward (10 pmol/ µL), 1.5 µL of primer reverse (10 pmol/ µL), 1 µL of template DNA and 8.5 µL of nuclease free water, in a final

volume of 25 µL. The primers used are listed in Table 4 and the amplification parameters used are listed in Table 5.

After evaluation of successful amplification by agarose gel electrophoresis (1% agarose) the amplicons were purified with the Nzytech purification kit, according to manufacturer's instructions. 1000 ng of each fragment and 1000 ng of the pTA131 plasmid were digested in independent reactions containing 3 µL of each enzyme, 1X of appropriated reaction buffer in a final volume of 40 µL for the fragments and 20 µL for the plasmid. The reactions were incubated at 37°C for 1 hour. All the digestions were separated in a 1% agarose gel and the digested DNA was excised from the gel and purified with the Nzytech purification kit, according to manufacturer's instructions.

To construct pKOM1 plasmid, a ligation reaction was prepared containing: 50 ng of the linearized pTA31, 17 ng of the digested M1 up fragment, 17 ng of the digested M1 down fragment and 1X Anza T4 DNA ligase master mix (Thermo) in a final volume of 20 µL. The reaction was incubated at room temperature for 15 minutes and 5 µL were used to transform chemically competent *E. coli* cells using the heat shock method. The transformation was plated in LB agar plates with ampicillin (100 µg/mL) and X-Gal (40 µg/mL) and grown in an incubator at 37°C overnight. White single colonies were selected and screened for the presence of the desired insert in the MCS (multiple cloning site) of pTA131 by colony PCR using the universal primers as described in Blaby et al. 2010. A positive colony was grown in LB containing 100 µg/mL of ampicillin at 37 °C, overnight and its plasmid was extracted with the Nzytech miniprep kit, according to manufacturer's instructions. The plasmid was subjected to Sanger sequencing (StabVida, Portugal) in order to confirm that the desired insert was obtained.The construction of pKOM2 and pKOM3 followed the same procedure described for pKOM1 but using the M2 up and M2 down fragments and M3 up and M3 down fragments, respectively.

Table 4: List of all the primers used in this study.

Name	Primer	Source
M1 up fw	5'-CGTCTAGACCTCTGTATCGCCGAAGATG-3'	This study
M1 up rv	5'-CAGGGATCCAACTACCCGATTCCGACAGA-3'	This study
M1 down fw	5'-CAGGGATCCGTTGTTCCAGCCCGTAATCT-3'	This study
M1 down rv	5'-GACAAGCTTCCACCCACTCACTTATGATTTG-3'	This study
M2 up fw	5'-CGTCTAGAACAGCTTGTCCGCTTCAGTC-3'	This study
M2 up rv	5'-CAGGGATCCATCGCGGTGAGTAGTTCTCC-3'	This study
M2 down fw	5'-CAGGGATCCCTATCAGGACCGCTGGTGTC-3'	This study
M2 down rv	5'-GACAAGCTTCGAACGTCGCATACAAAGTG-3'	This study
M3 up fw	5'-CGTCTAGACGGGAGGTGTTATGATGTCC-3'	This study
M3 up rv	5'-CAGGGATCCTCCACGACCAGTGGTGTTAC-3'	This study

M3 down fw	5'-CAGGGATCCGATGGCCTCATCTGAAAGGA-3'	This study
M3 down rv	5'-GACAAGCTTGAGATTCCGTGGCTCAGACT-3'	This study
medM1 fw	5'-AGTACATATGACACAGCAGCTTGCAGC-3'	(Costa 2017)
medM1 rv	5'-CGATCTCGAGTTACTCCAGCAGAAGCACACAG-3'	(Costa 2017)
medM2 fw	5'-AGTACATATGAACCGCGTCTACACGATC-3'	(Costa 2017)
medM2 rv	5'-CGATCTCGAGTTACTCAAGCAAGAGAACGGA-3'	(Costa 2017)
medM3 fw	5'-AGTACATATGGCGGGCGTATTTACTGAG-3'	(Costa 2017)
medM3 rv	5'-CGATCTCGAGTTATTCCCAGCGTAACACGTT-3'	(Costa 2017)
HalocheckM1 fw	5'-ACGATTATCGATGCGGAGAC-3'	This study
HalocheckM1 rv	5'-GGTCATCACGTCTGTGTTGG-3'	This study
HalocheckM2 fw	5'-GGTCATCACGTCTGTGTTGG-3'	This study
HalocheckM2 rv	5'-GGTCATCACGTCTGTGTTGG-3'	This study
HalocheckM3 fw	5'-TTGCATTGCGTATCTGCTTC-3'	This study
HalocheckM3 rv	5'-GCTGGTCGTCTTCTCTGGAC-3'	This study

Table 5: PCR conditions used for the formation of the upstream and downstream fragments

Primers	Steps	Temperature and Time	Cycles
M1_up_fw & M1_up_rv M1_down_fw & M1_down_rv	Initial denaturation	98°C - 30s	-
	Denaturation	98°C - 10s	30
	Annealing	63°C - 10s	
	Extension	72°C - 30s	
	Final extension	72°C - 5min	-
M2_up_fw & M2_up_rv M2_down_fw & M2_down_rv M3_up_fw & M3_up_rv M3_down_fw & M3_down_rv	Initial denaturation	98°C- 30s	-
	Denaturation	98°C - 10s	30
	Annealing	64°C - 10s	
	Extension	72°C - 30s	
	Final extension	72°C - 5min	-

2.2.4. Third step: transformation of *H. mediterranei* WR510 and selection of pop-in transformants

In order to obtain a triple knock-out, it was necessary to construct single mutants of *H. mediterranei* WR510 in the first place. *H. mediterranei* WR510 was grown in 10 mL of YPC broth in a 50 mL tube until it reached the exponential phase ($OD_{600}> 0.5$). At this stage, the strain was transformed with 1 µg of pKOM1, pKOM2 or pKOM3 plasmid using the PEG mediated protocol described by Allers *et al.* 2004. After transformation, the pop-in transformants were selected on CA media plates (media without uracil), after incubation

at 45 °C for 5 days. Colonies could only thrive on this media if the plasmid had integrated the chromosome in the region of the target gene.

2.2.5. Fourth step: selection of "true" pop-out transformants

Five pop-in transformants from the three transformations performed were grown on YPC medium supplemented with 5-FOA (250 μg/mL) at 45°C with 180 rpm. After four days the cultures were transferred to fresh medium (final $OD_{600nm:}$ of 0.05) and grown in the same conditions for four more days to force the excision of the plasmid from the chromosome. A dilution of this culture was spread on YPC agar plates containing 5-FOA (250 μg/mL) and about 50 single colonies were picked in two media: YPC agar with 5-FOA and CA agar plates. Colonies growing in the presence of 5-FOA that were not able to grow in CA agar were true pop-out transformants.

2.2.6. Fifth step: Screening of pop-out transformants to select single knock-out mutants

As above mentioned, the pop-out transformants can have the deletion of the target gene or, alternatively, they can have the wild-type target gene (Fig 17). Only the first are the desired transformants. Thus, a screening based in PCR was performed. 10 to 20 pop-out transformants were tested in each round of reactions. Colony PCR using primers "Halocheck" and "medM" was performed separately to screen the knock-out transformants. Each colony PCR reaction contained 6.25 μL of Platinum SuperFi PCR Master Mix (2X), 0.75 μL of primer forward (10 pmol/ μL), 0.75 μL of primer reverse (10 pmol/ μL), a variable amount of template DNA (biomass of the colony) and 4.25 μL of nuclease free water, in a final volume of 12.5 μL. The primers used are listed in Table 4 and the amplification parameters used are listed in Table 6.

Table 6: PCR conditions used for the screening of pop-out knock-out transformants.

Primers	Steps	Temperature and Time	Cycles
medM1* medM2 medM3	Initial denaturation	98°C - 30s	-
	Denaturation	98°C - 10s	
	Annealing	64*/62°C - 15s	30
	Extension	72°C - 2 min	
	Final extension	72°C - 5min	-
Halocheck M1 HalocheckM2 Halocheck M3	Initial denaturation	98°C- 30s	-
	Denaturation	98°C - 10s	
	Annealing	50°C - 15s	30
	Extension	72°C - 15s	
	Final extension	72°C - 5min	-

To identify *ΔpyrEΔM1* mutants, two pairs of primers were used in different reactions: i) halocheck M1 fw and halocheck M1 rv, designed to amplify a 212 bp fragment located in the middle of the *medM1* gene and ii) MedM1 fw and MedM2 rv, designed to amplify the full *medM1* gene (Table 5). The same rational was applied to identify *ΔpyrEΔM2* and *ΔpyrEΔM3* mutants and the primers used are listed also in Table 5. The result of PCR reactions was evaluated in agarose electrophoresis (1%) and single knock-out mutants were selected based on the molecular weight of their amplicons (Table 7).

Table 7: Expected size of amplicons from PCR using primers "halocheck" and "MedM" on *H. mediterranei* ATCC and the mutant strain *H. mediterranei ΔM1ΔM2ΔM3*. Abbreviations: NA, no amplification.

