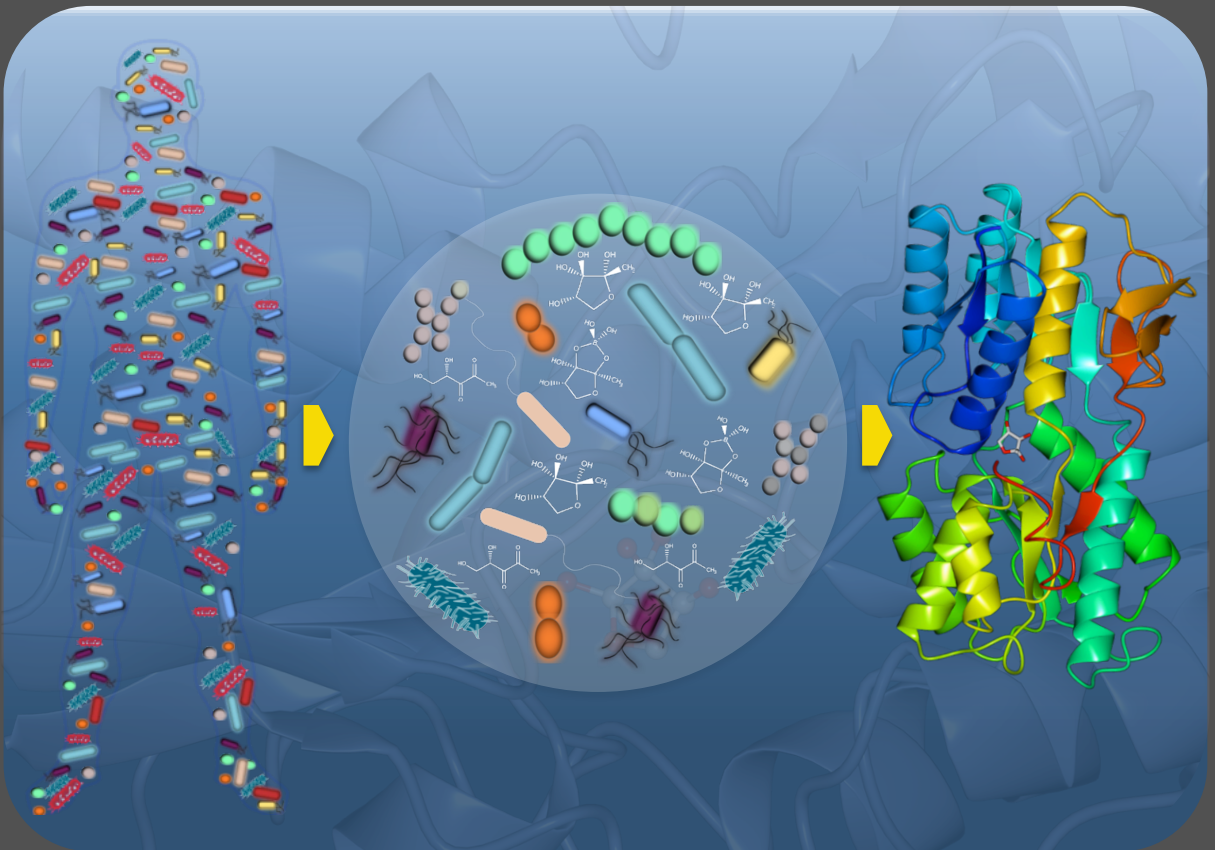


Discovery of novel autoinducer-2 receptors

Inês Mota Torcato



Dissertation presented to obtain the Ph.D degree in Biochemistry
Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
October, 2019



UNIVERSIDADE
NOVA
DE LISBOA

Discovery of novel autoinducer-2 receptors

Inês Mota Torcato

Dissertation presented to obtain the Ph.D degree in Biochemistry

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras, October, 2019

Discovery of novel autoinducer-2 receptors



From left to right: Mariana Pinho, Mónica Serrano, William Bentley, Inês Torcato, Karina Xavier, João Cabral e Sérgio Filipe.

Supervisor:

Karina de Bívar Xavier

Principal investigator in Instituto Gulbenkian de Ciência

President of the Jury:

Mariana Luisa Tomás Gomes Pinho

Associated professor from Instituto de Tecnologia Química e Biológica from Universidade Nova de Lisboa

Examiners:

William E. Bentley

Robert E. Fischell distinguished chair in engineering, University of Maryland

João Henrique Resende de Oliveira de Moraes Cabral

Principal investigator in Instituto de Investigação e Inovação em Saúde from Universidade do Porto

Sérgio Joaquim Raposo Filipe

Auxiliar professor in Faculdade de Ciências e Tecnologia from Universidade NOVA de Lisboa

Mónica Paula Fernandes Serrano

Researcher in Instituto de Tecnologia Química e Biológica António Xavier from Universidade NOVA de Lisboa

Acknowledgments

As in quorum sensing, a group of people and a set of experiences had to get together so I could bring this work to completion. So, I would like to acknowledge:

My supervisor Karina Xavier for all the support and guidance during these years. For allowing me to explore my ideas, for providing me good opportunities, for being able to motivate me when I needed, for always being available to discuss scientific or everyday subjects. Above all, thank you for being an understanding and deeply good person.

My friends and colleagues from MolBioS 2015, the shaking brains, for being the best support network I didn't know I wanted. This journey was infinitely lighter and happier with you.

Bacterial signaling laboratory members for making the laboratory what it is. Joana for always caring and being supportive. Filipe and Rita for being my companions on the PhD journey and for the countless conversations. Vítor for all the "jokes" that made me laugh because they were too bad. Miguel for always giving me a different perspective. Tanja for the joyful companionship and for all the help with the Clostridia. Carina and Margarida for bringing a breath of fresh air to the laboratory. Roberto, for being one of the best scientists I have met and for always having time to give good advice. Vanessa for the friendship, support, laughs and patience in our ventures. Rita Valente for the laughs and for sharing with me the "classics" of quorum sensing. Paulo for inventing the best conversation starter ever. André for the constant good mood. Jess, Joana Dias e Özhan for the good moments.

The members of the Evolution Biology laboratory for helpful advice and great company.

Stephen Miller for sharing his love for proteins with me. Thank you for welcoming me in your lab, for always letting me try every step of the processes and for teaching me how to make them better.

Professor Helena Santos for believing in my potential when many wouldn't. For all the opportunities that you gave me, all that you taught me and for always being a mentor, even when I was no longer in your laboratory. Thank you, also, for the guidance as a member of my thesis committee.

Adriano Henriques for all the fruitful discussions that shaped this work for the better. Thank you, also, for the guidance as a member of my thesis committee.

Rita Ventura for welcoming me in her laboratory and for a rewarding collaboration. Osvaldo for trying to turn me into a synthetic chemist and Eva for the care and advice.

Fundação para a Ciência e Tecnologia for funding my PhD fellowship (PD/BD/105736/2014) within the scope of the PhD program Molecular Biosciences (PD/00133/2012) and the research work through the projects PTDC/BIA-MIC/4188/14 and PTDC/BIA-BQM/29668/2017.

Instituto Gulbenkian de Ciência and Instituto de Tecnologia Química e Biológica for the great research conditions that allowed me to have the best of two worlds.

Ana Aranda and Élio Sucena for the help with my inclusion into the PhD community at IGC.

Raquel, Ioanna, Catalina and the remaining of the 2015 IBB class for the good moments.

My Swarthmore family: Meghann, Judy, Eli and Nick for the adventures and Heather for the affection and care. All of them for the friendship that filled my days.

Nuno, Carla, Ana, Dušica, Dário, Marta, Luís, Sara and Gonçalo from the Cell Physiology and NMR group that helped and encouraged me during my PhD application and that always treated me like a current member when I was at ITQB. A very special thank you to Nuno, Carla and Ana that thought me fundamental techniques that allowed me to hit the ground running on my PhD.

Soni for being the most patient, enthusiastic, kind and thorough first mentor I could have asked for and for the friendship that followed.

My friends that are always there. Particularly, Sara for being my person and Cat for saying the right thing at the right time without even realizing it.

E quero ainda agradecer:

Ao meu pai António, às minhas irmãs Carina e Sílvia, ao meu irmão emprestado Etienne, aos meus sobrinhos David, Rita, Eva, Alexandre e David, à minha família mais recente Natália, Carlos e Luís, à D. Maria, ao Sr. Diamantino e aos meus tios emprestados Zé Manel e Lena por viverem a vida ao meu lado.

À minha mãe que sempre me ensinou que eu podia ser o que eu quisesse.

Ao meu João pelo amor e apoio incondicional e por ter o condão de fazer tudo ficar bem.

"Nothing in life is to be feared, it is only to be understood."

Marie Curie

"Everybody has a plan until they get punched in the mouth."

Mike Tyson

Contents

ABSTRACT.....	XV
RESUMO.....	XVII
THESIS OUTLINE.....	XIX
ABBREVIATIONS AND ACRONYMS.....	XXI
CHAPTER 1. GENERAL INTRODUCTION.....	25
1.1. Strength in numbers	27
1.2. Quorum sensing systems in Gram-negative bacteria	29
1.3. Quorum sensing systems in Gram-positive bacteria.....	32
1.4. Diffusible signal factor mediated quorum sensing	33
1.5. Quorum sensing in multi-species environments	36
1.5.1. Interspecies communication mediated by autoinducer-2	36
1.5.1.1. AI-2 synthesis	37
1.5.1.2. Export of AI-2	39
1.5.1.3. AI-2 detection	40
1.5.2. Quorum sensing networks	50
1.5.3. Listening in on the neighbors	54
1.6. Cell-cell communication in the gut microbiota	56
1.6.1. AI-2 alters the composition of the microbiota in antibiotic-induced dysbiotic gut	57
1.6.2. Interkingdom communication	57
1.6.2.1. Autoinducer-3, host hormones and fucose	58
1.6.2.2. AI-2 mimic produced by the host	60
1.6.2.3. Host immune response to quorum sensing signals	61
1.7. Behaviors regulated by quorum sensing and heterogeneity	62
1.8. Final remarks and main goals	64
CHAPTER 2. IDENTIFICATION OF NOVEL LSRB-LIKE RECEPTORS IN CLOSTRIDIA.....	65
2.1. Abstract	67
2.2. Introduction	67
2.3. Results	72
2.3.1 Identification of LsrB orthologs and in vitro binding to AI-2	72
2.3.2. AI-2 internalization in <i>C. saccharobutylicum</i>	74
2.3.3. <i>C. saccharobutylicum</i> LsrB binds <i>R</i> -THMF	75
2.3.4. The LsrB binding site is more plastic than previously thought	77
2.3.5. LsrB receptors bind AI-2 with high affinity.....	79
2.3.6. CsLsrB shares a common ancestor with canonical LsrB receptors	81
2.3.7. Identification of additional putative noncanonical LsrB receptors	84

2.4. Discussion	85
2.5. Experimental procedures	89
2.5.1. Identification of LsrB orthologs.....	89
2.5.2. AI-2 synthesis	90
2.5.3. AI-2 internalization assays	90
2.5.4. Protein expression and purification	91
2.5.5. Crystallization studies	92
2.5.6. Site directed mutagenesis.....	93
2.5.7. <i>In vitro</i> AI-2 binding assay.....	93
2.5.8. Isothermal titration calorimetry.....	94
2.5.9. Phylogenetic analysis	95
2.5.10. Identification of noncanonical LsrBs	96
2.6. Acknowledgements	97
2.7. Supplementary information	98
2.7.1. Figures	98
2.7.2. Experimental procedures	103
2.7.2.1. Incubation of AI-2 in <i>C. saccharobutylicum</i> supernatants.....	103
2.7.2.2. AI-2 internalization in <i>C. saccharobutylicum</i> and <i>E. coli</i>	103
2.7.1.3. AI-2 internalization in <i>E. coli</i> under anaerobic and aerobic conditions.....	104
2.7.3. Tables	105
CHAPTER 3. DEVELOPMENT OF A CHEMICAL PROBE TO IDENTIFY AI-2 RECEPTORS.....	113
3.1. Abstract	115
3.2. Introduction	115
3.3. Results	118
3.3.1. Solvents used for synthesis of TentaGel based chemical probes denature the AI-2 receptor.....	118
3.3.2. D-Desthiobiotin-AI-2 binds both LuxP and LsrB receptors	121
3.3.3. D-Desthiobiotin-AI-2 can pull-down LuxP	123
3.3.4. AI-2 production profile in <i>C. difficile</i>	126
3.3.5. Strategy to pull-down the AI-2 receptor in <i>C. difficile</i>	127
3.4. Discussion	129
3.5. Experimental procedures	132
3.5.1. Chemical synthesis	132
3.5.2. Protein expression and purification	132
3.5.3. LuxP pull-down using the TentaGel based probe	133
3.5.4. FRET assays to determine LuxP viability	134
3.5.4. Native gel and western-blot	135
3.5.5. LuxP pull-down using the DB-AI-2 probe	136
3.5.6. Assessment of AI-2 production	137

3.5.7. Preparation of <i>C. difficile</i> cell lysates	138
3.5.8. Attempts to pull down <i>C. difficile</i> AI-2 receptor	138
3.6. Acknowledgments	139
3.7. Supplementary information	140
3.7.1. Figures	140
3.7.2. Experimental procedures	142
3.7.2.1. Attempt to pull-down LuxP using streptavidin bound to agarose	142
3.7.2.2. Attempt to pull-down LuxP using streptavidin bound to magnetic beads	143
3.7.2.3. Binding of commercial D-desthiobiotin to streptavidin magnetic particles	144
3.7.2.4. Reversion of the binding order to streptavidin	144
3.7.2.5. Determination of the AI-2 stability	145