Primers	*H. mediterranei* ATCC 33500 (wild-type)	*H. mediterranei ΔM1ΔM2ΔM3*
Halocheck M1	252 bp	NA
Halocheck M2	254 bp	NA
Halocheck M3	212 bp	NA
MedM1	≈ 3000 bp	≈ 1000 bp
MedM2	≈ 3000 bp	≈ 1000 bp
MedM3	≈ 3000 bp	≈ 1000 bp

2.2.7. Sixth step: Construction of a triple knock-out mutant

The single knock-out *ΔpyrEΔM1* was transformed with pKOM2 plasmid according to the procedure described in section 2.2.4. The resulting pop-in and pop-out transformants, were obtained with the procedures described in sections 2.2.5. and 2.2.6. In the end, the double knock-out *ΔpyrEΔM1M2* strain was obtained. This strain was then transformed with pKOM3 plasmid as described in section 2.2.4. and the triple knock-out *ΔpyrEΔM1M2M3* was successfully identified after following the procedures of sections 2.2.5 and 2.2.6.

2.2.8. Seventh step: Quantification of *medM1*, *medM2* and *medM3* copy number

The single mutants and the triple mutant were grown in YPC broth at 45 °C for 3 days with aeration (180 rpm). 1 mL of the culture was used to extract total DNA with the Promega™ Wizard™ SV Genomic DNA Purification kit (Fisher), according to manufacturer's instructions. Total DNA of *H. mediterranei* WR510 strain was also extracted following the same procedure. This DNA was used as template to obtain standard curve of the amplification of each *medM* gene.

qPCR reactions included 1X SYBR Green qPCR master mix (Thermo), 0.5 pmol/µL of each primer and 1 µL of DNA template in a final volume of 10 µL. The volume of DNA template used in the reactions of the standard curve contained from $7x10^6$ to 0 copies of the target DNA (in decimal serial dilutions). The volume of DNA template used in the test reactions contained 30 ng of DNA of single or triple knock-out mutants. All the reactions were performed in triplicate. The primers applied to each set of reactions corresponded to the "check primers" that are listed in Table 4. These primers were constructed to amplify a region of the target gene that was deleted in the knock-out mutants. Thus, we were expecting to obtain no amplification (i.e. 0 copies) of *medM1*, *medM2* and *medM3* genes in the test reactions. The parameters of amplification are listed in Table 8. A melt curve was obtained after the last extension step in the PCR reaction. The melt temperature ranged from 65 °C to 95 °C, with an increment value of 0.5 °C and an incubation time for each temperature increment of 10 seconds.

Table 8: Real-time PCR conditions using primers halocheckM1, halocheckM2 and halocheck M3.

Step	Temperature	Time	Cycles
UDG activation	50°C	2 min	-
Dual-Lock™ DNA polymerase	95°C	2 min	-
Denaturation	95°C	15s	
Annealing	51°C	15s	40
Extension	72°C	1 min	

2.3 Growth curve

Growth curves were evaluated in order to understand if the deletion of *medM* genes had some impact on the growth of *H. mediterranei* WR510 strain.

The 250 mL flasks used for the cultures were rinsed with abundant distilled water to prevent the impairment of growth due to detergent residues. A pre-inoculum of each strain was prepared in 50 mL tubes containing 10 mL of YPC broth and at 45°C with 180 rpm for approximately 48 hours. Five biological replicas were prepared (5 different colonies in 5 different tubes. The OD_{600nm} was measured and the cultures were prepared in order to have an OD_{600nm} of 0.8. 500µL of these cultures was used to inoculate 250 mL flasks containing 50 mL of YPC broth. The cultures were incubated at 45 °C with aeration (180 rpm) for 2 days.

2.4. Bioassays

Bioassays were performed to assess the antiarchaeal activity of all *H. mediterranei* mutants (Δ*M1*, Δ*M2*, Δ*M3*, and Δ*M1M2M3*), *H. mediterranei* wild-type ATCC 33500 and *H. mediterranei* WR510 strain. As bioindicators, the haloarchaeal strains *H. volcanii*, *H. denitrificans* and *H. salinarum* were used (Table 3). Two techniques were employed to evaluate such activity: i) co-culture and ii) double agar layer.

In the co-culture technique, producers and bioindicator strains were grown on YPC broth, at 45°C with aeration (180 rpm). The exception was *H. salinarum* that was grown on YPC super-salt at 37°C and 180 rpm. After two days, the OD $_{600nm}$ of cultures were measured. The bioindicator strains were incorporated in 20 mL of YPC agar at a final OD_{600nm} of 0.05, followed by pouring in petri dishes. After the agar was solidified, 10µL of the producer's culture was inoculated in the centre of the plate. The plates were incubated at 45°C for a maximum of 7 days and the inhibition halos were visualized.

In the double agar layer technique, YPC agar plates were prepared. The producer strains were cultivated as for co-culture assay and 10 µL of the cultures were inoculated in the centre of YPC agar plates. The strains were incubated at 45°C for a period of two days and the plates were submitted to sterilization with UV light treatment for 15 minutes. Afterwards, 15 mL of YPC agar (or YPC super salt in the case of *H. salinarum*) containing each bioindicator at a final OD_{600nm} of 0.05 were poured in the UV light treated agar plates. To prepare these cultures, bioindicators were grown as for co-culture procedure. The plates were incubated at 45 °C or 37 °C (*H. salinarum*) for a maximum of 7 days and the inhibition halos were visualized.

3. RESULTS & DISCUSSION

This chapter was presented as poster in an international meeting:
Castro I., Costa H., Turgeman-Grott I., Allers T., Mendo S., Caetano T. (2019) Can the haloarchaeal lanthipeptides be halocins? The case study of *Haloferax mediterranei*. GRC Archaea: Ecology, Metabolism and Molecular Biology. Les Diablerets (Switzerland)

This chapter will be included in the manuscript that is in preparation:
Castro I., Costa H., Turgeman-Grott I., Allers T., Osório H., Mendo S., Caetano T. (in preparation) The lanthipeptides of haloarchaea: are they halocins?

3.1. Construction of "suicide" plasmids

The "suicide" plasmids pKOM1, pKOM2 and pKOM3 were constructed based on pTA131 plasmid (Fig 21). Each plasmid was created to allow the recombination with the respective target gene (*medM1*, *medM2* and *medM3*), followed by integration in the chromosome or in the megaplasmid pHM3000 of *H. mediterranei* (Table 9). The construction of these plasmids was based on Naor *et al*. (2013) procedure, aiming the deletion of the halocin H4 gene (*halH4*) from *H. mediterranei* WR510. The *halH4* gene (1080bp) is approximately three times smaller than the lanthipeptide synthetase genes (≈ 3000bp; Table 9). However, the size of the upstream and downstream regions to be included in the UP and DOWN fragments (Fig. 20) was maintained (≈ 400 bp) as well as the 5' and 3' fragments to be included in the UP and DOWN fragments, respectively (≈ 400 bp).
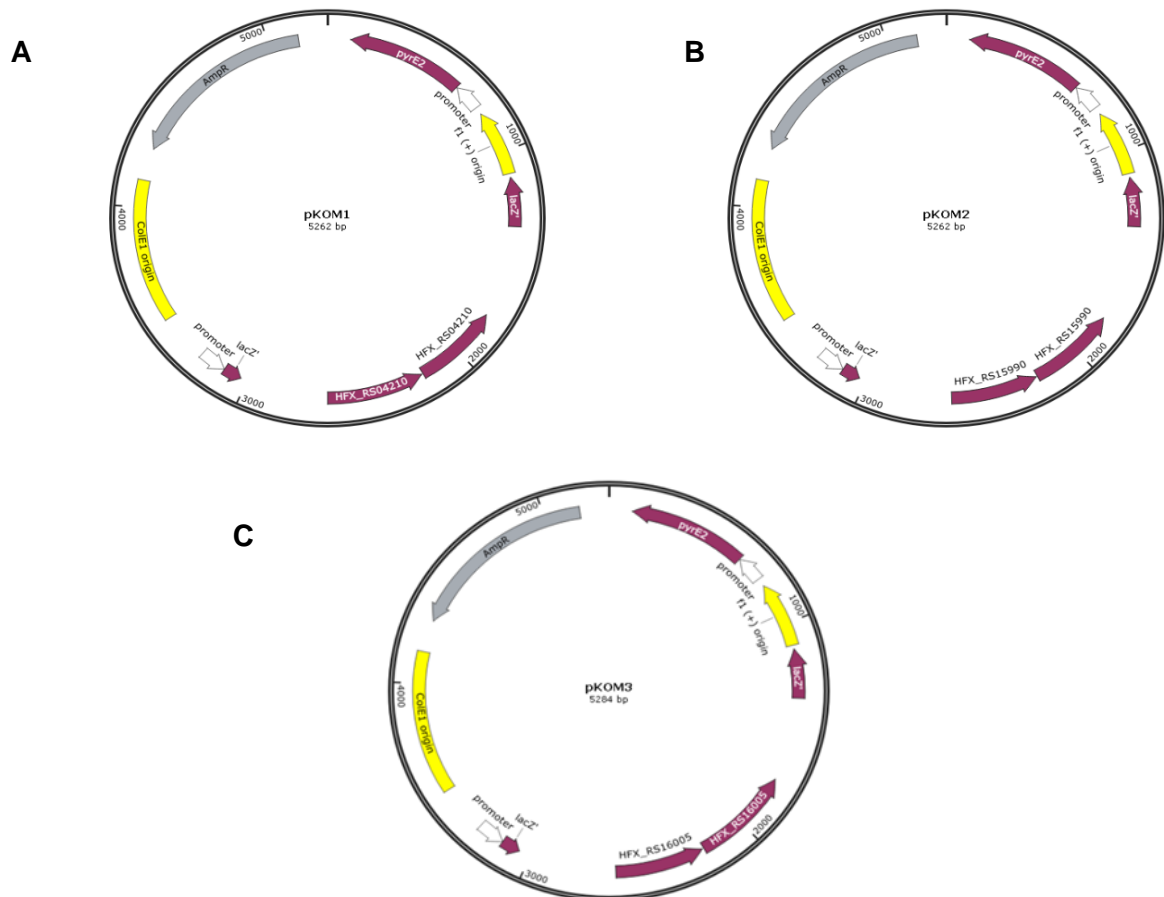
Figure 21: Knock-out plasmids: (A) pKOM1, (B) pKOM2 and (C) pKOM3. The knock-out plasmids pKOM1, pKOM2 and pKOM3 were used in the pop-in & pop-out assay to knock-out the class II lanthipeptide synthetase genes *medM1*, *medM2* and *medM3* in *H. mediterranei* WR510, respectively. The plasmids also contains the *pyrE2* gene as selective marker to allow haloarchaeal transformation and the truncated fragments of the respective *medM* gene.