CHAPTER 4. EVALUATION OF THE BIOLOGICAL ACTIVITY OF AN OPTICALLY PURE AUTOINDUCER-2 AGONIST.....147

4.1 Abstract	149
4.2. Introduction	149
4.3. Results	152
4.3.1. Chemical synthesis	152
4.3.2. Binding affinity of the DPD analogs to LuxP	153
4.3.3. Biological activity of the DPD analogs	154
4.3.4. Chemical stability	155
4.3.5. Internalization studies	157
4.4. Discussion	158
4.5. Experimental procedures	161
4.5.1. Chemical synthesis	161
4.5.2. LuxP-FRET assay	161
4.5.3. Bioluminescence assay in <i>Vibrio harveyi</i>	162
4.5.4. <i>In vitro</i> chemical stability testing	162
4.5.5. Internalization studies	163
4.6. Acknowledgments	164
4.7. Supplementary material	164
4.7.1. Autoinducer bioassay (AB) medium composition	164
4.7.2. Figures	164
4.7.3. Tables	165

CHAPTER 5. DETECTION AND QUANTIFICATION OF THE QUORUM SENSING SIGNAL AUTOINDUCER-2 IN MICE CAECUM.....167

5.1. Abstract	169
5.2. Introduction	169
5.3. Results and discussion	172

5.3.1. Development of a protocol for AI-2 purification from caecum content	172
5.3.2. Optimization of the removal of high molecular weight components from caecum extracts	176
5.3.3. AI-2 derivatization and ethyl acetate extraction for mass spectrometry analysis...	179
5.3.4. Choosing the collision energy	183
5.3.5. Specificity of AI-2 detection in the biological matrix	185
5.3.6. Improvement of the liquid chromatography and mass spectrometry methods	189
5.4. Conclusion	191
5.5. Experimental procedures	191
5.5.1. Optimized preparation of caecum samples for UHPLC-MS/MS analysis	191
5.5.2. Dimethyl 2,2'-((4,5-diamino-1,2-phenylene)bis(oxy))diacetate synthesis	193
5.5.3. <i>Vibrio harveyi</i> BB170 bioluminescence assay	193
5.5.4. HPLC-MS/MS method used during sample optimization	194
5.5.5. Final UHPLC-MS/MS method used to quantify AI-2	194
5.6. Acknowledgments	195
CHAPTER 6. GENERAL DISCUSSION.....	197
References.....	210

Abstract

Quorum sensing is a bacterial cell-cell communication mechanism mediated by the production and detection of small chemical molecules named autoinducers. This communication mechanism allows bacteria to regulate group behaviors, through synchronization of gene expression, in response to variation in cell density and species composition. The possibility of interfering with quorum sensing to control bacterial behaviors is an attractive alternative to the commonly used antibiotics, to which bacteria develop resistances. Autoinducer-2 (AI-2) mediated quorum sensing is especially interesting as it is involved in interspecies communication in complex biological niches, like the mammalian gut microbiota, where it was shown to affect the bacterial species composition. The AI-2 synthesis pathway is widely conserved, with homologs for the AI-2 synthase, LuxS, identified in more than 500 bacterial species. Two classes of AI-2 receptors have been identified: LuxP and LsrB. However, these receptors are not ubiquitous and there are reports of bacterial species, in which no known AI-2 receptor was identified, that possess phenotypes regulated by AI-2. Given this observation, novel AI-2 receptors are hypothesized to exist. Their identification is the first step to understand the molecular mechanisms that underlie AI-2 mediated communication and which are fundamental for modulation of this process.

The main objective of this work was to identify novel AI-2 receptors. We identified a noncanonical, functional LsrB-type receptor in *Clostridium saccharobutylicum*. Subsequent bioinformatic analysis suggests that these noncanonical LsrB receptors exist in bacterial species in which no receptors were known, as in bacteria belonging to the Spirochaetes, Actinobacteria and predominantly Firmicutes phyla. Moreover, we show that the binding site of LsrB-like receptors is more plastic than previously thought, as naturally occurring substitutions in AI-2-binding amino acid residues was observed. To advance on the identification of other AI-2 receptors that might not be homologous to LsrB, we developed a chemical probe. This probe consists of AI-2 bound to D-desthiobiotin

and is the first probe that binds all the known AI-2 receptors, which suggests that will also be able to bind other unknown receptors. To confirm this hypothesis, preliminary tests to pull-down an unknown AI-2 receptor from the important human pathogen, *Clostridioides difficile*, were performed.

Lastly, due to our interest in the ability of modulating the communication mechanism mediated by AI-2, we studied an AI-2 agonist and started the optimization of a method to quantify AI-2 in mice caecum. The optically pure AI-2 analog is two times more active than AI-2 and cannot be degraded by LsrB-harboring bacteria. Thus, it has the potential to, in lower doses, be effective in modulating AI-2 mediated communication in bacteria with LuxP receptors as the pathogen *Vibrio cholerae* without the interference from LsrB-containing *Enterobacteriaceae*. The processes of extraction and derivatization of AI-2 from mice caecum for mass spectrometry analysis were optimized and further samples will be tested to validate this method to quantify AI-2 directly in the gut.

Overall, this work brings significant developments towards the understanding of the molecular mechanisms that underlie AI-2 mediated quorum sensing and that might be modulated for the development of novel antibacterial approaches.

Resumo

Quorum sensing é um mecanismo bacteriano de comunicação celular mediado pela produção e detecção de pequenas moléculas químicas chamadas de autoindutores. Este mecanismo de comunicação permite que as bactérias regulem comportamentos de grupo, através da sincronização da expressão genética, em resposta à variação da densidade celular ou da composição bacteriana. Assim, a possibilidade de interferir com o processo de *quorum sensing* é uma alternativa bastante atractiva ao uso de antibióticos, aos quais as bactérias desenvolvem resistências. O processo de *quorum sensing* mediado pelo autoindutor-2 (AI-2) é especialmente aprazível, pois está envolvido na comunicação entre bactérias de diferentes espécies, em nichos biológicos complexos, como a microbiota do intestino de mamíferos onde foi mostrado que causa uma alteração na composição bacteriana. A via de síntese do AI-2 é amplamente conservada tendo sido identificados homólogos da sua sintase, LuxS, em mais de 500 espécies bacterianas. Duas classes de receptores para o AI-2 foram identificadas: LuxP e LsrB. Contudo, estes receptores não estão presentes em todas as bactérias que possuem comportamentos regulados por AI-2. Assim, existem bactérias com fenótipos regulados por AI-2 em que nenhum dos receptores conhecidos foi identificado. Estas observações conduzem à hipótese de que novos receptores para o AI-2 existam. A sua identificação será o primeiro passo para perceber os mecanismos moleculares que são a base da comunicação mediada pelo AI-2 e será fundamental para a modulação deste processo.

Desta forma, o principal objectivo deste trabalho foi a descoberta de novos receptores de AI-2. De acordo com este objectivo, identificámos um receptor da família dos receptores LsrB, funcional e não canónico em *Clostridium saccharobutylicum*. Uma análise bioinformática subsequente sugere que estes receptores LsrB, não canónicos, existem em espécies bacterianas com receptores de AI-2 até agora desconhecidos, que pertencem predominantemente ao filo Firmicutes

mas também aos filos Spirochaetae e Actinobacteria. Adicionalmente, mostrámos que o sítio de ligação dos receptores LsrB é mais plástico do que o esperado, pois os receptores não canónicos apresentam substituições nos aminoácidos que ligam AI-2. Desenvolvemos também uma sonda química para identificar novos receptores de AI-2. Esta sonda consiste na molécula de AI-2 ligada a D-destiobiotina e é a primeira sonda a ser capaz de ligar todos os tipos de receptores de AI-2 conhecidos. Esta observação sugere que também será capaz de ligar receptores de AI-2 desconhecidos. Para confirmar esta hipótese, testes preliminares para identificar o receptor de AI-2 do patógeno humano *Clostridioides difficile*, foram levados a cabo.

Por fim, devido ao nosso interesse em modular a comunicação mediada pelo AI-2, estudámos a acção de um análogo de AI-2 e começámos a optimização de um método para quantificar a concentração de AI-2 no ceco de ratinhos. O enantiómero puro do análogo de AI-2 é duas vezes mais activo que AI-2 e não pode ser degradado por bactérias com receptores LsrB. Deste modo, com concentrações mais baixas das que seriam necessárias para o AI-2, poderá modular o processo de comunicação mediado pelo AI-2 em bactérias com os receptores LuxP, como o patógeno *Vibrio cholerae*. Os processos de extracção e derivatização do AI-2 a partir de ceco de ratinhos para análise por espectrometria de massa foi optimizado e mais amostras serão testadas para validar a eficiência deste método que irá permitir a quantificação de AI-2 presente nos intestinos.

Globalmente, este trabalho fornece desenvolvimentos significativos para a compreensão de mecanismos moleculares que permitem a comunicação mediada pelo AI-2 e que no futuro poderão ser modulados para o desenvolvimento de novas terapias antibacterianas.

Thesis outline

In complex, multi-species niches, bacteria need to sense and integrate information coming from various sources. As interspecies communication mediated by autoinducer-2 (AI-2) allows bacteria to sense self and non-self produced AI-2, this communication mechanism might be advantageous in these settings. However, the study of the molecular mechanisms underlying AI-2 mediated quorum sensing has been hindered by the difficulty in identifying AI-2 receptors.

This work was designed to improve the knowledge of AI-2 receptors and of the consequent modulation of bacterial niches, specifically the gut microbiota.

In chapter 1, a review of quorum sensing discoveries with an emphasis on AI-2 receptors was performed. The main goals of the present work are also described in this chapter.

In chapter 2, novel LsrB-like receptors were identified. These new receptors differ from canonical LsrB receptors through naturally occurring mutations on the amino acids involved in AI-2 binding, previously thought as essential. Moreover, novel putative LsrB-like receptors were identified in bacterial species in which no AI-2 receptors had been discovered before, as in species from the Spirochaetes and Actinobacteria phyla.

In chapter 3, a chemical probe able to bind both LuxP and LsrB receptors was developed. This probe has the potential to be a tool for identification of unknown AI-2 receptors. Accordingly, preliminary studies to identify an AI-2 receptor in *Clostridioides difficile*, a bacterium with phenotypes regulated by AI-2, are presented.

In chapter 4, the biological activity of an optically pure AI-2 agonist was assessed. From our results, we propose that this agonist might modulate LuxP mediated responses at lower concentrations than AI-2.

In chapter 5, a procedure to isolate AI-2 from mice intestines was optimized. Additionally, quantification of the isolated AI-2 by ultrahigh pressure liquid

chromatography-tandem mass spectrometry was performed. Upon validation, this method will allow the quantification of AI-2 in the mammalian gut to study physiological responses dependent on AI-2.

The results of this Doctoral work are presented in chapters 2 to 5. Each chapter can be read separately and contains: abstract, introduction, results, discussion and experimental procedures sections.

In chapter 6, the new knowledge produced in this doctoral work was integrated with what is already described in the literature. Furthermore, experiments that could provide relevant additions to the field are proposed in this chapter.

Abbreviations and Acronyms

4-ethoxy-HPD	4-ethoxy-5-hydroxy-2,3-pentadione
4-propoxy-HPD	4-propoxy-5-hydroxy-2,3-pentadione
Acyl-ACP	Acylated acyl carrier protein
AHL	Acyl-homoserine lactones
AI-2	Autoinducer-2
AI-3	Autoinducer-3
AibA, AibB	Autoinducer-2 binding proteins A and B
APCI	Atmospheric pressure chemical ionization
API	Atmospheric pressure ionization
BDSF	<i>cis</i> -2-dodecainoic acid
BSA	Bovine serum albumine
c-di-GMP	Cyclic dimeric guanosine monophosphate
CAI-1	Cholera autoinducer-1
CaLsrB	LsrB from <i>Clostridium autoethanogenum</i>
cAMP	Cyclic adenosine monophosphate
CFP	Cyan fluorescent protein
CLPY	LuxP bound to cyan and yellow fluorescent proteins
CRP	cAMP receptor protein
CsLsrB	LsrB from <i>Clostridium saccharobutylicum</i>
DB	D-desthiobiotin
DB-AI-2	D-desthiobiotin bound to AI-2
DCM	Dichloromethane
DHAP	Dihydroxyacetone phosphate
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DPD	(<i>S</i>)-4,5-dihydroxypentane-2,3-dione

DPO	3,5-dimethylpyrazin-2-ol
DSF	Diffusible signal factor
DTT	Dithiothreitol
EC50	Half maximal effective concentration
EHEC	Enterohemorrhagic <i>Escherichia coli</i>
ESI	Electrospray ionization source
FRET	Förster resonance energy transfer
GF	Germ free
GMP	Guanosine monophosphate
HAI-1	3OHC4-homoserine lactone
HPLC	High-pressure liquid chromatography
HSL	Homoserine lactone
IL	Interleukin
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IQS	Integrating quorum sensing signal
ITC	Isothermal titration calorimetry
KEGG	Kyoto Encyclopedia of Genes and Genomes
<i>m/z</i>	Mass-to-charge ratio
MS/MS	Tandem mass spectrometry
Ni ⁺ -NTA	Nickel-nitrilotriacetic
OD	Optical density
P-AI-2	Phosphorylated autoinducer-2
P-HPD	3-hydroxy-2,4-pentadione-5-phosphate
P-TPO;	3,4,4-trihydroxy-2-pentanone-5-phosphate
PAS	Per/Arnt/Sim
PDB	Protein Data Bank
PQS	Pseudomonas quinolone signal

PTS	Phosphotransferase system
<i>R</i> -THMF	(2 <i>R</i> ,4 <i>S</i>)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran
RMSD	Root mean square deviation
RNA	Ribonucleic acid
<i>S</i> -THMF-borate	(2 <i>S</i> ,4 <i>S</i>)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran-borate
S/N	Signal-to-noise ratio
SAH	<i>S</i> -adenosylhomocysteine
SAM	<i>S</i> -adenosylmethionine
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SPF	Specific pathogen free
SRH	<i>S</i> -ribosylhomocysteine
TBAF	Tetrabutylammonium fluoride
Tdh	Threonine dehydrogenase
TEV	Tobacco etch virus
TG	TentaGel resin
TG-AI-2	TentaGel derivatized with AI-2
THF	Tetrahydrofuran
UHPLC	Ultrahigh performance liquid chromatography
UPLC	Ultra-pressure liquid chromatography
YFP	Yellow fluorescent protein

Chapter



General introduction

1.1. Strength in numbers

Just like a single human, a small, isolated bacterium is not able to thrive successfully. Nevertheless, it has been shown that, as with humans, strength in numbers can come a long way. This is the reason why bacteria wait until they have the right numbers to engage in group behaviors that would be otherwise unsuccessful. This coordination is only possible because bacteria are able to sense and communicate with each other through a “language” named quorum sensing.