Table 9: Characteristics of the genes encoding the three class II lanthipeptide synthetases found in *H. mediterranei*.

Gene	Locus tag	Size	Location
medM1	HFX_RS04210	3143 bp	Chromossome
medM2	HFX_RS15990	3191 bp	Megaplasmid pHM3000
medM3	HFX_RS16005	3221 bp	Megaplasmid pHM3000

3.2. Pop-in and Pop-out strategy

The pop-in and pop-out strategy was employed to delete the class II lanthipeptide synthetases, MedMs, from the *Haloferax mediterranei* genome (Costa 2017). One of the specific objectives of this thesis was to implement this strategy in the laboratory since it was never performed before.

In the end, five mutants were successfully created: *H. mediterranei ΔM1*, *H. mediterranei ΔM2*, *H. mediterranei ΔM3*, *H. mediterranei ΔM1M2, H. mediterranei ΔM1M2M3*. To obtain pop-in transformants, the "suicide" plasmids constructed weretransformed into *H. mediterranei* WR510. It was found that, after 3 to 5 days of incubation, the pop-in colonies were visible on the plates. However, pop-in colonies from transformation of *H. mediterranei ΔM1M2* with pKOM3 appeared only after 10 days of incubation. Thus, when no pop-in colonies are visible after a few days, the incubation period should be prolonged. The WR510 strain is auxotrophic to uracil being unable to grow on media lacking it. So, in CA agar, the cells are obligated to incorporate the plasmid by homologous recombination into their chromosome and/or megaplasmid. This will create a tandem repeat of wild-type gene, selectable markers (*bla* and *pyrE2*) and truncated fragments of the target gene (MxUP and MxDOWN) (Fig. 17). In these circumstances, strains will be able to express the *pyrE2* gene (Fig. 17 and Fig. 18). So, the pop-in strains will be able to produce dUTP from Gln and their pyrimidine metabolism will not be dependent on the uracil availability. The success of this step is correlated to the amount of plasmidic DNA used in the transformation procedure.

To obtain pop-out transformants, pop-in colonies selected in CA plates were grown in YPC broth containing 5-FOA (designated as first passage) and plated in the same media. This is a selective pressure agent that causes intrachromosomal crossing-over between the target gene and the truncate fragments, forcing the excision of the plasmid. Plasmid excision can restore the target gene or induce the deletion of the target gene (knock-out mutation) (Fig. 17). The two outcomes are correlated with the direction of crossing-over events: i) if they occur on the left of the deletion, the wild type gene

sequence is restored, ii) if they occur at the right side, the result is the desired mutation (Allers et al. 2004). This means that pop-out wild type and the pop-out knockout events are random and unpredictable. So, the colonies that grow in YPC supplemented with 5-FOA require confirmation by PCR before being selected as pop-out strains with the target gene deleted. This is the main disadvantage of the pop-in and pop-out strategy. It was found that the sub-culture of the first passage in fresh YPC broth with 5-FOA increases the number of pop-out knock-out detected, but this step is not mandatory. After plating in YPC plates containing 5-FOA, colonies were normally visible after 2-3 days of incubation. Among the pop-out transformants, the percentage of knock-outs obtained was not uniform and ranged from 2% to 35%. At least one knock-out was identified after the analysis of 60 colonies growing in YPC agar with 5-FOA. So, its advisable to perform the screening of 20 colonies in each PCR round, up to a maximum of 60 colonies/gene to be deleted.

Finally, the strategy that should be applied to successfully generate *H. mediterranei* WR510 knock-out mutants is summarized in Fig. 22. In order to obtain a single mutation, the whole process takes around 20 days.
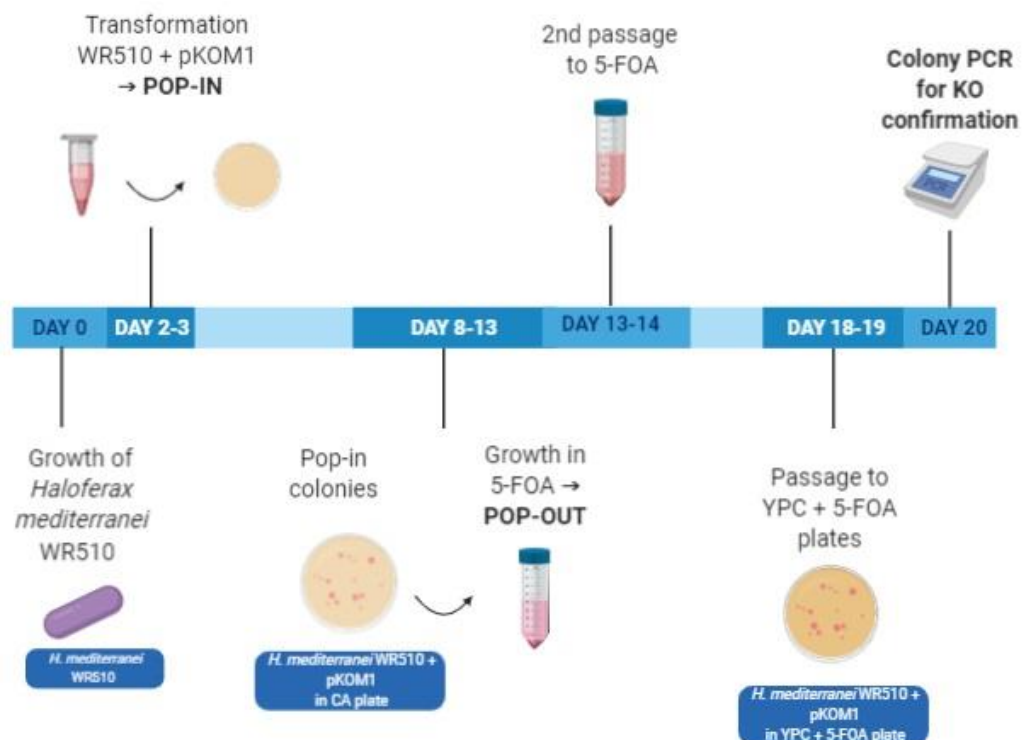

Figure 22: Haloarchaeal cell transformation time schedule using the *medM1* single knock-out as example. To obtain a single mutation, the whole procedure lasts approximately 20 days. Abbreviations: pKOM1, knock-out plasmid pKOM1; CA, casamio acids; 5-FOA, 5-fluorotic acid; YPC, yeast-peptone-casamino acis; KO, knock-out (biorender.com).

3.2.1. Quantification of *medM* genes in the single knock-outs

As abovementioned, before obtaining a *H. mediterranei* strain without the three *medM* genes, single knock-outs were created. Single knock-outs can also be very important to determine the involvement of each lanthipeptide in the halocin activity of *H. mediterranei*.

When manipulating the genome of haloarchaea, one should be aware, with time, that deletion mutations can be reverted. Polyploidy is an inherent characteristic of haloarchaea and the phenomena of gene conversion was already reported among polyploid haloarchaea (Ludt and Soppa 2019). They have multiple copies of both chromosome and plasmids. Gene conversion is defined as "the non-reciprocal transfer of information from one DNA sequence to a homologous but not an identical sequence" and it is present in all three domains of life and some viruses (Ludt and Soppa 2019). There are three types of gene conversion in haloarchaea: i) the first occurs between expression sites and non-expression sites of a given gene leading to a variation in sequences and phenotypes (antigenic variation in pathogens for example), ii) the second is termed "concerted evolution" and conversely to the previous one leads to the equalization of sequences and is often reported when a higher similarity is observed between paralogous genes from the same family in a strain compared to the orthologous genes from other strains and iii) the third type is the intermolecular gene conversion that occurs between non identical loci (Ludt and Soppa 2019). The last can be observed when, for instance, mutated copies of a gene are restored to the wildtype. Haloarchaea employ this state to repair damaged or mutated copies of their genes. The wild type copy function as the donor and transfers information to the mutated copy (the acceptor) via intramolecular gene transfer (Zerulla and Soppa 2014).