Quorum sensing is a cell-cell communication mechanism mediated by the production, release, accumulation and detection of small, extracellular molecules. These molecules, named autoinducers, accumulate in the environment as the density of the population increases (1,2). When a threshold concentration is reached, the autoinducers are detectable by bacterial populations through interaction with a receptor protein, allowing bacteria to monitor their cell numbers and to alter global patterns of gene expression, accordingly (1,3). This change in gene expression enables bacteria to regulate group behaviors like bioluminescence, biofilm formation, virulence factors expression and antimicrobial compounds production (2,3).

However, almost 300 years passed from the discovery of bacteria by Antony van Leeuwenhoek in 1684 to the first reports of intercellular communication within a bacterial population (4). During this time, bacteria were regarded as asocial beings with more interest as microbiology tools than as object of organizational-based studies. It was with the discovery that emergence of competence during growth of *Streptococcus pneumoniae* was regulated by two macromolecular cell products by Alexander Tomaz in 1964 (5) and the coining of the term autoinduction by Kenneth Nealson in 1970 (1) when studying bioluminescence in *Vibrio fischeri* and *Vibrio harveyi*, that the concept of intercellular communication in bacterial suspensions became known. Nevertheless, for two decades more this communication process was considered

to be an anomaly that only existed in a small subset of bacteria (4). In the 1980's, identification of the genes necessary for quorum control of luminescence (6) and identification of the first chemical structure of an autoinducer molecule, the quorum sensing signal from *V. fischeri* (7) were important landmarks in the field. However, only in the 1990's a significant expansion was observed. This was a result of the introduction of DNA sequencing and comparative sequence analysis as everyday laboratory tools (8). The utilization of these tools enabled the identification of LuxI (autoinducer synthase) and LuxR (transcriptional regulator of luminescence) homologs in diverse bacterial species by William Claiborne Fuqua in 1994 (9). From this point on, it was clear that many Gram-negative bacterial species controlled genes for group behaviors through LuxI-LuxR systems. Also, small peptides were identified as autoinducer molecules in *S. pneumoniae* (10) and *Staphylococcus aureus* (11), Gram-positive bacteria. Additionally, the observation that *V. harveyi* responded to two different autoinducers led to the discovery of autoinducer-2 (AI-2)(12). This molecule was postulated to be an interspecies quorum sensing molecule as AI-2 produced by one bacterium could be recognized by other species of bacteria (13,14). Additionally, the machinery for production and detection of AI-2 was identified in both Gram-negative and -positive bacteria indicating communication across species boundaries (15-17).

With these findings, it became established that bacteria communicate with their neighbors through sophisticated communication mechanisms. It is now known that quorum sensing is much more complex than previously thought and that is only one of the cell-cell mechanisms that bacterial cells use to interact and regulate group-behaviors. Recent reports of interkingdom quorum sensing mechanisms and identification of quorum sensing like mechanisms in archaea, fungi, viruses and even in ants (3,8,18) shows the relevance of the role of density dependent responses across different domains of life and that we are still only in the initial steps of understanding how microorganisms communicate. Moreover,

exciting studies on quorum sensing systems in dynamically changing environments and in multi-species bacterial communities are being conducted (19-21) and will most certainly provide exciting new developments.

1.2. Quorum sensing systems in Gram-negative bacteria

The quorum sensing system considered the paradigm for Gram-negative bacteria was also the first to be discovered (22). *V. fischeri*, a marine bacterium, controls the synthesis of luciferase through a quorum sensing system regulated by a LuxI-LuxR (synthase/receptor) pair. The production of light by this bacterium is very interesting, as it allows *V. fischeri* to have a symbiotic relationship with the Hawaiian squid *Euprymna scolopes*. *V. fischeri* colonizes the light organ of this squid and is able to grow to densities unachievable in sea water due to the richness in nutrients present in this organ. In turn, it prevents the squid from being detected by predators because the light emitted by bacteria will counterbalance the light from the moon, avoiding the formation of the squid's shadow (2,23).

Many systems that use LuxI-LuxR homologs have been identified in Gram-negative bacteria (2,9). LuxI catalyzes the synthesis of the most common class of autoinducers in Gram-negative bacteria - acyl-homoserine lactones (AHL) - using S-adenosylmethionine (SAM) and an acylated acyl carrier protein (acyl-ACP) as the substrates (24,25). These autoinducers have an invariable homoserine-lactone ring (HSL; acquired from SAM) attached to acyl chains of variable length that can comprise 4-18 carbon atoms (3). The acyl chains, which usually are derived from fatty acid biosynthesis, can have modifications or different saturation states (2,9). Each Gram-negative bacterial species mainly produces only a specific type of AHL, although, some species can also produce additional AHLs in smaller quantities (2,24). Upon production, the majority of these molecules can diffuse freely through the bacterial membrane and their concentration will increase in the environment with the increase of bacterial density (24). Further, these systems

usually possess a feed-forward loop of activation of the autoinducer synthase. Thus, the detection of autoinducer will promote the synthesis of more autoinducer (2,3).

Although this is what is predominantly observed, there are some exceptions. For example, LuxM and AinS catalyze the same reaction as LuxI in *V. harveyi* and *V. fischeri*, respectively, but do not share homology with this class of enzymes (24). *Rhodopseudomonas palustris*, *Bradyrhizobium* spp. and *Aeromonas* spp. produce the uncommon autoinducers p-coumaroyl-homoserine lactone, isovaleryl-HSL and cinnamoyl-HSL, respectively. Interestingly, the synthesis of p-coumaroyl-homoserine lactone, where the acyl group is derived from plant-derived p-coumarate, allows *R. palustris* to combine the detection of bacterial population density to availability of plant consumables (3,18). Moreover, it has been suggested that several AHLs may be actively transported across the cell membranes (24,26).

After accumulating in the environment, the AHL produced will bind its cognate receptor, usually a LuxR receptor (Fig. 1.1). LuxR-type receptors are cytoplasmic transcription factors that bind both AHLs and DNA (3,24). Upon binding of the AHL signal to the amino-terminal domain of the receptor, it will form stable homodimers and will bind *lux boxes* upstream of target genes, regulating their expression (3). However, in the absence of the AHL signal most LuxR receptors fail to fold and are degraded (3,27). The AHL-LuxR binding is generally highly specific due to a combination of amino acid variation and flexibility in the binding pocket of the LuxR receptor. This allows specific intraspecies communication even in the presence of molecules with only small differences in the acyl chain (3,24). It has been shown that polar residues in the ligand binding site, namely three tryptophan residues, interact with the HSL moiety to define binding orientation. Moreover, it has been observed that the acyl chains can either extend and point toward the solvent (short acyl chains) or bend and face the interior of the pocket (long acyl chains) (3). Alternatively, AHLs can

be detected by two-component membrane-bound histidine kinases that signal to cytoplasmic transcription factors through phosphorylation, such as LuxN of *V. harveyi* (Fig. 1.1). These membrane receptors are strictly specific for their cognate ligand as well (24). In fact, it has been shown that LuxN is not activated by any variant of the cognate ligand and that AHLs with longer acyl chains act as potent antagonists (3). The observation that other AHL molecules block the quorum sensing response of *V. harveyi* further exemplifies the complexity of bacterial communication. It suggests that in the presence of bacterial competitors, *V. harveyi* stops the quorum sensing process to avoid exploitation of produced public goods (3).

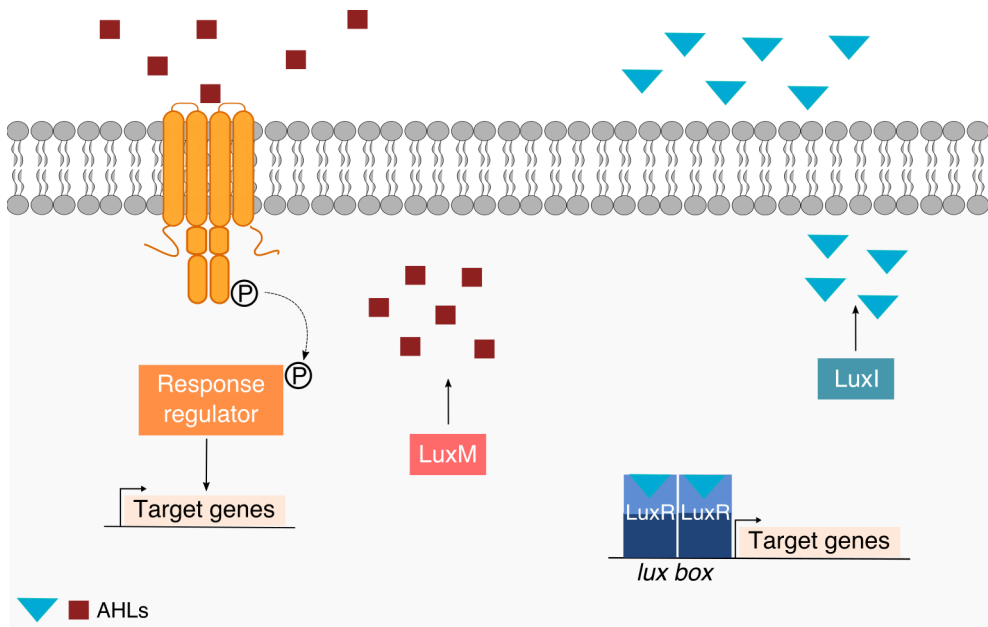


Figure 1.1 - Quorum sensing in Gram-negative bacteria. Acyl-homoserine lactones (AHLs) are the most common quorum sensing signals used for communication in Gram-negative bacteria. They are usually synthesized by LuxI synthases and recognized by LuxR receptors, cytoplasmic transcription factors. However, some exceptions exist. For example, in *Vibrio harveyi*, an AHL is produced by the LuxM synthase, which is not a homologue of LuxI synthases. This class of AHLs is then recognized by a histidine kinase two component system. Due to their chemical properties, most AHLs can diffuse freely across the membrane accumulating both inside and outside the cell.

1.3. Quorum sensing systems in Gram-positive bacteria

Quorum sensing in Gram-negative and -positive bacteria is mediated by completely different autoinducers. Gram-positive bacteria use modified peptides, also named pheromones, as autoinducers (2,24,28). These signal peptides are mainly encoded by short open reading frames and are translated by ribosomes inside the cell (24,28). Factors that regulate the expression of these peptides are still not well known but it was shown that, at least in some cases, there is a positive feedback loop as it is observed for AHL signaling. More recently, a new source of autoinducing peptides was found in *Enterococcus faecalis* and *Streptococcus pyogenes*. These bacterial species possess short, hydrophobic peptides that can be cleaved and repurposed as signal peptides (28). Thus, proteolytic degradation of surface proteins, more specifically of the signaling sequences that signal the export of proteins, can also be a source of signaling peptides (28).

Each bacterial species produce a specific signal that is recognized with high specificity by cognate receptors and thus, similarly to AHL communication, autoinducing peptides are utilized in intraspecies communication (2,24,28). Since these molecules are unable to permeate the cell membrane they have to be exported through specialized transporters that are usually encoded in close proximity to the peptide encoding genes (28). In between translation, export and detection these peptides are subject to various modification events that include cleavage from larger precursor peptides, cyclization and/or inclusion of chemical groups like lactone and thiolactone rings (2,24). After export, the concentration of signaling peptides will increase until the concentration required for detection is achieved. At this point, the signaling peptides can be detected through either, an extracellular or an intracellular detection pathway (Fig. 1.2) (28). In the extracellular pathway the signaling peptides interact with a membrane sensor histidine kinase of a two-component signal transduction system. This will induce the autophosphorylation of the histidine kinase, which will induce the phosphorylation of the response regulator, activating it. The activated response

regulator will then modulate the expression of the target genes. In the intracellular pathway, linear autoinducer peptides are re-imported into the cell through an oligopeptide transporter system and either activate transcriptional regulators that will then modulate target gene expression or they inhibit aspartyl phosphate phosphatase thus modifying the phosphorylation state of regulators and consequently the expression of their target genes (28).

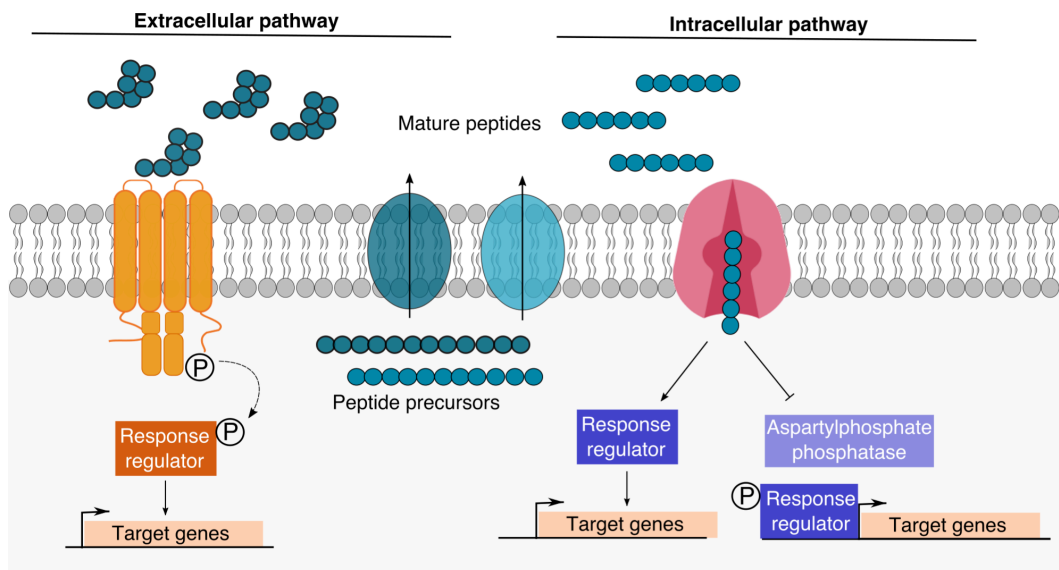


Figure 1.2 - Quorum sensing in Gram-positive bacteria. Autoinducing peptides are the most common quorum sensing signals used for communication by Gram-positive bacteria. They are usually subjected to modifications like cleavage from larger precursor peptides and cyclization. Due to their chemical properties they are actively exported from the cell and are usually recognized by histidine kinase two component systems. The binding of these peptides will elicit a phosphorelay response that leads to regulation of target genes (extracellular pathway). Alternatively, linear peptides can be re-imported to the cell where they either inhibit the activity of aspartylphosphate phosphatase or directly bind to transcriptional regulators to regulate the expression of target genes (intracellular pathway).

1.4. Diffusible signal factor mediated quorum sensing

The diffusible signal factor (DSF) comprises a family of cis-2-unsaturated fatty acid molecules that are responsible for regulation of motility, virulence, biofilm formation, iron uptake, extracellular polysaccharide and extracellular

enzyme production (29). Although widely conserved in Gram-negative bacteria like *Xanthomonas campestris*, *Pseudomonas aeruginosa* and *Burkholderia cenocepacia*, emerging evidence revealed that this mechanism is involved not only in intraspecies communication but also in interspecies and interkingdom communication mechanisms. In *X. campestris* pv. *campestris*, DSF signaling seems to give this bacterium a fitness advantage in interspecies competition (30). Production of DSF-type molecules allowed *X. campestris* to interfere with morphological transition and sporulation in *Bacillus thuringiensis* by modulating the expression of *ftsZ*, which encodes a key player in bacterial cell division (30). Additionally, BDSF (*cis*-2-dodecenoic acid) produced by *Burkholderia* species has been shown to stimulate biofilm dispersion, production of molecules involved in stress response and upregulation of iron uptake by inducing siderophore production in *Francisella novicida* (29,31). Moreover, it has been shown that DSF signals produced by *Streptococcus mutans* and by *B. cenocepacia* inhibits *Candida albicans* morphological transitions (32,33).

DSF-mediated quorum sensing can be grouped into three categories taking into account their genomic context. The first is the paradigm for DSF-based quorum sensing and has been functionally verified in *Xanthomonas* sp., *Xylella fastidiosa*, *Lysobacter enzymogenes* and *Stenotrophomonas maltophilia* (29). In this category the genes that code for key signaling components RpfF, RpfC and RpfG are co-localized in the *rpf* gene cluster (29). *rpfF* encodes the DSF synthase and *rpfC* and *rpfG* encode the components of a two-component system responsible for detection of the signals (34,35). Similarly to what occurs with other quorum sensing mechanisms, DSF will accumulate in the extracellular environment and will bind the RpfC sensor histidine kinase. Upon binding, a phosphorelay cascade will activate the GYP domain of the response regulator RpfG that functions as a phosphodiesterase degrading cyclic dimeric guanosine monophosphate (c-di-GMP) to GMP (29,36). Since c-di-GMP is an inhibitory ligand of global transcription factor Clp (for CRP-like protein), the decrease in c-di-GMP levels will

lead to de-repression of Clp that will modulate the expression of several hundred genes including those encoding virulence factor production (29,36). The observation that RpfF has a broad substrate specificity as it uses acyl-ACP substrates with 8 to 14 carbon atoms, might explain the production of multiple DSF signaling molecules by the same bacterial species (37). The second category of DSF-mediated quorum sensing has been functionally characterized in the opportunistic pathogens *B. cenocepacia* (38) and *Cronobacter turicensis* (39). In this category, there is no *rpf* cluster, only the gene *rpfF*, which encodes the synthase and a novel receptor gene *rpfR* in the same locus. This receptor protein is a modular protein that possesses a Per/Arnt/Sim (PAS) domain where the BDSF signal binds and GGDEF and EAL domains that have diguanylate cyclase and phosphodiesterase activities, respectively (40). Thus, RpfR is a one-component signal transduction system that can sense the quorum sensing signal and directly modulate the levels of c-di-GMP (40). Although there is a different detection system, both categories lead to the same end result: altered levels of c-di-GMP (30). Accordingly, in *B. cenocepacia*, upon accumulation of BDSF, this signal binds the RpfR receptor stimulating its phosphodiesterase activity (38). The third category of DSF signaling has been identified in *P. aeruginosa*. In this category, the biosynthesis of the DSF-type molecule has been attributed to DspI a putative enoyl-coenzyme A hydratase. Although no mechanism of detection has been identified, a cluster of five genes (PA4978–PA4983) has been proposed to be important for the DSF mediated quorum sensing in this organism (41,42).

Interestingly, a turnover mechanism for DSF production has been proposed (29). RpfB, encoded upstream of RpfF, is proposed to be a fatty acid Acyl-CoA ligase involved in DSF processing (43). It has been shown that this enzyme is regulated by DSF in a concentration-dependent manner and that Clp represses its expression through a direct interaction with a conserved DNA motif on the *rpfB* promoter (44). Moreover, homologues for *rpfB* have been found in bacteria with the three categories of DSF mediated quorum sensing (29). More studies have to

be performed to fully understand the mechanism behind this turnover process and its biological importance, but the rapid elimination of the signal might suggest that this process could be important to quickly alternate between behaviors induced or repressed by DSF mediated quorum sensing (29).

1.5. Quorum sensing in multi-species environments

Bacteria in complex multi-species environments will probably have contact with a complex mixture of autoinducers self-produced, produced by their clonal siblings, by other cells from the same bacterial species and also by members from distantly related competitors belonging to other bacterial species (3). The necessity of understanding and differentiating information coming from various sources, might explain the presence of different autoinducer production and detection systems in the same bacterium (3). The initial studies on quorum sensing systems were fundamental to discover the key players, how they are regulated and the behaviors that they regulate. They were, however, usually performed in well-mixed pure laboratory cultures (19). Currently, a concern for studying quorum sensing processes in naturally occurring situations exists. These situations include mixtures of bacterial species, non-ideal conditions with internal (such as transcript levels and protein numbers) and external (temperature, pH, osmolarity) fluctuations or/and bacterial structured surface-bound communities (biofilms) (3,19). Thus, in this next section, communication mechanisms important for complex, multi-species environments will be explored.

1.5.1. Interspecies communication mediated by autoinducer-2

In the early 1990's, Bassler et al proposed that two signal-response pathways would converge to regulate luminescence expression in *V. harveyi* (12,45). The second signal was named AI-2 and it was proposed to be involved in interspecies communication. This role in interspecies communication was proposed due to the

observation that cell-free supernatants of different bacterial species could induce bioluminescence in *V. harveyi*, suggesting that different bacterial strains produced AI-2 (14). Nowadays, it is known that AI-2 actually comprises a family of molecules that derive from a linear precursor synthesized by LuxS, a widely conserved synthase. These molecules can be rapidly interconverted in aqueous solution and are fundamental for interspecies communication in many multi-species environments such as the mammalian gut microbiota (approached in section 1.6.1) (21).

1.5.1.1. AI-2 synthesis

SAM, a metabolite from central metabolism, is the substrate for the synthesis of AI-2. The consumption of SAM as a methyl donor results in S-adenosylhomocysteine (SAH) that is metabolized to adenine and S-ribosylhomocysteine (SRH). SRH is then cleaved by LuxS generating homocysteine and 4,5-dihydroxy-2,3-pentanedione (DPD). DPD is the linear precursor of AI-2 that, in solution, cyclizes spontaneously to form (2S, 4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran-borate (S-THMF-borate) and (2R,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran (R-THMF) (16). Since these molecules are interconvertible, an equilibrium between the different molecular forms exists in solution (Fig. 1.3). This equilibrium made the identification of the molecular forms active in the signaling process very challenging (46). However, crystallization of AI-2 receptors with bound ligand enabled the identification of S-THMF-borate and R-THMF as the active forms of AI-2 (46,47). It had been previously suggested that a mix of furanones could result from spontaneous cyclization of DPD as it is a highly reactive molecule that should not be stable in aqueous solution (48). Still, the identification of a borated furanone was surprising as the conjugation of furanoid rings with boron is rarely observed in nature (47). Nevertheless, S-THMF-borate is recognized by the marine bacterium *V. harveyi* and in marine environments the concentration of borate is abundant (approximately 0.4 mM)

(47). As borate can readily react with adjacent hydroxyl groups on the furanosyl rings, *S*-THMF-borate can be easily formed (46). Further studies showed that addition/removal of boron could dislocate the equilibrium. Thus, in the presence of boric acid, the predominant form of AI-2 is *S*-THMF-borate whereas removal of boron dislocates the equilibrium in the direction of *R*-THMF and *S*-THMF (46).

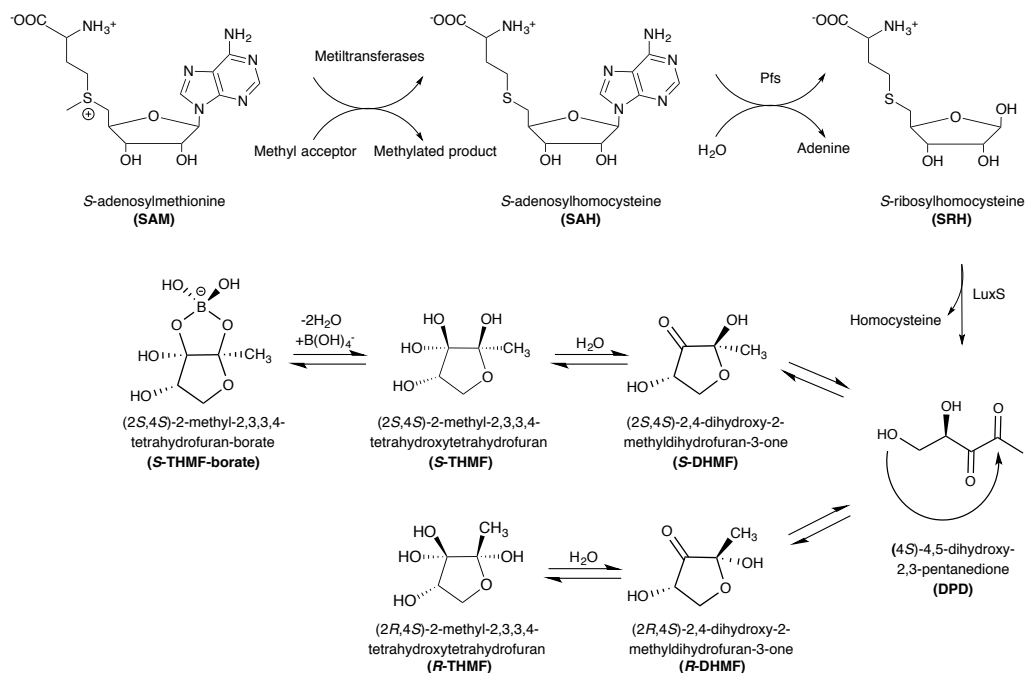


Figure 1.3 - AI-2 synthesis. The removal of a methyl group from SAM results in SAH that is metabolized to adenine and SRH. SRH is then cleaved by LuxS to homocysteine and DPD, the precursor molecule of the signaling forms of AI-2: *R*-THMF and *S*-THMF-borate.

The observation that AI-2 was a product from a crucial central metabolic pathway (activated methyl cycle) and that the substrate of LuxS is originated from the toxic intermediate SAH gave rise to the possibility that AI-2 was a by-product of metabolism (49). However, improvements in AI-2 synthesis and discoveries regarding AI-2 receptors and other players in the signaling process support the role of AI-2 as a signal molecule (50). It was shown that construction of *luxS* mutants could promote changes related to metabolic processes like methionine

metabolism, recycling of homocysteine or accumulation of intermediates of the SAM metabolism (49). Thus, to dissociate the metabolic effects from the signaling ones, complementation of the observed effect by addition of the signal was fundamental (50). The advances on the synthesis of DPD and even the commercial availability of this signal made this process much easier (51). In addition, advances on the understanding of the detection and processing mechanisms of the signal enabled the construction of mutants in key players other than the synthase (52). Still, SAM is also the substrate for the synthesis of AHLs and cholera autoinducer-1 (CAI-1). This common link between quorum sensing and SAM suggests a benefit in connecting the central metabolism of the cell and quorum sensing responses (50).

AI-2 synthase, LuxS, is widely conserved both in Gram-negative and -positive bacteria and homologs have been identified in more than 500 bacterial species (3). This makes AI-2 the most common bacterial autoinducer identified to date (3). Moreover, the observation that the linear form of DPD can originate different signaling forms of AI-2 further supported AI-2 as an interspecies communication signal as many species can synthesize the linear form that will then acquire the conformations recognized by the different receptors.

1.5.1.2. Export of AI-2

As mentioned before, autoinducers accumulate in the extracellular environment. The hydrophilic nature of AI-2 suggests the existence of a transporter responsible for transporting AI-2 across the membrane (53). YdgG, also named transport of quorum sensing signal TqsA, from *Escherichia coli* K-12 BW25113 was proposed to export AI-2 (54,55). This comes from the observation that in an *ydgG* mutant there was less extracellular AI-2 activity and more intracellular AI-2 activity than in the wild type (54). However, a study following the extracellular AI-2 accumulation through time in both wild type and the *ydgG* mutant was not included. This study would provide more accurate evidence of

the effect of *ydgG* deletion in AI-2 extracellular concentration. Moreover, the observation that deleting an *ydgG* homolog gene in *Klebsiella pneumoniae* would only promote a non-statistically significant decrease in the extracellular AI-2 concentration (56) and the lack of biochemical proof of AI-2/YdgG binding, suggests that further studies are needed to clarify the exporting mechanism for AI-2.