Because *H. mediterranei* strains are polyploid, each cell can have up to 30 copies of their chromosome and other 30 of their plasmids. So, if one copy of a gene remains intact it can be enough to restore the whole genome, making it an excellent mechanism of DNA repair and also a mechanism that counteracts spontaneous mutations (Zerulla and Soppa 2014). This is also the reason why a second growth of the first passage culture in medium with 5-FOA allows the identification of pop-out knock-out colonies where no copy of the *medMs* gene was left intact. So, intermolecular gene conversion can be an important biological phenomenon in the context of this study: if a single copy of the *medMs* genes remains intact in the pop-outs selected as knock-outs, gene conversion can occur and result in pop-outs with restored *medMs* genes. Consequently, it is of upmost relevance to assure that the *H. mediterranei* knock-outs constructed in this study had no

copies of the gene region that was intentionally deleted. Some authors defend that PCR is not a reliable technique to prove the inexistence of a gene and demand southern-blot analysis (personal communication). However, there are studies that applied PCR for this purpose (S. Chen et al. 2019).

qRT-PCR (quantitative real-time PCR) allows to accurately establish the gene copy number of a given sample. So, qRT-PCR was used in this study to determine the copy number of *medM* genes in the total DNA of the single and triple knock-outs. The standard curves of defined copy number/ng of DNA were constructed with total DNA extracted from three different *H. mediterranei* WR510 colonies (Fig. 23).

When performing qRT-PCR, one of the key aspects is the analysis of the results obtained. Although the main purpose is to determine the number of copies of a given sample, one must guarantee that the reaction ran properly which translates in an authentic result. To do so, we can look at the efficiency (E) and the R^2 (R^2) in Fig. 23 to verify the quality of the qRT-PCR. The assay, the master mix performance and sample quality are all variables that influence the PCR efficiency. In general, a good reaction has an efficiency of around 90-110%. The efficiency is related to amplification process, meaning that a perfect reaction, E=100%, the fold increases by 2 at each cycle. The R^2 describes how truthful the prediction of the quantity of the sample is, based on the C_T (threshold cycle) value. An R^2 value >0.99 translates in a good correlation between two values (Real-Time PCR: Understanding CT 2008).

The qRT-PCR reactions of *medM1* and *medM3* were determined successful based on the above-mentioned parameters. The E and R^2 values obtained for *medM2* standard curve determined that the reaction was not successful as for *medM1* and *medM3* (Fig. 23).

None of the *medM* under evaluation was amplified in the test conditions (DNA of *H. mediterranei ΔM1*, *H. mediterranei ΔM2*, *H. mediterranei ΔM3* or *H. mediterranei ΔM1M2M3*) up to 30 cycles of amplification. These results confirmed that the copy number of *medM* genes in the knock-outs is 0 copies/ng DNA.
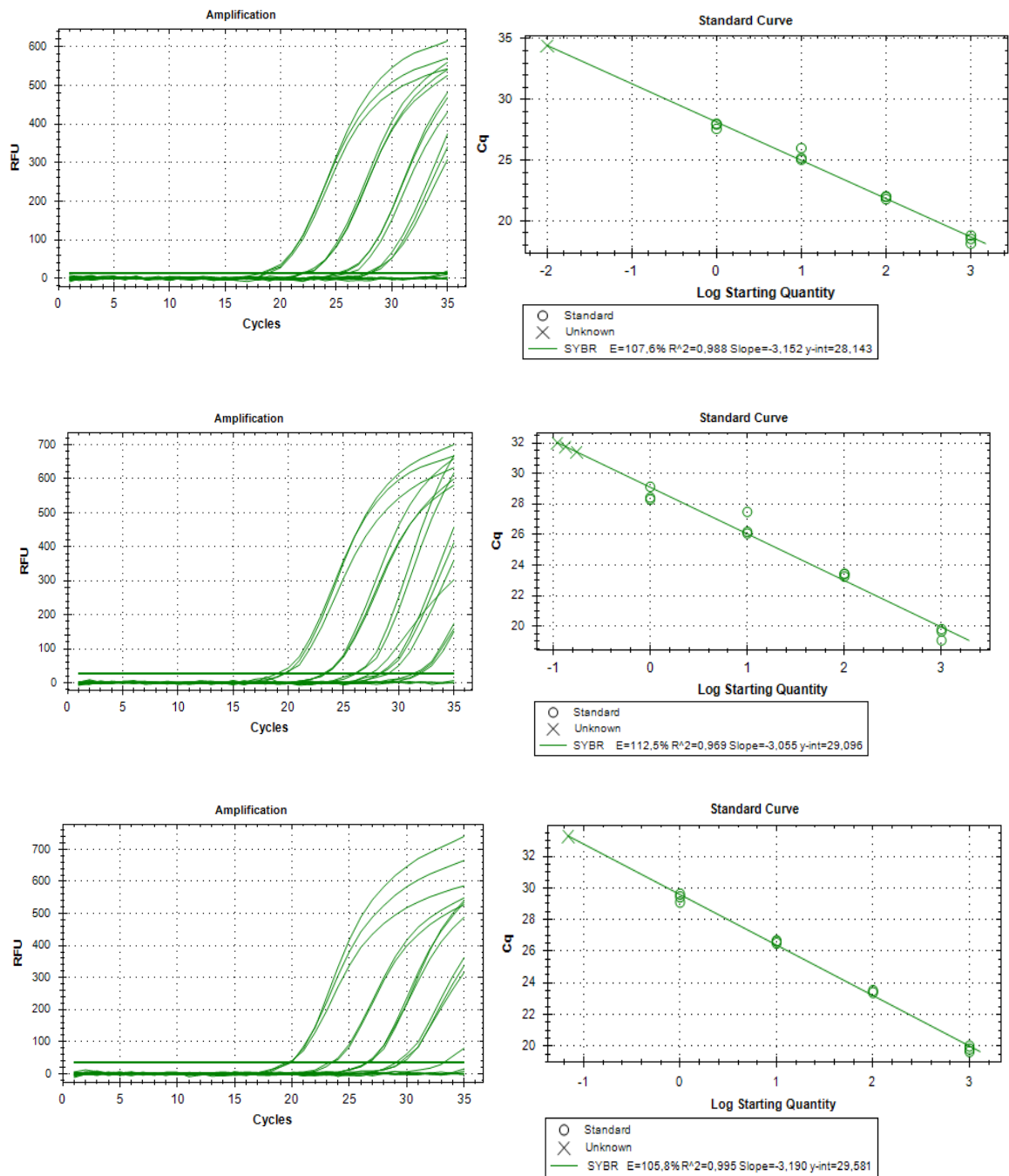
Figure 23: qRT-PCR of *medM1*, *medM2* and *medM3* genes from *H. mediterranei* ΔM1, *H. mediterranei* ΔM2 and *H. mediterranei* ΔM3 respectively.

3.2.2. The triple mutant *H. mediterranei ΔM1M2M3*

The triple mutant *H. mediterranei ΔM1M2M3* was successfully obtained from a series of mutants and mutations. Firstly, strain *H. mediterranei* WR510 was transformed with pKMO1 and subjected to the pop-in and pop-out assay, giving rise to the pop-outs. The single knock-outs *H. mediterranei ΔM1* were firstly identified by PCR with the MedM1 (forward and reverse) primers. They target the boundaries of *medM1* gene, amplifying both wild-type pop-outs as well as knockout pop-outs, with differences in the molecular weight of amplicons (example in Fig. 24). The pop-out knockouts were subsequently confirmed by other PCR with HalocheckM1 (forward and reverse) primers that only amplify when the wildtype gene is intact because this set of primers amplifies the middle of the gene (example in Fig. 25).
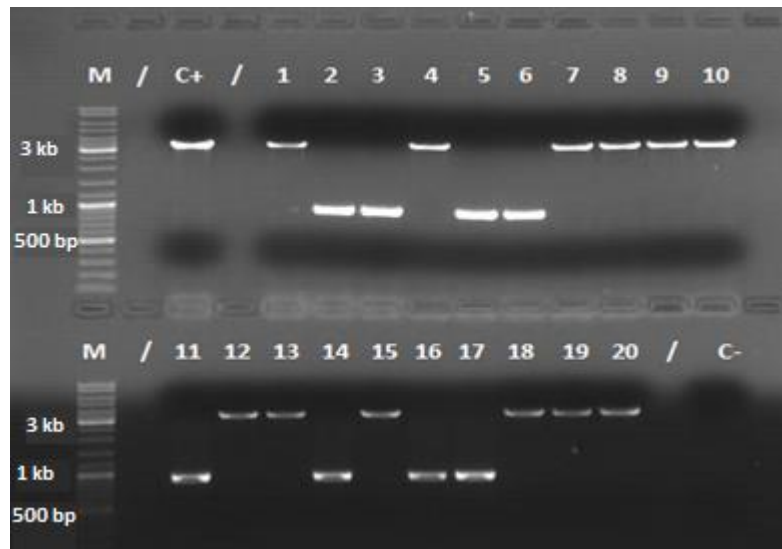
Figure 24: Example of an agarose gel showing the result of pop-out colonies screening. In this case, the primers MedM3 were used to screen pop-out colonies resulting from *H.mediterranei ΔM1M2* + pKOM3 pop-in. Legend: M, Gene Ruler DNA Ladder Mix marker; C+, positive control; C -, negative control; 1 to 20: pop-out colonies.