1.5.1.3. AI-2 detection

Although the synthetic pathway of AI-2 is widely conserved and LuxS homologs have been found in more than 500 bacteria, the same is not true for the AI-2 receptors. Two types of receptors have been identified: LuxP and LsrB (46,47,50). They bind the signaling forms of AI-2, S-THMF-borate and R-THMF, and interact with different effectors, transducing the information through different pathways (46,47,50). Moreover, there are many bacteria with phenotypes regulated by AI-2 that do not possess these receptors. For example, the production of toxins A and B in *Clostridioides difficile* (57) and phage release in *Enterococcus faecalis* (58) have been proposed to be regulated by AI-2. However, neither LuxP nor LsrB receptors have been identified in these bacterial species. Hence, variants of these types of receptors or even novel families of receptors might exist. The low sequence homology between LsrB and LuxP (11%; (46)) suggests that other receptors with low homology and thus, unable to be identified through sequence analysis, might exist. Additionally, although LuxP and LsrB share the same fold, so do many sugar binding proteins. This makes fold comparison insufficient for identification of AI-2 receptors. So, new strategies to identify AI-2 receptors must be developed. A review of what is known for LuxP and LsrB receptors is presented next.

LuxP receptors

The first AI-2 receptor to be identified and to have its structure solved was LuxP from *V. harveyi* (12,47). LuxP belongs to the substrate-binding protein family that has high affinity for its cognate ligands, which explains how it was possible to identify *S*-THMF-borate as the signaling molecule bound to this receptor after extensive purification. LuxP has a classic class I substrate-binding protein fold where two similar α/β domains are connected through a three-stranded hinge (Fig. 1.4, left) (47).

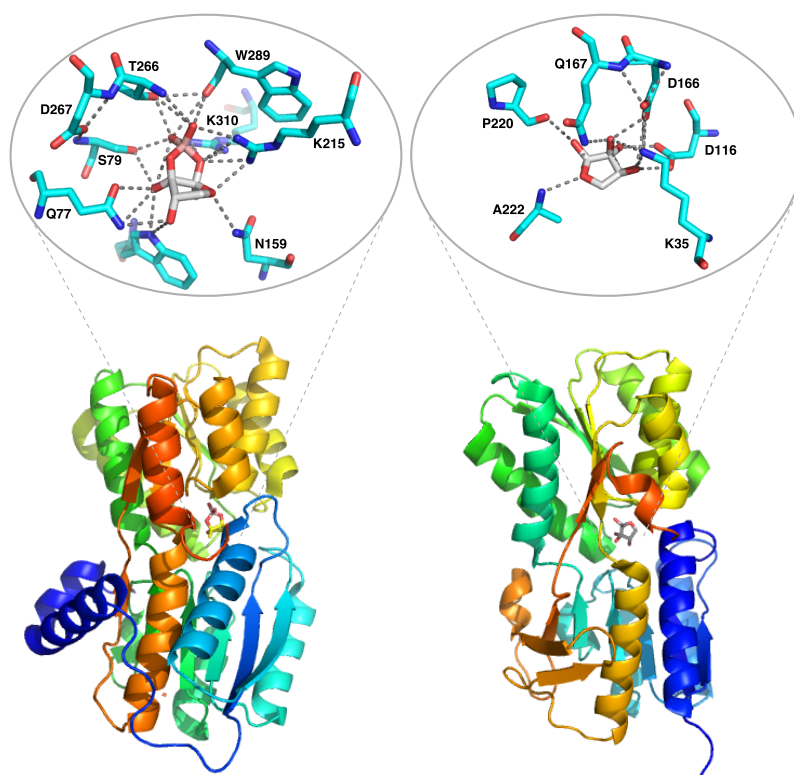


Figure 1.4 - AI-2 receptors. LuxP receptor from *Vibrio harveyi* (left) and LsrB receptor from *Salmonella enterica* serovar Typhimurium (right). LuxP receptors bind *S*-THMF-borate whereas LsrB receptors bind *R*-THMF. A close up of the binding of amino acid residues from the binding site to the active forms of AI-2 are shown. The proteins are presented as cartoon, colored from blue to red, from the amino terminus to the carboxyl terminus. *S*-THMF-borate and *R*-THMF are presented as stick.

The binding site is situated near the hinge between the two domains and binding of AI-2 promotes significant conformational changes. So, in the absence of AI-2, the domains are in an open conformation leaving the binding site exposed to the solvent. However, upon binding of AI-2 the domains close around the binding site (46). The binding site of LuxP has two positively charged amino acid residues (R215 and R310) that stabilize the negative charge on S-THMF-borate. Moreover, residues Q77, S79, T266, and W82 are also key residues in the binding site of LuxP that form hydrogen bonds with S-THMF-borate (Fig. 1.4, left)(46).

Curiously, LuxP interacts with a histidine kinase two-component system. LuxQ, the membrane sensor histidine kinase has two PAS domains (59). One of them usually binds the ligand(s). However, in this case, it binds LuxP instead. Although both PAS domains and substrate binding proteins are ubiquitous, they were never found to interact until the discovery of this quorum sensing pathway (59). LuxP exists in complex with the periplasmic domain of LuxQ. In the absence of bound AI-2, at low cell density, LuxP is in its open conformation and its amino terminus is in close contact with the LuxQ groove, forming a clasp that locks the LuxPQ complex into a conformation unable to form stable dimers with another LuxPQ. In this conformation, LuxQ is locked in its kinase activity. Binding of AI-2, at high cell density, promotes a conformational change in LuxP, as mentioned above, that is proposed to release the clasp (60). This release allows LuxPQ to form an asymmetric dimer with other LuxPQ, promoting the change from kinase to phosphatase activity and the ultimate regulation of target genes (this signaling pathway is approached in depth in section 1.5.2) (60). The dimer (LuxPQ)₂ formed will be asymmetric due to the fact that only one of the LuxP subunits is able to bind both LuxQ proteins simultaneously (60).

Until now, LuxP has only been identified in the Vibrionales, where AI-2 has been shown to regulate behaviors like bioluminescence, biofilm formation, competence and production of proteases, siderophores and virulence factors (3,50).

LsrB receptors

After identification of the LuxP receptor in *V. harveyi*, genes involved in AI-2 detection, transport and metabolism in *E. coli* K-12 MG1655 and *Salmonella enterica* serovar Typhimurium were identified (52,61-63). The homology of LsrB to a substrate-binding protein made it the perfect candidate for AI-2 receptor in these organisms. Indeed, the crystal structure of the LsrB from *Salmonella* serovar Typhimurium was the first tridimensional structure of an LsrB receptor to be solved, confirming its function as an AI-2 receptor (46). Later, structures of the LsrBs from *Salmonella* serovar Typhi (PDB ID 5GTA), *Yersinia pestis* (64), *Sinorhizobium meliloti* (65) and *Bacillus anthracis* (PDB ID 4PZ0) have also been solved. LsrB has a fold identical to LuxP, with the two α/β domains connected through a three-stranded hinge. However, LsrB recognizes R-THMF as the signaling form of AI-2, instead of S-THMF-borate, and thereby important differences exist in the binding site (Fig. 1.4, right)(46). The overall charge of the LsrB binding site is negative, contrarily to LuxP that has a positive charge that stabilizes S-THMF-borate (46). Additionally, the conserved amino acids in the binding site of LuxP are not conserved in the binding site of LsrB. Instead, six amino acid residues, K35, D116, D166, Q167, P220 and A222 (using the LsrB from *Salmonella* serovar Typhimurium as reference) were shown to be key, conserved residues that form hydrogen bounds with R-THMF (17,46,65).

Unlike bacteria with LuxP receptors, the signaling mediated by AI-2 in bacteria with LsrB receptors involves AI-2 internalization through an ABC transporter (Fig. 1.5) (63). Thus, AI-2 will accumulate in the extracellular milieu, in proportion to bacterial cell density, until a concentration high enough to allow binding of AI-2 to LsrB expressed at basal levels, is reached. At this point, AI-2 will bind LsrB that will transfer AI-2 to the transmembrane domain of the ABC transporter formed by LsrC and LsrD. Upon ATP hydrolysis by LsrA, a conformational change in the ABC transporter will allow the entry of AI-2 in the cell (52,63). After entering the cell, AI-2 will be phosphorylated by LsrK and thus,

effectively trapped inside the cell (62,66,67). Phosphorylated AI-2 (P-AI-2) will bind LsrR, a transcriptional regulator, blocking repression and leading to the expression of LsrB and of the transporter (68,69). The increase in expression of these genes promotes AI-2 internalization and consequent depletion of extracellular AI-2. In *E. coli* P-AI-2 is further processed by LsrG which catalyzes the isomerization of P-AI-2 to 3-hydroxy-2,4-pentadione-5-phosphate (P-HPD), an isomer that exists in equilibrium with its hydrated form 3,4,4-trihydroxy-2-pentanone-5-phosphate (P-TPO) (70). LsrF then catalyzes the transfer of an acetyl group from P-HPD to coenzyme A, producing dihydroxyacetone phosphate (DHAP) and acetyl-CoA, which are key metabolites used by the cell in metabolic pathways such as glycolysis and the citric acid cycle (Fig. 1.5) (71).

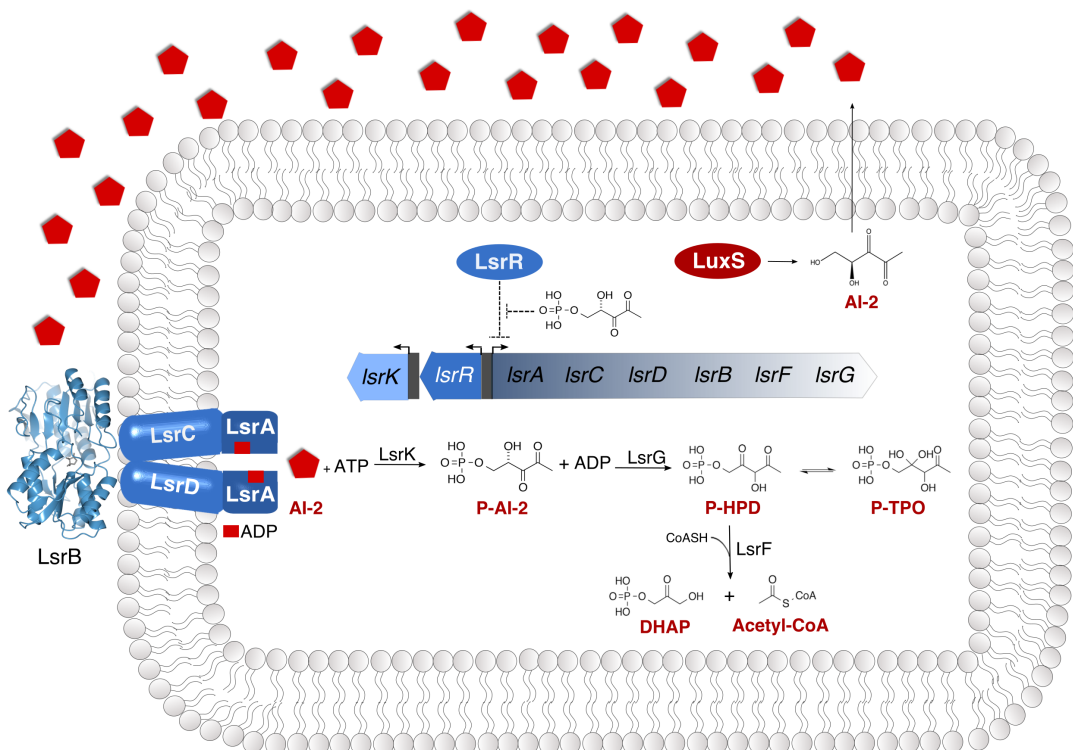


Figure 1.5 - Representation of the detection, internalization and processing of AI-2 in bacteria with LsrB receptors. LuxS catalyzes the synthesis of AI-2, in its linear form. AI-2 is then actively transported to the extracellular environment possibly through YdgA. Extracellular AI-2 will accumulate with the increase in bacterial density until a threshold concentration is reached. At this concentration, AI-2 will be internalized and phosphorylated by LsrK. Upon phosphorylation (P-AI-

2), AI-2 is trapped inside the cell and will alleviate the repression exerted by LsrR. The consequent expression of the LsrB receptor and of the associated ABC transporter (*lsrACD*) promotes AI-2 internalization and depletion of extracellular AI-2. In bacteria with LsrG and LsrF enzymes, P-AI-2 is further processed. LsrG catalyzes the isomerization of P-AI-2 to 3-hydroxy-2,4-pentadione-5-phosphate (P-HPD), an isomer that exists in equilibrium with its hydrated form 3,4,4-trihydroxy-2-pentanone-5-phosphate (P-TPO). LsrF then catalyzes the transfer of an acetyl group from P-HPD to coenzyme A, creating dihydroxyacetone phosphate (DHAP) and acetyl-CoA, which are key metabolites used by the cell in metabolic pathways like glycolysis and the citric acid cycle.

This pathway has been identified in many bacterial species and the genes for detection, transport, regulation and phosphorylation are usually conserved and encoded in the *lsr* (for LuxS regulated) operon (50). In bacteria that also have the enzymes involved in AI-2 metabolism, their genes are also encoded in the *lsr* operon (70,71). In *E. coli*, the Lsr system besides being regulated by LsrR, it is also regulated by catabolite repression and by the phosphoenolpyruvate phosphotransferase (PTS) system (52,72,73). Thus, in the presence of bacterially favored sugars like glucose, the Lsr system is not activated and AI-2 will remain in the extracellular environment. Additionally, it was shown that the phosphorylation state of Enzyme I from the PTS system was important for the regulation of this system (73) and that the phosphoenol pyruvate-dependent sugar PTS system protein, HPr, regulates activity of LsrK (72). These observations together with the fact that AI-2 synthesis is catalyzed by an enzyme involved in the central metabolism (discussed above) reveal a link between quorum sensing and metabolism that might be important for bacteria ability of rapidly respond to changes in nutrient levels (49,50,72).

The discovery of AI-2 internalization was surprising, at the time, because the majority of the quorum sensing signals would accumulate and remain in the extracellular environment (49,50,52,72). AHLs for example, due to their chemical properties can diffuse freely through the membranes and accumulate both in the inside and on the outside of the cell (74). So, interaction with either histidine kinase two-component systems or cytoplasmic transcriptional regulators presents no cost to the cells. In addition, LuxP also interacts with a histidine kinase two-component system. Thus, the question of why would the cell export AI-2 to later

import it again, both times with cost to the cell arose (49). Currently, it has been proposed that the internalization of AI-2 might enable differential regulation of behaviors through intra- and extracellular AI-2 (75). In fact, internal AI-2, through LsrK and LsrR, was shown to affect the expression of downstream genes that are involved in the regulation of aggregation, attachment and biofilm formation in *E. coli* whereas extracellular AI-2 has been shown to act as a chemoattractant necessary for aggregation and biofilm formation in enterohemorrhagic *E. coli*, in a process independent of AI-2 uptake into the cell (75-79). Additionally, it has been shown that AI-2 internalization can enable interference with AI-2 mediated quorum sensing from other bacteria (80). Accordingly, an Lsr detection and transporter system has been identified in *Sinorhizobium meliloti* although no LuxS homolog was identified nor AI-2 production was detected (65). As internalization of AI-2 provides no metabolic gain for *S. meliloti* (as it is also the case for *Salmonella* serovar Typhimurium (63)), it has been proposed that this bacterium senses AI-2 either to interfere with the quorum sensing of bacterial competitors or to distinguish between niches such as the soil where other bacterial species exist or nodules inside the host plant where *S. meliloti* is the only colonizer (65). Besides, the discovery of other quorum sensing systems that actively export and import the signals as in the case of linear autoinducing peptides in Gram-positive bacteria showed that this was not a one of a kind occurrence (28).

After the discovery of LsrB in *Salmonella* serovar Typhimurium and *E. coli*, other functional LsrB receptors were also identified in other Gram-negative Proteobacteria like *Photobacterium luminescens* (81), *Aggregatibacter actinomycetemcomitans* (82) and *S. meliloti* (65). The identification of LsrB receptors only in members of Proteobacteria could suggest that this system was present only in this phylum. However, in 2009, Pereira et al showed that *Bacillus cereus* and *Bacillus anthracis*, Gram-positive bacteria that belong to the Firmicutes, had functional Lsr systems, being able to produce and internalize AI-2 (17). A phylogenetic study proposed that events of lateral gene transfer were likely

responsible for the appearance of the *lsr* operon in Firmicutes (17). These findings were further supported by the elucidation of the crystal structure of the *B. anthracis* LsrB, which was solved in 2014 (PDB ID: 4PZ0). Moreover, other bacterial species, besides the Gram-positive *Bacillus*, were suggested to have these systems. For example, *Yersinia pestis* has a LuxS homolog and produces AI-2 (83) and Pereira et al proposed that it would have a functional LsrB receptor (17). Two years later, the crystal structure of the LsrB receptor from *Y. pestis* complexed with R-THMF was obtained (64). This reinforces the validity of the criteria utilized to identify functional LsrB receptors. LsrB homologs were proteins i) with amino acid sequence similarity higher than 60% with the LsrB from *Salmonella* serovar Typhimurium, ii) with the six conserved amino acid residues that hydrogen bound R-THMF and iii) encoded in an operon with homologs of the proteins involved in transport and regulation (17). With the rapid development in sequence techniques that has occurred in the last years, many other bacterial genomes became available. Probably a new screen using these criteria would identify additional functional LsrB receptors in other bacterial species, showing that the Lsr system is more widespread than what is thought.

In bacteria with LsrB systems, AI-2 has been shown to regulate biofilm formation, motility, virulence, aggregation and attachment (50,75-79).

Other proteins recognize AI-2

Proteins that are neither LuxP nor LsrB receptors have been proposed to interact with AI-2. Given the similarity of AI-2 with pentose sugars, the majority of these proteins are (or are annotated as) sugar-binding proteins.

The substrate binding protein of the ribose transporter, RbsB has been shown to interact with AI-2 in *A. actinomycetemcomitans*. However, this organism also has an LsrB receptor as mentioned before. Thus, both proteins seem to be involved in AI-2 recognition and accordingly, they compete for AI-2 binding *in vitro* (82). Determination of the half maximal inhibitory concentration (IC₅₀) by assessing

inhibition of luminescence produced by *V. harveyi* in response to a mix of LsrB or RbsB with AI-2 from *V. harveyi* or from *A. actinomycetemcomitans* showed a difference in affinities between the two proteins (82). Moreover, depletion of extracellular AI-2 was only sternly impaired in a $\Delta lsrB\Delta rbsB$ double mutant strain. The apparent redundant binding of LsrB and RbsB to AI-2 might explain the slower internalization rate of AI-2 that was previously observed in the absence of the LsrB receptor (62) or in the absence of each of the AI-2 binding proteins in *A. actinomycetemcomitans* (82).

In *Haemophilus influenza* strain 86-028NP, a protein annotated as RbsB, with homology for the LsrB receptor of *E. coli* and for the RbsB protein of *A. actinomycetemcomitans* was proposed to sense AI-2. *rbsB* mutants had impaired biofilm formation that could be complemented by addition of synthetic AI-2. Also, AI-2 uptake was decreased in this mutant, AI-2 internalization could not be competitively inhibited by ribose or other pentose sugars and AI-2 was shown to regulate AI-2 internalization as in Lsr systems (84,85). Therefore, it is possible that this protein was missannotated as a ribose-binding protein.

The D-galactose binding protein of *Fusobacterium nucleatum* has also been proposed to be an AI-2 receptor. The authors propose that D-galactose functions as an AI-2 inhibitor. In the presence of D-galactose, biofilm formation, which is induced by AI-2 in this organism, was inhibited. Also, AI-2 activity was repressed in the presence of D-galactose (86). Nevertheless, the observation that it also inhibits AI-2 activity in *V. harveyi* that has a LuxP receptor and that this repression could also be induced by D-ribose makes the existence of further studies critical.

The fructose-specific phosphotransferase enzyme IIB, FruA, has also been proposed as an AI-2 detection system in *Streptococcus pneumoniae*. This is based on the observation that a *fruA* mutant is unable to respond to exogenous AI-2 and that expression levels of *fruA* were altered in wild type and *fruA* mutant strains upon addition of AI-2 (87). While *fruA* might belong to an AI-2 detection system, there is also the possibility that this PTS system is regulating the expression of an

Lsr system. In *E. coli*, it has been shown that the PTS system regulates Lsr, linking quorum sensing to nutrient availability. The phosphorylation state of the subunit IIA^{Glc} regulates the activity of adenylate cyclase, which synthesizes cyclic adenosine monophosphate (cAMP) that in complex with its receptor (CRP) regulates transcription of the Lsr promoter (52,66,69,88). Furthermore, phosphorylated EI is required for initiating AI-2 internalization and the activity of LsrK is regulated by the PTS protein HPr (72,73). Thus, the hypothesis that the *fruA* mutant is unable to respond to synthetic AI-2 due to repression of the Lsr system should be considered. Moreover, the authors highlight that no Lsr systems have been identified in Gram-positive bacteria to reinforce their hypothesis. However, that is not the case, as identification of Lsr systems in *B. cereus* and *B. anthracis* precedes this work. Also, although a quick search does not return LsrB homologs in *S. pneumoniae* D39, the studied strain, in other strains of *S. pneumoniae*, *Salmonella* serovar Typhimurium LsrB homologs can be found with approximately 60% of homology and with the six conserved amino acids. They are, however, annotated as HTH-type transcriptional repressor PurR, a DNA binding domain. However, by submitting the amino acid sequence of these LsrB homologues to Phyre2, a protein fold recognition server, the first three results obtained are structures of LsrB receptors. These observations strongly suggest the existence of LsrB systems in at least some strains of *S. pneumoniae*.

Additionally, two substrate-binding proteins AibA and AibB (for AI-2 binding proteins A and B) were shown to independently bind AI-2 in *Helicobacter pylori* (89). The molecular details of binding have not yet been identified but these proteins were shown to interact with the TlpB chemoreceptor to induce biofilm dispersal (89,90). In enterohemorrhagic *E. coli*, a similar interaction was shown. LsrB binds AI-2 and interacts with the chemoreceptor Tsr to mediate cell aggregation (75,77). No significant similarity was found when comparing the amino acid sequence of these proteins with the one from the LsrB of *E. coli*. Thus,

if the functionality of these proteins is confirmed, a new class of receptors might be identified.

All together, these observations suggest that some bacterial species developed alternative mechanisms to detect AI-2. Furthermore, observation that these alternatives exist even in strains that possess LsrB receptors, might explain why strains mutated for LsrB can still internalize AI-2, albeit at a much slower rate. To distinguish between AI-2 receptors and proteins that are able to bind AI-2 but did not evolve to be AI-2 receptors, optimization of assays that allow determination of precise binding affinities like isothermal titration calorimetry and surface plasmon resonance might be key.

1.5.2. Quorum sensing networks

Two of the most studied network architectures exist in *P. aeruginosa* and in *Vibrio* spp. These bacterial species use different quorum sensing signals and response mechanisms and are able to integrate all the information from these different systems to accurately regulate quorum sensing behaviors.

P. aeruginosa has four quorum sensing pathways known to interact and regulate virulence as well as biofilm formation (Fig. 1.6). Two of them are LuxI/LuxR like pathways where the receptors are cytoplasmic transcription factors, which regulate many downstream genes, including the genes of the corresponding synthases creating a feed forward loop of autoinduction (3,91). These are the pairs LasI/LasR and RhII/RhlR. It also possesses a quinolone (2-heptyl-3-hydroxy-4-quinolone; Pseudomonas quinolone signal, PQS) and a non-ribosomal peptide (2-(2-hydroxyphenyl)-thiazole-4-carbaldehyde; integrated quorum sensing signal, IQS) mediated quorum sensing systems (3). These systems function in a hierarchical way. LasR bound to 3-oxo-C12-HSL activates the expression of *rhlR* and *rhlI* and the *pqsR* and *pqsABCDH* genes that encode the quinolone mediated system. The PqsR bound to PQS activates *rhlR* and *rhlI* connecting the three quorum sensing systems. In turn, RhlR inhibits the

expression of *pqsR* and *pqsABCDH* genes, presumably to maintain the ratio between 3-oxo-C12-HSL (AHL that binds LasR) and C4-HSL (AHL that binds RhlR) and thereby the correct activation of the PQS system (3,92). This tight control has been shown to regulate at least 13 Rhl and/or Las-directed functions and will fine-tune the response, enabling robust cell-cell communication under diverse conditions (92).

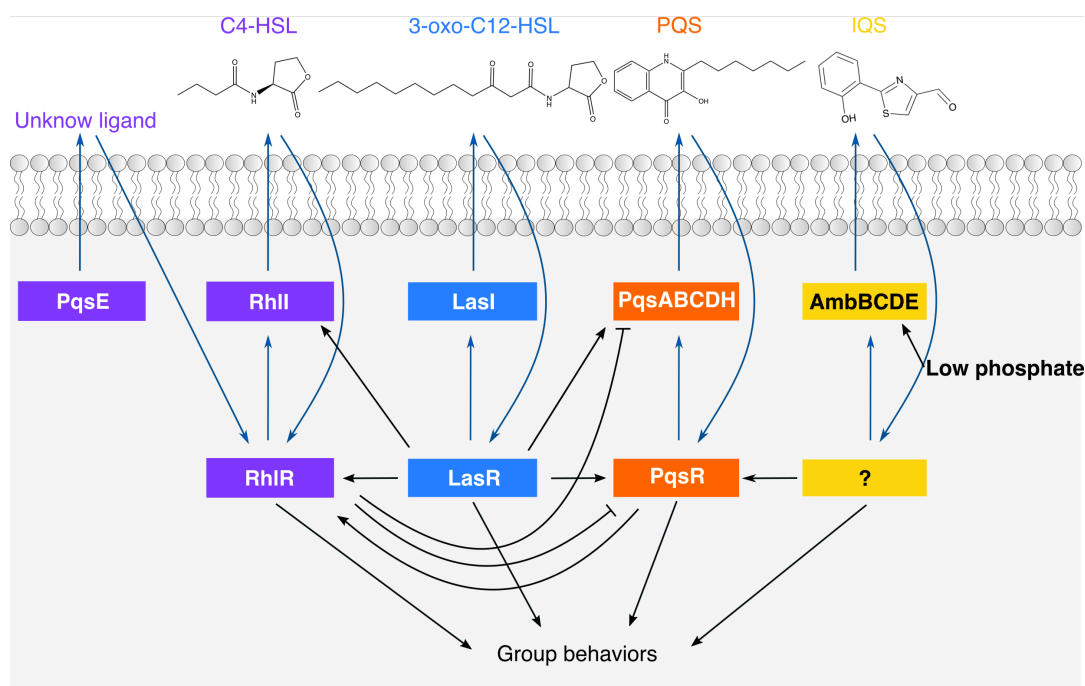


Figure 1.6 - Hierarchical quorum sensing pathways in *Pseudomonas aeruginosa*. Four quorum sensing pathways that work in a hierarchical order to regulate group behavior are highly interconnected in *P. aeruginosa*. RhlI, LasI, PqsABCDH and AmbBCDE are the enzymes that catalyze the synthesis of the signals C4-HSL, 3-oxo-C12-HSL, PQS and IQS, respectively. These signals interact with the cytoplasmic transcription factors RhlR, LasR, PqsR and an unknown receptor of IQS. The synthesis of IQS is activated in low concentrations of phosphate. Moreover, PqsE catalyzes the synthesis of an autoinducer of unknown structure that is able to bind RhlR. Adapted from (3,19).

The IQS mediated system can substitute the control exerted by LasR when in low phosphate concentrations (93). When the concentrations of phosphate are low, the phosphate starvation protein activates the *ambBCDE* genes responsible for synthesizing IQS, which will in turn activate the expression of *pqs* genes (93).

Under these conditions, the activation of the quinolone mediated system will provide the extra regulation layer needed for the production of virulence factors regulated by RhlR, usually mediated by the *las* system (3). The discovery of this fourth quorum sensing system has answered the question of how LasR mutants, frequent in contexts of *P. aeruginosa* chronic infection (94), can still regulate the production of virulence factors (3). Recently, PqsE has been identified as a quorum sensing synthase, which is also encoded in the *pqs* operon and that produces an alternative autoinducer, of unknown structure, that binds RhlR in the absence of C4-HSL. Moreover, it was described that PqsE is required and RhlI is dispensable for RhlR-directed virulence (95). This discovery unveils a new layer of complexity of the cell-cell communication network from *P. aeruginosa*.