The strain *H. mediterranei ΔM1* was subsequently transformed with the pKOM2 plasmid to obtain the mutant *H. mediterranei ΔM1M2* applying the same technique as for *H. mediterranei ΔM1*. The same PCR strategy was applied to select a pop-out knock-out *H. mediterranei ΔM1M2*, but with primers targeting the flanking regions or the middle of *medM2* gene (example in Figs. 24 and 25). Lastly, the strain *H. mediterranei ΔM1M2* was transformed with pKMO3 and the triple mutant *H. mediterranei ΔM1M2M3* was selected after pop-in and pop-out methodology and the screening of *medM3* gene as mentioned for

medM1 and *medM2* (Figs. 24 and 25). The triple mutant was successfully constructed after approximately 6 months.
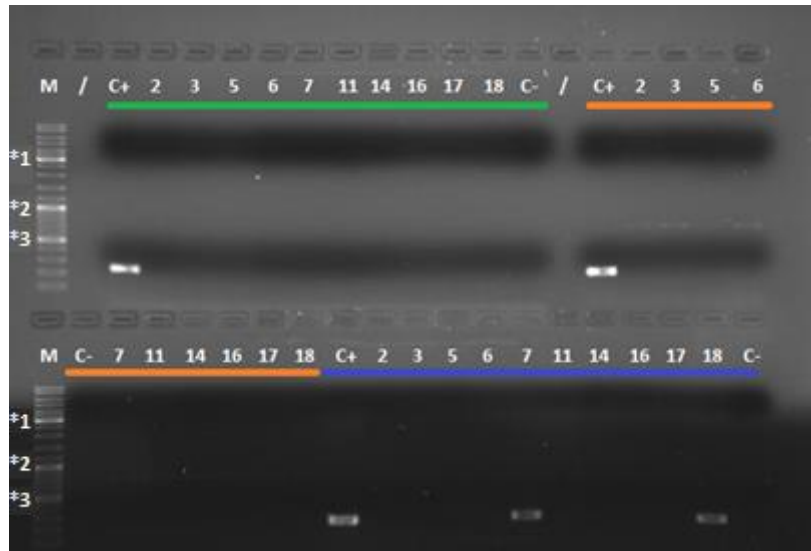
Figure 25: Example of an agarose gel showing the result of confirmation of pop-out knock-outs by PCR using primers HalocheckM1 (green), HalocheckM2 (orange) and HalocheckM3 (blue). The results demonstrated that all putative pop-out knockout colonies were indeed *H. mediterranei ΔM1M2M3*. Legend: M, Gene Ruler DNA Ladder Mix marker; C+, positive control; C -, negative control; 2, 3, 5, 6, 11, 14, 16, and 17, putative triple knockouts; 7 and 18, pop-out wild-type (internal controls).

3.2.3. Quantification of *medM* genes in the triple knock-out

After obtaining the triple knock-out mutant of *H. mediterranei* lacking all class II lanthipeptide synthetase genes, qRT-PCR analysis was performed to validate all mutations (Fig. 26) The qRT-PCR reactions of *medM1* and *medM2* were determined successful based on the E and R^2 values observed. The E and R^2 values of *medM3* amplification were above of the expected (Fig. 26). None of the *medM* under evaluation were amplified using DNA of *H. mediterranei ΔM1M2M3* up to 30 cycles of amplification. These results confirmed that the copy number of *medM* genes in the triple knock-out is 0 copies/ng DNA (Fig. 26).
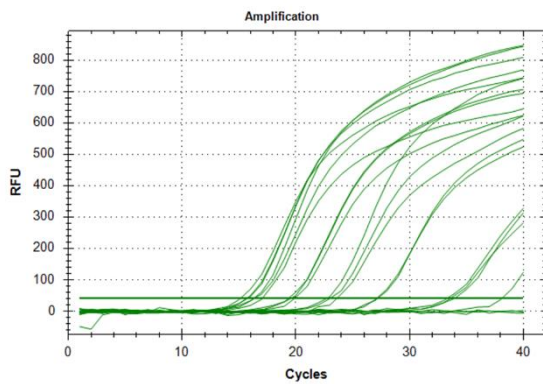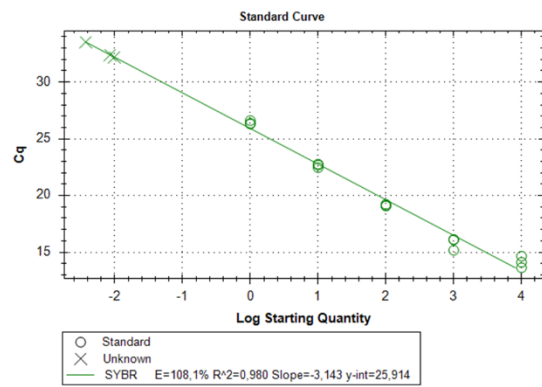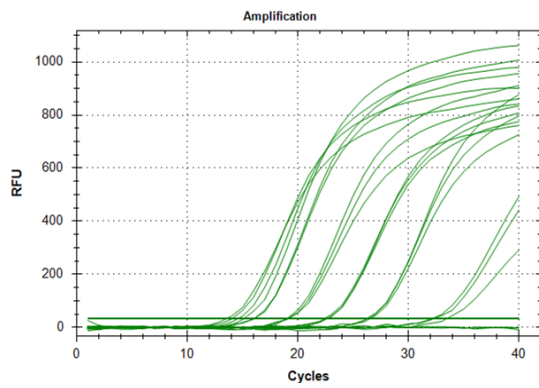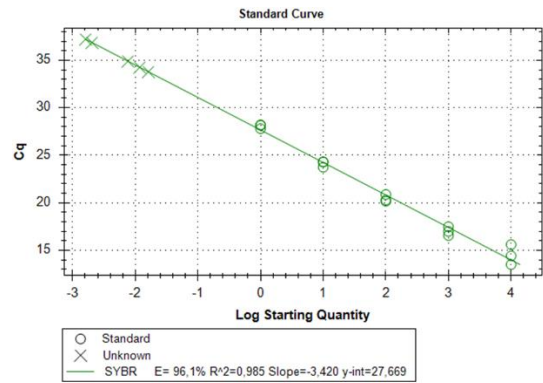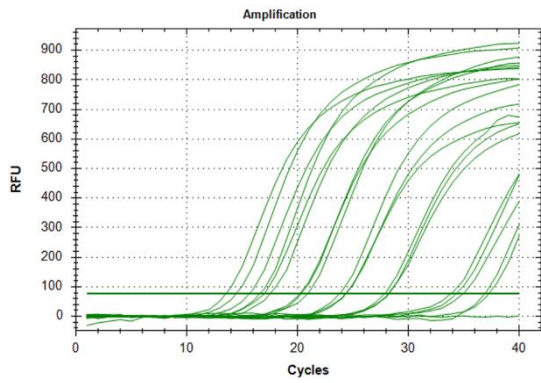
Figure 26: qRT-PCR of *medM1*, *medM2* and *medM3* genes from *H. mediterranei* ΔM1M2M3, respectively.
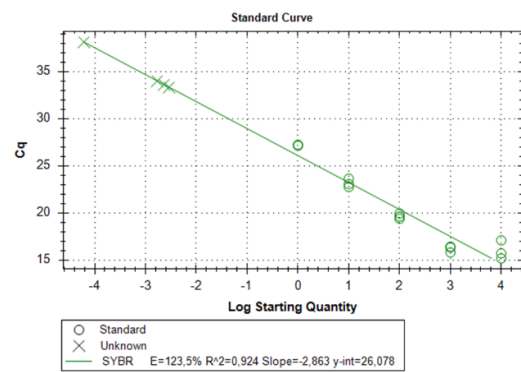
3.3. Impact of triple gene deletion in the growth of *H. mediterranei*

The growth of *H. mediterranei* WR510 and *H. mediterranei* Δ*M1M2M3* in broth was monitored for two days (Fig.27), but no major differences were observed. Thus, deletion of the class II lanthipeptide synthetase genes, *medM1*, *medM2* and *medM3* did not interfere with *H. mediterranei* growth.
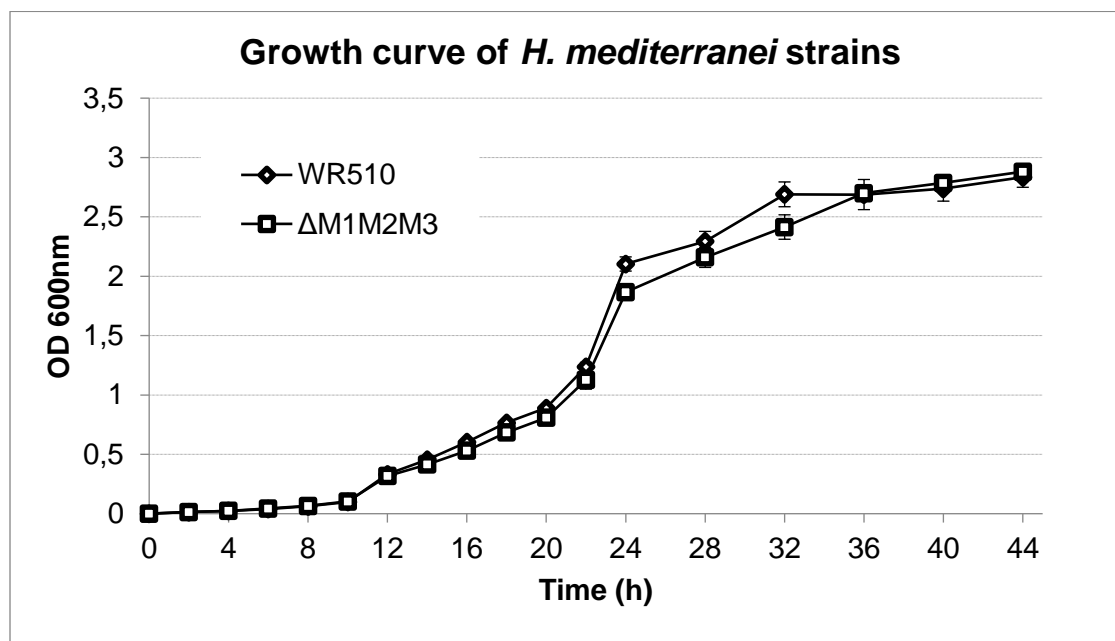
Figure 27: Growth curve of two strains of *H. mediterranei* WR510 and the triple knock-out Δ*M1M2M3*.