V. cholerae and *V. harveyi* also have multiple quorum sensing systems. In this case, two are conserved in both species and work in parallel (Fig. 1.7). Both *Vibrios* recognize AI-2 through LuxPQ and CAI-1 ((S)-3-hydroxytridecan-4-one, produced by CqsA) through CqsS. At low cell density, the membrane bound receptors CqsS and LuxQ catalyze the transfer of a phosphate to the phosphotransfer protein LuxU, which further transfers it to the response regulator LuxO (96). Phosphorylated LuxO together with the alternative sigma factor σ_{54} (97), activates the expression of genes encoding four (in *V. cholerae*) or five (in *V. harveyi*) homologous regulatory small RNAs (3). These are named Qrr and are dependent on Hfq, a RNA-binding protein that facilitates the base pairing between small RNAs with their mRNAs targets (98). Qrr will activate translation of the low cell density master regulator AphA and destabilize the mRNAs encoding the master high cell density regulators HapR (in *V. cholerae*) or LuxR (in *V. harveyi*) (3,98). Thus, at low cell density, AphA, which activates expression of genes required for biofilm formation and pathogenicity, is produced while HapR/LuxR are not.

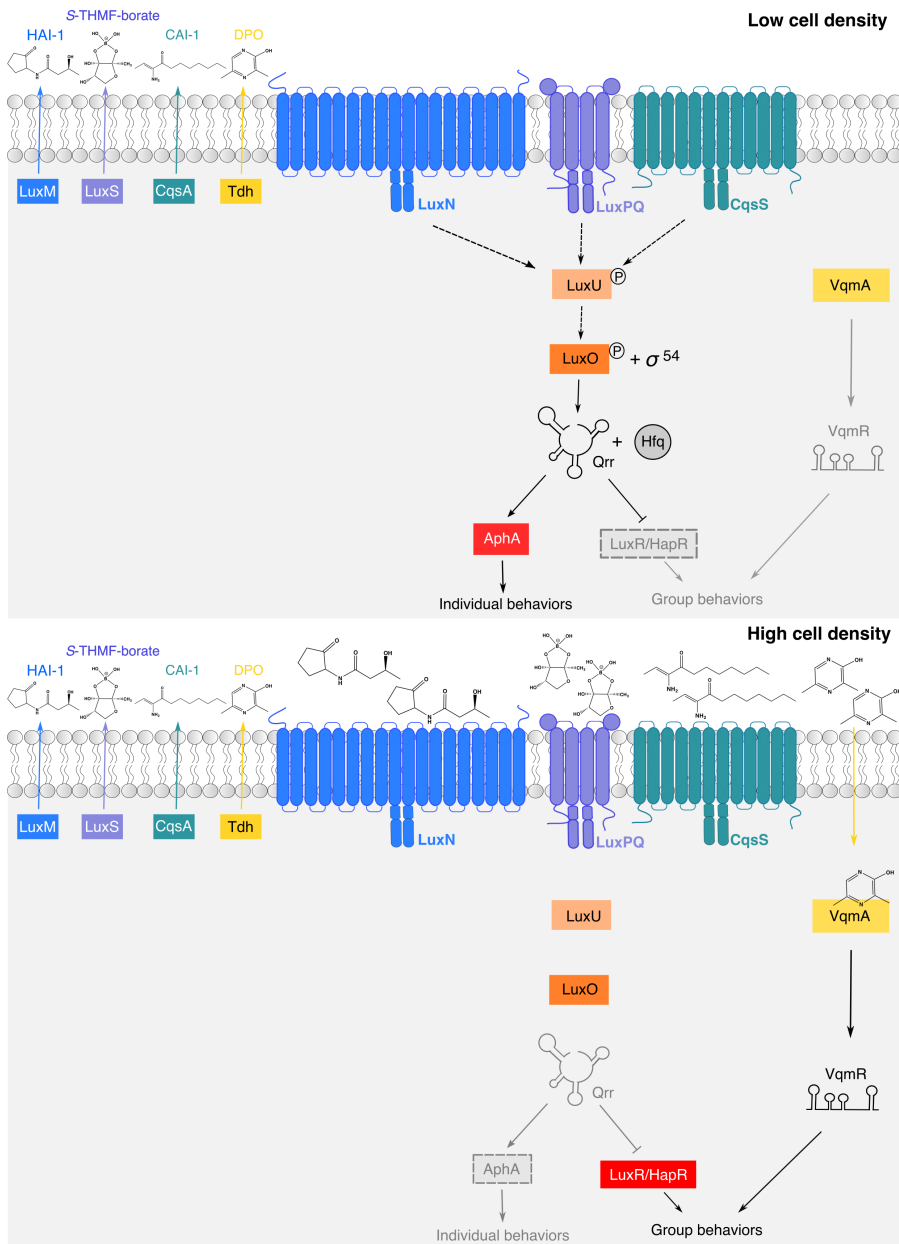


Figure 1.7 - Parallel quorum sensing pathways in *Vibrio* spp. At low cell density (top panel) when the concentrations of autoinducers are low, their cognate receptors will induce a phosphorylation cascade that results in the activation of AphA, a transcription factor that regulates individual behaviors. At high cell density (bottom panel), high concentrations of HAI-1, AI-2, CAI-1 and DPO will be present due to the action of the synthases LuxM, LuxS, CqsA and Tdh, respectively. Upon binding to the corresponding receptors LuxN, LuxPQ and CqsS, a shift from kinase to phosphatase activity will occur leading to the inactivation of LuxO and ultimately to the activation of LuxR (*V. harveyi*) or HapR (*V. cholerae*), which regulate group behaviors. Binding of DPO to VqmA also contributes to the regulation of group behaviors. Adapted from (3,19).

In turn, when the concentration of autoinducers increase, at high cell density, their binding to the receptors switches their activity from kinases to phosphatases, resulting in dephosphorylation and inactivation of LuxO, and termination of Qrr expression, thereby inactivating AphA production. HapR and LuxR will be produced and will activate genes involved in group behaviors (98). Namely their activation will lead to activation of bioluminescence and inhibition of biofilm formation (3). Moreover, two additional interacting quorum sensing pathways exist, one in each of these organisms. In *V. cholerae*, 3,5-dimethylpyrazin-2-ol (DPO), a newly discovered autoinducing molecule made from threonine and alanine, in a reaction catalyzed by threonine dehydrogenase (Tdh), will bind to the cytoplasmic transcription factor VqmA (99). This transcription factor will, in turn, activate expression of *vqmR*, which encodes the VqmR regulatory small RNA. VqmR controls a downstream regulon of 17 genes. One of these genes is *vpsT* that is repressed by both VqmR and HapR and that is required for biofilm formation (99). In *V. harveyi*, 3OHC4-homoserine lactone (HAI-1) produced by LuxM, an AHL synthase with no homology to the LuxI-type of enzymes, also binds a histidine kinase two-component system and interacts with LuxU like CqsS and LuxPQ (12,24). Thus, all of the known pathways work in parallel to control quorum sensing regulated behaviors. This arrangement, which requires detection of multiple signals, might be important to avoid noise from other molecules in the environment due to coincidence detection of the autoinducer molecules. Moreover, five (in *V. cholerae*) and six (in *V. harveyi*) feedback loops ensure optimal dynamics, fidelity and smooth transitions between quorum sensing states (3).

1.5.3. Listening in on the neighbors

Up to now, many examples of production and detection of quorum sensing molecules by different bacterial species have been described. However, what happens when a bacterial species is able to detect a signal that it does not

produce? Or if it can promiscuously detect other signals? When considering multi-species environment, the ability to detect molecules produced by neighboring bacteria might convey important information for cooperative or competitive behaviors. Here, some examples of “eavesdropping” will be discussed.

LuxR solo receptors are LuxR receptors that do not possess a cognate LuxI synthase and they comprise approximately 76% of the annotated LuxR proteins (3). The large number of solo LuxR receptors might indicate that unknown autoinducers or autoinducers produced by other bacteria might modulate their activity. In fact, this class of receptors has been associated with intra and interspecies communication as well as interkingdom communication (100). The LuxR solo receptor QscR from *P. aeruginosa* is known to respond to nanomolar concentrations of C8-HSL, C10-HSL, 3-oxo-C10-HSL, C12-HSL, 3-oxo-C12-HSL and C14-HSL and consequently to endogenously produced AHLs and those produced by cohabitating species like *Burkholderia cepacia* (101,102). Accordingly, QscR has lower substrate specificity than LasR and RhIR (101). SidA from *E. coli* and *Salmonella* serovar Typhimurium are known to respond to exogenously produced AHLs (100). More recently, LuxR solo receptors that do not bind AHLs but plant-derived molecules have also been identified in pathogenic and beneficial plant-associated bacteria (100).

In *Pectobacterium wasabiae*, a plant pathogen, the quorum sensing receptor ExpR2 responds to its cognate AHL but also to AHLs produced by other bacterial species. This allows *P. wasabiae* to eavesdrop on neighboring bacteria and join their efforts on exploitation of the host resources (103).

In *P. aeruginosa*, pathogenicity is increased both *in vitro* and *in vivo* in the presence of AI-2, even though it does not produce AI-2 (104). Similarly, *S. meliloti* has an AI-2 inducible Lsr-like system able to recognize and internalize AI-2 but does not have the LuxS synthase and thereby does not produce AI-2 (65).

Enteropathogenic *E. coli* was shown to have its virulence upregulated in the presence of synthetic or *V. cholerae* produced CAI-1 (105). However, this was not

observed for closely related *E. coli* strains or in *Salmonella* serovar Typhimurium suggesting that this detection system has evolved in enteropathogenic *E. coli* due to its common niche with *V. cholerae*. The ability of sensing CAI-1 might improve the fitness of enteropathogenic *E. coli* as it allows the timing of the infection to increase the chances of successfully colonizing the host (105).

Additionally, autoinducers involved in intraspecies communication might be the activating signal in one strain and inhibit the quorum sensing system of a closely related strain. For example, four different strains of *S. aureus* produce specific autoinducer peptides and so, each peptide activates its cognate AgrC receptor but inhibits activation of the others by competitive binding to the non-cognate receptors (2). This leads to intraspecies competition and, as each *S. aureus* strain is the primary causative agent of a specific type of *S. aureus* disease, seems to be fundamental to establish a specific niche for each strain (2). Also, in *V. harveyi*, the inhibition of LuxN receptor by longer AHLs suggests that this bacterial species shuts down quorum sensing in response to autoinducers produced by competitor bacterial species to evade utilization of its public goods (3).

1.6. Cell-cell communication in the gut microbiota

Eukaryotes shelter a large number of bacterial communities that comprise the microbiota. Actually, the ratio of human to bacterial cells is estimated to be close to 1:1 (106). These communities exist mainly in the mammalian skin, oral cavity and the gut (19). The mammalian gut microbiota has generated ample interest due to its beneficial role on the maturation of the host's immune system, modulation of host metabolism, development and physiology (107). Moreover, there is growing evidence that gut microbiota dysbiosis is associated with the appearance of human disease such as obesity, diabetes and colorectal cancer (107). In this section we will approach communication mechanisms that shape the gut microbiota.

1.6.1. AI-2 alters the composition of the microbiota in antibiotic-induced dysbiotic gut

AI-2 is involved in interspecies communication being recognized by different species of bacteria. Thus, it is not surprising that AI-2 communication occurs in a multispecies environment like the gut microbiota. It has been shown that the increased production of AI-2 in the gut microbiota of mice promoted the expansion of Firmicutes over Bacteroidetes after antibiotic-induced dysbiosis (21). Firmicutes and Bacteroidetes are the two major phyla of bacteria in the gut and exist in a ratio of approximately 1:1. However, after treatment with the antibiotic streptomycin, nearly all Firmicutes are eliminated and Bacteroidetes expand (21). By adding an *E. coli* engineered to produce high levels of AI-2, a significant increase in the abundance of Firmicutes was observed. Notably, approximately 80% of the Firmicutes have homologs for the AI-2 synthase, LuxS, whereas only 20% of the Bacteroidetes have homologs for LuxS. This suggests that AI-2 is able to promote the colonization of AI-2 producing species. Moreover, this suggests a role for AI-2 signaling in accelerating the recovery of gut microbiota after antibiotic-induced dysbiosis (21).

1.6.2. Interkingdom communication

Several examples of interkingdom communication are known. Either with plants, fungi or even with the human host, bacteria are able to sense the surrounding environment to maintain successful mutualistic or host-pathogen relationships. In the gut microbiota there are several examples of host-bacteria interactions and an intersection with quorum sensing and nutritional cues seems to be the recurrent trend in interkingdom signaling (108). In this section a brief overview of these processes will be presented.

1.6.2.1. Autoinducer-3, host hormones and fucose

Autoinducer-3 (AI-3) is an aminated aromatic compound whose structure has not been identified (109). Several members of the human gut and pathogens produce AI-3, which was first identified in enterohemorrhagic *E. coli* (EHEC) where it was shown to induce virulence (110). The remarkable aspect of this quorum sensing system is that the host-produced hormones epinephrine and norepinephrine bind the same receptor as AI-3 (110). The co-evolution of bacteria and host for billions of years have exposed bacteria to host hormones used for communication in eukaryotic cells (111). This might explain the presence of bacterial systems able to recognize host hormones. Induction of this system by AI-3 or the hormones (the signals seem to be interchangeable) is needed for *in vivo* induction of virulence in EHEC as it was demonstrated in rabbit and bovine infection models (112,113). Thus, it is thought that this quorum sensing system is involved in interkingdom communication enabling EHEC to recognize that it has reached the intestine and consequently, induce virulence (108). Similar to other quorum sensing signals, AI-3 interacts with a histidine kinase two-component system (Fig. 1.8). AI-3, epinephrine and norepinephrine bind QseC, the membrane-bound receptor, promoting its autophosphorylation. QseC will then phosphorylate its cognate response regulator QseB as well as two others: KdpE and QseF. This signaling cascade will activate the locus of enterocyte effacement genes (pathogenicity island that encodes the type III secretion system, associated chaperones and effector proteins responsible for attaching and effacing lesions in the large intestine), *flhDC* motility genes, Shiga toxin production, stress responses, iron uptake amongst others (109,114).