3.4 Bioassays

The biosynthetic clusters of RiPPs, which includes the precursor peptide, modification enzymes and transport and cleavage proteins, are not intrinsically exclusive. In class II lanthipeptides, the latter statement means that the absence of a precursor peptide, LanA in a given cluster does not obligatorily translate in the absence of lanthipeptide production: if the synthetase LanM is intact, it can use as substrate other LanAs located in different biosynthetic clusters, as it was observed for the class II synthetase ProcM (Knerr and van der Donk 2012; Yu, Mukherjee, and van der Donk 2015). This possible promiscuity in substrate usage by the LanMs requires the deletion of all *lanM* genes for the functional study of lanthipeptides *in vivo*. So, this was the strategy applied in this work.

After obtaining the triple knock-out, *H. mediterranei* ΔM1M2M3, bioassays were performed to evaluate the antiarchaeal character of the strain compared to its parental strain WR510. Firstly, *H. salinarum* was used as bioindicator strain. Both co-culture and double agar layer technique were employed, but the inhibition halos were clearer in the second technique. It was observed that both strains have antiarchaeal activity against *H. salinarum* (Fig. 28). The same result was observed when other two haloarchaea were used as bioindicators: *H. volcanii* and *H. denitrificans* (Table 3).

Thus, the lanthipeptides putatively produced by *H. mediterranei* are not involved in its microbial activity towards *H. salinarum*, *H. volcanii* and *H. denitrificans*.

Figure 28: Growth inhibition of *H. salinarum* by *H. mediterranei* WR510 (*ΔpyE*) and *H. mediterranei* *ΔM1M2M3*. Double agar layer technique to test the antiarchaeal profile of strains WR510 and Δ*M1M2M3* against the strain *H. salinarum*.

4. CONCLUSIONS AND FUTURE CHALLENGES

4.1. Conclusion

Archaea was established as a new Domain 42 years ago (Woese and Fox 1977). At that time, archaeal strains were almost a synonymous of extremophile micro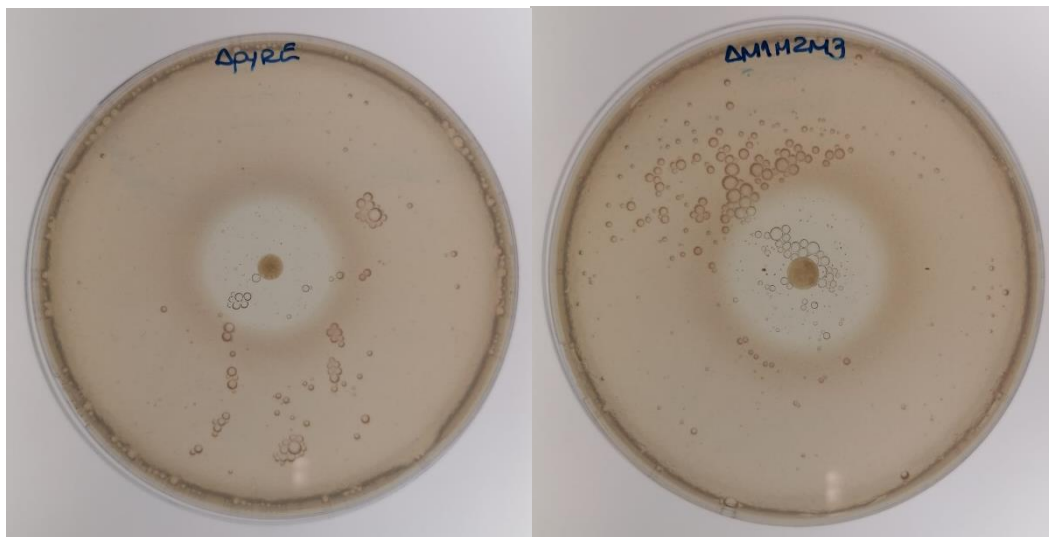organisms. Nowadays, it is recognized that the ecological niches of Archaea are wide, and may even be found in the human microbiome (Nkamga, Henrissat, and Drancourt 2017). However, the knowledge on archaeal secondary metabolites is still lagging behind. In particular, the potential of haloarcheocins is an unexplored research area. Although several haloarcheocins have been described, most of them were characterized solely based on the bioactivity of their supernatants or the bioactivity detected in cross-inhibition tests. Few studies were conducted with purified compounds and, to the best of our knowledge, none of the haloarcheocins described so far were structurally characterized. The best well-studied haloarcheocin is HalC8, which has a proposed biosynthetic cluster and some data on its biosynthetic model (Besse et al. 2017; Mei et al. 2008). More recently it was shown that HalC8 and HalC8-like peptides are not confined to haloarchaeal genomes, they can be detected in other Archaea and in some Bacteria (Makarova et al. 2019). Due to this haloarcheocin being the one with more information available, it was possible to investigate the putative biosynthetic clusters involved in the production of HalC8 and HalC8-related peptides by Haloarchaea throughout comparative genomics analysis. Our analysis suggests the division of HalC8 and HalC8-like peptides in, at least, three groups. As already described by Besse et al. (2017), HalC8 is only found in *Natrinema* spp. and *Haloterrigena* spp. All the others are HalC8-like peptides, being the ones found on *Haloferax* spp. the most closely-related with HalC8. Comparative genomics showed that the biosynthetic cluster of HalC8 should include a protein of unknown function (*halU*), two membrane-located peptides (*halP1* and *halP2*) and a transcriptional regulator (*halR*). The biosynthetic cluster of some HalC8 have also some ABC transport protein, that are also present in strain not containing the HalC8 gene. Thus, the ABC transporters identified previously, albeit intriguing, are most probably not required for HalC8 biosynthesis.

The role of haloarcheocins such as HalC8 is still to be unravelled. The search for the peptide and/or proteins responsible for antiarchaeal activity displayed by some haloarchaeal strains is intriguing. Following the example of the peptide HalH4 of *H. mediterranei*, more putative haloarcheocins should be submitted to gene knock-out to further prove their mode of action and whether or not they are "true" haloarcheocins. Lanthipeptide synthetases from class II, LanMs, were also found encoded in the genome of *H. mediterranei*. However, their role in the antimicrobial activity displayed by this strain

is still to be unravel, as their elimination from *H. mediterranei* genome did not translate in the loss of antimicrobial activity. So, in this study we demonstrated that the putative class II lanthipeptides are not the major contributors to the cross-inhibition observed. We cannot, however, rule out the possibility of them being haloarcheocins. Further studies need to be conducted in order to evaluate whether the putative lanthipeptides are being expressed in the parental strain WR510 as well as the wild-type, *H. mediterranei* ATCC. The role of the putative class II lanthipeptides in *H. mediterranei* remains unknown.

The discovery of cross-inhibition between haloarchaeal strains was attributed to haloarcheocins, antimicrobial peptides with antiarchaeal activity. The protein origin of such activity was proved when protease treatment of supernatants led to a loss of antiarchaeal properties (Meseguer 1986). However, the contribution of other substance cannot be ruled out. New genome mining technologies result in more genomes being sequenced and also in the discovery of other biosynthetic clusters, encoding other secondary metabolites such as terpenes (Wang et al. 2019), and siderophores (Dave, Anshuman, and Hajela 2006; Patil, Suryawanshi, and Bajekal 2016). These molecules are gaining popularity and are candidates to antimicrobials produced by Archaea. The genome of *H. mediterranei* includes two clusters involved in the biosynthesis of terpenes and one cluster for a siderophore. These finding are rather recent and further studies need to be conducted in order to understand their role in bioactivity and overall function on *H. mediterranei*.

4.2. Future challenges

In an era where Archaea is becoming more and more relevant, it is imperative that research in archaeocins, lanthipeptides and other biosynthetic clusters, benefit from today's technology to go a step further. Future challenges will include the generation of knock-out mutants for the genes that encode putative antimicrobials in order to establish a gene-antiarchaeal activity direct relationship. More specifically, in *H. mediterranei*, the generation of terpene and siderophore mutants will enlighten their role in the bioactivity displayed. These systems will also be essential for the investigation of biosynthetic models and ecological roles. The strategies of purification and peptide identification are challenging, but this area is crucial to provide more insights into the bioactivity, the mode of action and the biotechnological potential haloarchaeal secondary metabolites.

BIBLIOGRAPHY

Allers, Thorsten, Hien-ping Ngo, Moshe Mevarech, and Robert G Lloyd. 2004. "Development of Additional Selectable Markers for the Halophilic Archaeon." *Applied and Environmental Microbiology* 70(2).

Atanasova, Nina S., Maija K. Pietilä, and Hanna M. Oksanen. 2013. "Diverse antimicrobial interactions of halophilic Archaea and Bacteria extend over geographical distances and cross the domain barrier." *MicrobiologyOpen* 2(5): 811–25.

Barbosa, Joana, Tânia Caetano, and Sónia Mendo. 2015. "Class I and Class II Lanthipeptides Produced by *Bacillus* Spp." *Journal of Natural Products* 78(11): 2850–66.

Besse, Alison et al. 2017. "Halocin C8: an antimicrobial peptide distributed among four halophilic archaeal genera: *Natrinema*, *Haloterrigena*, *Haloferax*, and *Halobacterium*." *Extremophiles* 21(3): 623–38.

Besse, Alison, Jean Peduzzi, Sylvie Rebuffat, and Alyssa Carré-Mlouka. 2015. "Antimicrobial peptides and proteins in the face of extremes: lessons from archaeocins." *Biochimie* 118: 344–55.

Bitan-banin, Gili, Ron Ortenberg, and Moshe Mevarech. 2003. "Development of a gene knockout system for the halophilic archaeon." *Journal of Bacteriology* 185(3): 772–78.

Blaby, Ian K. et al. 2010. "Towards a systems approach in the genetic analysis of archaea: accelerating mutant construction and phenotypic analysis in *Haloferax volcanii*." *Archaea* 2010.

Caetano, Tânia et al. 2011. "Heterologous expression, biosynthesis, and mutagenesis of type ii lantibiotics from *Bacillus licheniformis* in *Escherichia coli*." *Chemistry and Biology* 18(1): 90–100.

Castro, I et al. *The Lanthipeptides of Haloarchaea: Are They Halocins? (In Preparation)*.

Castro, Inês, Sónia Mendo, and Tânia Caetano. "Antibiotics from haloarchaea: what can we learn from comparative genomics? (Under Revision)." *Marine Biotechnology*.

Chatterjee, Champak et al. 2005. "Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production." *Journal of the American Chemical Society* 127(44): 15332–33.

Chen, Ping et al. 2001. "Effect of amino acid substitutions in conserved residues in the leader peptide on biosynthesis of the lantibiotic mutacin II." *FEMS Microbiology Letters* 195(2): 139–44.

Chen, Shaoxing et al. 2019. "A halocin promotes DNA uptake in *Haloferax mediterranei*." *Frontiers in Microbiology* 10.

Cheung, Joyce et al. 1997. "Isolation, sequence, and expression of the gene encoding halocin H4, a bacteriocin from the halophilic archaeon *Haloferax mediterranei* R4." *Journal of Bacteriology* 179(2): 548–51.

Costa, Hugo. 2017. "Lanthipeptides of Archaea: the case study of *Haloferax mediterranei*." MSc thesis of University of Aveiro.

Dave, B. P., Kena Anshuman, and Puja Hajela. 2006. "Siderophores of halophilic archaea and their chemical characterization." *Indian Journal of Experimental Biology* 44(4): 340–44.

Dilks, Kieran, M. I. Giménez, and Mechthild Pohlschröder. 2005. "Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic Archaea." *Journal of Bacteriology* 187(23): 8104–13.

Dong, Shi Hui et al. 2015. "The enterococcal cytolysin synthetase has an unanticipated lipid kinase fold." *eLife* 4(JULY2015): 1–18.

van der Donk, Wilfred A., and Satish K. Nair. 2014. "Structure and mechanism of lanthipeptide biosynthetic enzymes." *Current Opinion in Structural Biology* 29: 58–66.

Gaglione, Rosa et al. 2017. "Insights into the anticancer properties of the first antimicrobial peptide from Archaea." *Biochimica et Biophysica Acta - General Subjects* 1861(9): 2155–64.

Haseltine, C. et al. 2001. "Secreted euryarchaeal microhalocins kill hyperthermophilic Crenarchaea." *Journal of Bacteriology* 183(1): 287–91.

Havarstein, Leiv Sigve, Dzung Bao Diep, and Ingolf F. Nes. 1995. "A family of bacteriocin abc transporters carry out proteolytic processing of their substrates concomitant with export." *Molecular Microbiology* 16(2): 229–40.

Hill, Colin, Lorraine Draper, R. Ross, and Paul Cotter. 2008. "Lantibiotic immunity." *Current Protein & Peptide Science* 9(1): 39–49.

Karr, Elizabeth A. 2014. 89 Advances in Applied Microbiology. *Transcription Regulation in the Third Domain*. 1st ed. Elsevier Inc.

Karthikeyan, P., Sarita G. Bhat, and M. Chandrasekaran. 2013. "Halocin SH10 production by an extreme haloarchaeon *Natrinema* sp. BTSH10 isolated from salt pans of south India." *Saudi Journal of Biological Sciences* 20(2): 205–12.

Kavitha, P., A P. Lipton, A R. Sarika, and M S. Aishwarya. 2011. "Growth characteristics and halocin production by a new isolate, *Haloferax volcanii* KPS1 from kovalam solar saltern (India)." *Research Journal of Biological Sciences* 6(January): 257–62.

Kis-Papo, Tamar, and Aharon Oren. 2000. "Halocins: are they involved in the competition between halobacteria in saltern ponds?" *Extremophiles* 4(1): 35–41.

Knerr, Patrick J., and Wilfred A. van der Donk. 2012. "Discovery, biosynthesis, and engineering of lantipeptides." *Annual Review of Biochemistry* 81(1): 479–505.

Kumar, Vijay, and Santosh Kumar Tiwari. 2017a. "Activity-guided separation and characterization of new halocin HA3 from fermented broth of *Haloferax larsenii* HA3." *Extremophiles* 21(3): 609–21.

Kumar, Vijay, and Santosh Kumar Tiwari. 2017b. "Halocin HA1: An archaeocin produced by the haloarchaeon *Haloferax larsenii* HA1." *Process Biochemistry* 61(2017): 202–8.

Lagier, Jean Christophe et al. 2016. "Culture of previously uncultured members of the human gut microbiota by culturomics." *Nature Microbiology*.

Lequerica, J. L. et al. 2006. "A halocin acting on Na+/H+ exchanger of haloarchaea as a new type of inhibitor in nhe of mammals." *Journal of Physiology and Biochemistry* 62(4): 253–62.

Li, Yun et al. 2003. "Purification and biological characterization of halocin C8, a novel peptide antibiotic from *Halobacterium* strain AS7092." *Extremophiles* 7(5): 401–7.

Ludt, Katharina, and Jörg Soppa. 2019. "Polyploidy in halophilic archaea: regulation, evolutionary advantages, and gene conversion." *Biochemical Society Transactions* 47(3): 933–44.

Makarova, Kira S et al. 2019. "Antimicrobial peptides, polymorphic toxins, and self-nonself recognition systems in archaea : an untapped armory for intermicrobial conflicts." *American Society for Microbiology* 10(3).

Mei, Shuangshuang et al. 2008. "The helix-loop-helix motif at the n terminus of hall is essential for its immunity function against halocin C8." *Journal of Bacteriology* 190(19): 6501–8.

Meseguer, I., and F. Rodriguez-Valera. 1986. "Effect of halocin H4 on cells of *Halobacterium halobium*." *Journal of General Microbiology* 132(11): 3061–68.

Meseguer, Inmaculada, and F. Rodriguez-Valera. 1985. "Production and purification of halocin H4." *FEMS Microbiology Letters* 28(2): 177–82.

Meseguer, Inmaculada, Marina Torreblanca, and Tetsuya Konishi. 1995. "Specific inhibition of the halobacterial Na+/H+ antiporter by halocin H6." *Journal of Biological Chemistry* 270(12): 6450–55.

Naor, Adit, Yael Yair, and Uri Gophna. 2013. "A halocin-H4 mutant *Haloferax mediterranei* strain retains the ability to inhibit growth of other halophilic Archaea." *Extremophiles* 17(6): 973–79.

Nkamga, Vanessa Demonfort, Bernard Henrissat, and Michel Drancourt. 2017. "Archaea: essential inhabitants of the human digestive microbiota." *Human Microbiome Journal*

3: 1–8.

Notomista, Eugenio et al. 2015. "The identification of a novel *Sulfolobus islandicus* CAMP-like peptide points to archaeal microorganisms as cell factories for the production of antimicrobial molecules." *Microbial Cell Factories* 14(1): 1–11.

O'Connor, EM, and RF Shand. 2002. "Halocins and sulfolobicins: the emerging story of archaeal protein and peptide antibiotics." *Journal of Industrial Microbiology and Biotechnology* 28(1): 23–31.

Oren, Aharon. 2014. "Taxonomy of halophilic archaea: current status and future challenges." *Extremophiles* 18(5): 825–34.

Pašić, L., B. H. Velikonja, and N. P. Ulrih. 2008. "Optimization of the culture conditions for the production of a bacteriocin from halophilic archaeon Sech7a." *Preparative Biochemistry and Biotechnology* 38(3): 229–45.

Patil, Jaysing, Priyanka Suryawanshi, and Shyam Bajekal. 2016. "Siderophores of haloalkaliphilic archaea from Lonar Lake, Maharashtra, India." *Current Science* 111(4): 621–23.

Paul, Moushumi, Gregory C Patton, and Wilfred A van der Donk. 2007. "mutants of the zinc ligands of lacticin 481 synthetase retain dehydration activity but have impaired cyclization activity." *Biochemistry* 46(21): 6268–76.

Platas, Gonzalo, Inmaculada Meseguer, and Ricardo Amils. 2002. "Purification and biological characterization of halocin H1 from *Haloferax mediterranei* M2a." *International Microbiology* 5(1): 15–19.

Price, L. B., and R. F. Shand. 2000. "Halocin S8: a 36-amino-acid microhalocin from the haloarchaeal strain S8a." *Journal of Bacteriology* 182(17): 4951–58.

"Real-Time PCR: understanding CT." 2008. *Retrieved from: https://www.gene-quantification.de/ab-application-note-understanding-ct.pdf.*

Repka, Lindsay M., Jonathan R. Chekan, Satish K. Nair, and Wilfred A. Van Der Donk. 2017. "mechanistic understanding of lanthipeptide biosynthetic enzymes." *Chemical Reviews* 117(8): 5457–5520.

Rodriguez-Valera, F, G Juez, and D J Kushner. 1982. "Halocins: salt-dependent bacteriocins produced by extremely halophilic rods." *Canadian Journal of Microbiology*: 3–6.

Rose, R. Wesley, Thomas Brüser, Jessica C. Kissinger, and Mechthild Pohlschröder. 2002. "Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway." *Molecular Microbiology* 45(4): 943–50.

Sakaguchi, Kouta et al. 2013. "The *pyrE* gene as a bidirectional selection marker in *Bifidobacterium longum* 105-A." *Bioscience of Microbiota, Food and Health* 32(2): 59–68.

Seck, E. H. et al. 2019. "salt in stools is associated with obesity, gut halophilic microbiota and akkermansia muciniphila depletion in humans." *International Journal of Obesity* 43(4): 862–71.

Soppa, J, and D Oesterhelt. 1989. "*Halobacterium* sp. GRB: a species to work with!?" *Canadian Journal of Microbiology*.

Sun, Chaomin et al. 2005. "a single gene directs both production and immunity of halocin C8 in a haloarchaeal strain AS7092." *Molecular Microbiology* 57(2): 537–49.

Tang, Weixin et al. 2015. "Applications of the class II lanthipeptide protease licp for sequence-specific, traceless peptide bond cleavage." *Chemical Science* 6(11): 6270–79.

Torreblanca, M., I. Meseguer, and F. Rodriguez-Valera. 1989. "Halocin H6, a bacteriocin from *Haloferax gibbonsii*." *Journal of General Microbiology* 135(10): 2655–61.

Torreblanca, M., I. Meseguer, and F. Rodriquez-Valera. 1990. "Effects of halocin H6 on the morphology of sensitive cells." *Biochemistry and Cell Biology* 68(1): 396–99.

Torreblanca, Marina, Inmaculada Meseguer, and A Ventosal. 1994. "Production of halocin is practically universal feature of archaeal halophilic rods." *Letters in Applied Microbiology* 19: 201–5.

Ulrich, L E, and E V Koonin. 2005. "one-component systems dominate signal transduction in prokaryotes." *Trends in Microbiology* 13(2): 52–56.

Untergasser, Andreas et al. 2012. "Primer3 - new capabilities and interfaces." *Nucleic acids research* 40(15): e115–e115.

Wang, Shengqin et al. 2019. "characterization of the secondary metabolite biosynthetic gene clusters in Archaea." *Computational Biology and Chemistry* 78: 165–69.

Woese, Carl R, and George E Fox. 1977. "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." *Proceedings of the National Academy of Sciences* 74(11): 5088 LP – 5090.

Yang, Xiao, and Wilfred A. Van Der Donk. 2015. "The Michael-type cyclizations in lantibiotic biosynthesis are reversible." *ACS Chem Bio* 10(5): 1234–38.

You, Young Ok, and Wilfred A. van der Donk. 2007. "Mechanistic investigations of the dehydration reaction of lacticin 481 synthetase using site-directed mutagenesis." *Biochemistry* 46(20): 5991–6000.

Yu, Yi, Subha Mukherjee, and Wilfred A van der Donk. 2015. "product formation by the

promiscuous lanthipeptide synthetase ProcM is under kinetic control." *Journal of the American Chemical Society* 137(15): 5140–48.

Zerulla, Karolin, and Jörg Soppa. 2014. "Polyploidy in haloarchaea: advantages for growth and survival." *Frontiers in Microbiology.*5.

Zhang, Qi, Yi Yu, Juan E. Vélasquez, and Wilfred A. Van Der Donk. 2012. "Evolution of lanthipeptide aynthetases." *Proceedings of the National Academy of Sciences of the United States of America* 109(45): 18361–66.

APPENDICES

Appendix 1.

Table S 1: Results of the TBLASTN analysis, using the *halC8* gene from *Natrinema* sp. AS7092 as query. *N. altunense* AJ2 and *Natrinema* sp. J7-2 are identical to *N. altunense* JCM 12890 and *N. gari* JCM 14663, respectively. *N. altunense* 1A4-DGR genome is not annotated.

| *ORGANISM*	LOCATION	LOCUS TAG
Natrinema altunense JCM 12890	Contig 35	C485_RS13570
Natrinema altunense AJ2	Contig 4	ALTAJ2_RS10490
Natrinema altunense 1A4-DGR	Contig00002	Nc 315052 to 315900
Natrinema gari JCM 14663	Contig 52	C486_RS11235
Natrinema pellirubrum DSM 15624	Complete genome	NATPE_RS21555
Natrinema sp. J7-2	Complete genome	NJ7G_RS01900
Haloterrigena salina JCM 13891 (A)	Contig 17	C477_RS04595
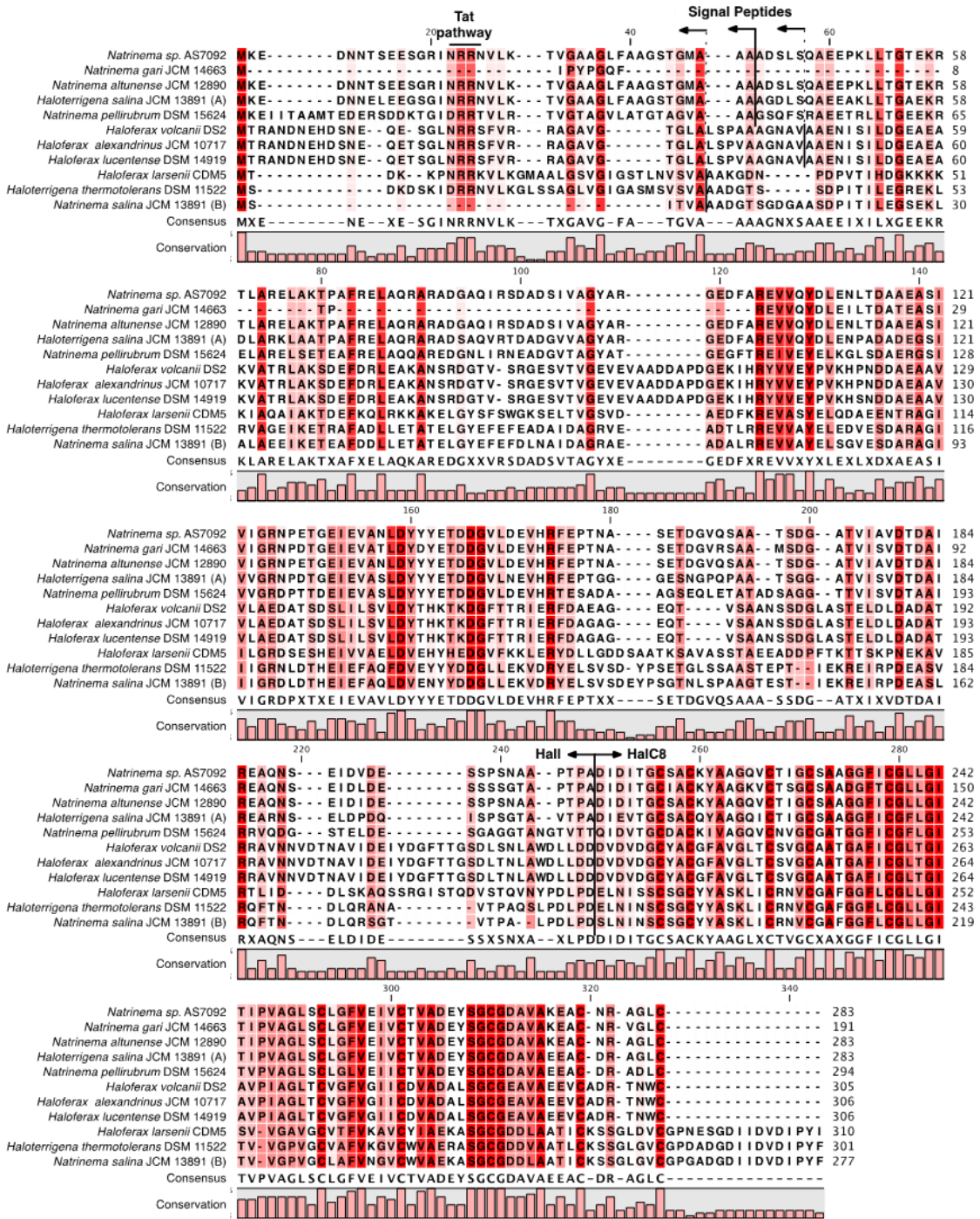
Figure S 1: Alignment of ProC8 peptides with reference to the proteolysis sites that define the Tat pathway signal peptides, the HaII and the HaIC8 peptides. The conserved residues are highlighted in red colour. In ProC8 of N. gari and N. salina (B) no signal peptides were detected with SignalP-5.0 server