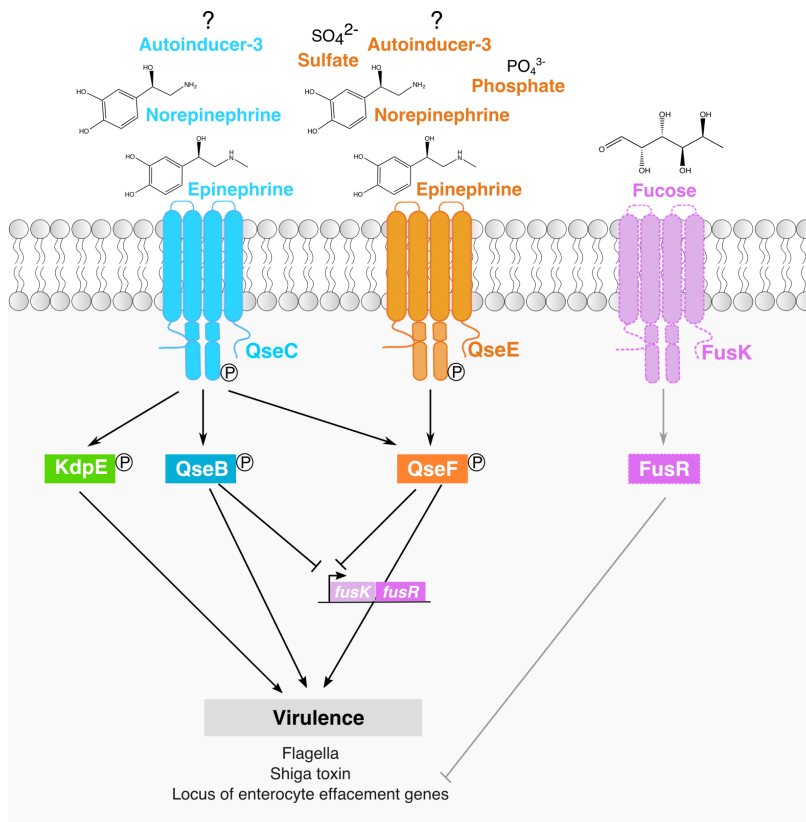


Figure 1.8 - Autoinducer-3, epinephrine and norepinephrine signaling. The host-produced hormones epinephrine and norepinephrine can bind the histidine kinase two component systems QseC and QseE. These receptors also bind autoinducer-3, whose structure is unknown, and QseE can also be activated by sources of phosphate and sulfate. The bindings of these molecules to the cognate receptors leads to the induction of virulence through the regulators KdpE, QseB and QseE. Moreover, QseB and QseF can also repress the expression of *fusR/fusK*, the histidine kinase two component system that upon binding of fucose, inhibits the expression of the genes in the locus of enterocyte effacement genes.

QseF is a transcription factor that belongs to a histidine kinase two-component system and can be phosphorylated by its cognate receptor QseE. QseE senses AI-3, epinephrine, norepinephrine and also sources of phosphate and sulfate (Fig. 1.8). Moreover, QseC and QseF can inhibit expression of *fusK/fusR*, a two-component system activated by fucose that represses expression of the locus of enterocyte effacement (109,115). The modulation of the FusK/FusR pathway might provide colonization advantage to EHEC. Commensal *E. coli* is the main competitor of EHEC in the lumen and it favors the use of fucose as carbon source.

By repressing the use of fucose, EHEC will rewire its metabolism to favor the use of D-galactose that is also highly prevalent. Since EHEC is more efficient in consuming D-galactose than the commensal *E. coli* it will gain a metabolic advantage (76,109).

The observations that α -adrenergic antagonists can bind QseC and inhibit the quorum sensing system, and that epinephrine and norepinephrine are able to bind QseC, suggests that AI-3 might have a conformation similar to these hormones (108). Recently, it was proposed that another host hormone, the mammalian stress hormone, dynorphin, released from the mice intestinal mucosa, activated quorum sensing in *P. aeruginosa* enhancing its virulence (111,116).

The prolonged co-evolution of bacteria and host has also exposed gut epithelial cells to quorum sensing molecules (111). So, host receptors might be able to recognize quorum sensing signals. Specifically, it has been proposed that AI-3 might be able to bind mammal α -adrenergic receptors and contribute for the maintenance of intestinal homeostasis. This hypothesis arose from the observations that mammals do not possess histidine kinase two-component systems, α -adrenergic agonists have been used to treat chronic diarrhea and AI-3 is produced not only by pathogens but also by members of the gut microbiota (108).

1.6.2.2. AI-2 mimic produced by the host

Another example of interkingdom communication in the gut microbiota is the production of an AI-2 mimic by epithelial cells in response to interaction with bacteria (117). It is recognized by both LuxP and LsrB AI-2 receptors and induces quorum sensing-dependent regulons in bacteria like *Salmonella* serovar Typhimurium and *V. cholerae*. This mimic is induced when epithelia are directly or indirectly exposed to bacteria suggesting that a secreted bacterial component stimulates its production. Still, it is not expressed in hematopoietic cells present in peripheral blood (117). This suggests that the mimic can have a role in host-

microbiota symbiosis as epithelial cells directly interact with colonizing bacteria. Also, the AI-2 mimic production was also increased following epithelial tight junction damage. This might indicate that the mimic might have an epithelial repairing role. The structure of this signaling molecule has not yet been identified. However, given its similarity to AI-2, involved in interspecies communication, it is tempting to predict that this mimic enables interaction between the host and a large range of bacterial species present in the gut (19,108,117,118).

1.6.2.3. Host immune response to quorum sensing signals

The gut houses the majority of the immune cells in the body and thus, it is expectable that bacteria elicit immune responses in the gut (111). Nevertheless, besides eliciting inflammatory responses that enable the gut to eliminate undesirable bacteria, autoinducers can also modulate the immune behavior of the gut. Injected AHLs induced inflammation by stimulating the expression of many cytokines such as interleukin 1 α (IL-1 α) and IL-6 in mice (119). AI-2 induced IL-8 expression in HCT-8 colon cancer cells (120). 3O-C12-HSL has been reported to enter mammalian immune cells and decreased the ability of dendritic cells derived from bone marrow to induce T-cell proliferation activation *in vitro* (121). Additionally, this AHL was shown to attenuate innate immune responses via disruption of NF- κ B, which potentially enables persistence of infection in the host (122). All the host receptors for these autoinducer molecules have not been identified yet, but it has been shown that a nuclear hormone receptor, the peroxisome proliferators-activated receptor, a transcription factor that modulates expression of proinflammatory genes, recognizes 3O-C12-HSL (123). However, this detection system does not fully account for all the biological effects that 3O-C12-HSL elicits in mammalian cells suggesting that other receptors exist (108). Moreover, autoinducing peptides can activate an intracellular signaling cascade that leads to tumor metastasis through the epidermal growth factor receptor (124).

In turn, host cells also developed quorum sensing-interfering mechanisms

like paraoxonase enzymes production, which act as lactonases and degrade AHLs within mammalian cells. Also, changes in the host conditions like pH, bile acid, temperature or osmotic pressure can affect autoinducers activity and consequently modulate the density and structure of gut microbiota affecting the health of the host in return (108,111,125).

1.7. Behaviors regulated by quorum sensing and heterogeneity

Quorum sensing has been shown to regulate behaviors like competence, motility, biofilm formation, virulence, bioluminescence and production of public goods like proteases or antibiotics (2,24). However, the more traditionally definition that quorum sensing promotes the synchronous expression of target genes across a bacterial population has been broaden due to new observations that include responses through intracellular chemical signals, interkingdom communication and heterogeneity in gene expression, all controlled by quorum sensing (3). The DSF-mediated communication approached in section 1.4 is a good example of quorum sensing controlling responses through intracellular chemical signals. Binding of DSF to its cognate receptor leads to a decrease in c-di-GMP levels, which will affect motility, biofilm formation and virulence (29). Linking quorum sensing to nucleotide-based second messengers allows the conversion of complex information encoded in autoinducer merged into a single, general intracellular signaling molecule (3). The detection of epinephrine and norepinephrine produced by the host by enteric bacteria and of an AI-2 mimic produced by mammalian epithelial cells in response to bacteria or disruption of tight junctions (explored in section 1.6) denotes modulation of the gut microbiome in result of interkingdom autoinducer signaling (117).

Also, heterogeneity in quorum sensing, where a subpopulation of cells can have quorum sensing activated and the remaining population has not, has been observed. The majority of the molecular mechanisms that cause heterogeneity are

not defined, but there are some factors that might instigate it like flow, surface topography, appearance of quorum sensing mutants and bet-hedging strategies for example (19). Under flow, the biomass required to initiate quorum sensing in a particular bacterial population increases with the increase in fluid flow rate. This is easily explained by the autoinducer removal induced by flow (126). Following the same logic, bacteria that inhabit secluded areas, like bacteria that colonize the intestinal crypts where little or no flow exists will turn quorum sensing on faster than the cells above the crypts that have their autoinducers washed away (126). Similarly, in thick biofilms, the protected cells inside the biofilm, will reach the concentration of autoinducers needed to induce quorum sensing faster than outer bacterial cells that have their autoinducers washed away (126). Also, appearance of quorum sensing mutants can promote quorum sensing heterogeneity as a part of the population will have the quorum sensing system on and the other will not (19,127). When the mutations occur in genes that lead to the production of public goods, i.e. secreted (thereby accessible to all) biomolecules that can capture nutrients from the environment, hydrolyze solid food sources and construct biofilms, cheaters will arise. Cheaters are able to use public goods produced by wild type cells but do not have the fitness cost to produce them and thus, overall, will have a fitness advantage over the wild type population being able to outcompete them (19). Since this conflict reduces population fitness, several processes to prevent cheating exist. These include spatial structure and quorum sensing mediated co-regulation of two enzymes: one important for the production of a public good and the other for the production of a private good (19,128). Additionally, quorum sensing regulation itself has been shown to contribute to prevent fixation of cheaters in *P. aeruginosa* (129).

Heterogeneity has been linked with low cell density as in the presence of fewer cells, higher noise is more likely to occur. An example would be distinctive expression of autoinducer synthases that cause phenotypic heterogeneity (19,130). Nevertheless, maintaining phenotypic heterogeneity in high cell densities could

allow bacteria to increase fitness under temporally varying conditions. Thus, some cells would engage in group activities while others perform individual behaviors, decreasing the overall risk of death of the bacterial population (bet-hedging) (19,131).

1.8. Final remarks and main goals

The acceptance of bacteria as social microbes was slow, but an explosion of discoveries on quorum sensing systems and an attempt to interpret them in light of what exists in nature has occurred. We know now that multiple systems that can interact exist and that extracellular factors like flow or structure can influence the efficiency of these systems. All this information and the increase in bacterial resistance to known antibiotics make the possibility of interfering with quorum sensing a very attractive strategy. AI-2 mediated quorum sensing is a particularly appealing system, as it can be used to target many different species and is involved in interspecies communication in complex biological niches like the mammalian gut microbiota. However, to manipulate this communication process we need to understand the molecular mechanisms that underlie these interactions. And the first step to understand them is to identify the AI-2 receptors in the species involved.

Hence, in this work our main goal was to identify novel AI-2 receptors. The advances in DNA sequencing allowed us to explore the large number of bacterial genomes available and, consequently, identify LsrB-like receptors with an alternate binding pocket. Also, we developed a new chemical probe that has the potential to enable the discovery of novel AI-2 receptors in organisms that have AI-2 regulated phenotypes and unknown receptors.

Finally, we were also interested in the ability of manipulating the communication mechanism mediated by AI-2. To achieve this goal, we studied the biological activity of AI-2 analogs with agonist activity for LuxP and developed a method to quantify AI-2 in the gut of mice.

Chapter



Identification of novel LsrB-like receptors in Clostridia

Unpublished data and data published in:

Torcato, I. M., Kasal, M. R., Brito, P. H., Miller, S. T. and Xavier, K. B. (2019) Identification of novel autoinducer-2 receptors in Clostridia reveals plasticity in the binding site of the LsrB receptor family. *J. Biol. Chem.* **294**(12) 4450-4463. DOI: 10.1074/jbc.RA118.006938.

Author contributions:

Inês M. Torcato, Karina B. Xavier and Stephen T. Miller conceptualized the study. Stephen T. Miller performed the crystal cryopreservation and obtained and analyzed the X-ray data. Patrícia H. Brito performed and analyzed the phylogenetic study. Meghann R. Kasal purified the proteins used for ITC experiments and with Stephen T. Miller performed the AI-2 *in vitro* binding assay of the LsrB-like proteins of *Clostridium saccharobutylicum* and *Clostridium autoethanogenum*. Inês M. Torcato performed the remaining experiments.

2.1. Abstract

Autoinducer-2 (AI-2) is unique among quorum sensing signaling molecules, as it is produced and recognized by a wide variety of bacteria and thus facilitates interspecies communication. To date, two classes of AI-2 receptors have been identified: the LuxP-type, present in the Vibrionales, and the LsrB-type, found in a number of phylogenetically distinct bacterial families. Recently, AI-2 was shown to affect the colonization levels of a variety of bacteria in the microbiome of the mouse gut, including members of the genus *Clostridium*, but no AI-2 receptor had been identified in this genus. Here, we identify a noncanonical, functional LsrB-type receptor in *Clostridium saccharobutylicum*. This novel LsrB-like receptor is the first one reported with variations in the binding site amino acid residues that interact with AI-2. The crystal structure of the *C. saccharobutylicum* receptor, determined at 1.35 Å resolution, revealed that it binds the same form of AI-2 as the other known LsrB-type receptors, and isothermal titration calorimetry (ITC) assays showed that binding of AI-2 occurs at a submicromolar concentration. Using phylogenetic analysis, we inferred that the newly identified noncanonical LsrB receptor shares a common ancestor with known LsrB receptors and that noncanonical receptors are present in bacteria from different phyla. This led us to identify putative AI-2 receptors in bacterial species in which no receptors were known, as in bacteria belonging to the Spirochaetes and Actinobacteria phyla. Thus, this work represents a significant step towards understanding how AI-2-mediated quorum sensing influences bacterial interactions in complex biological niches.

2.2. Introduction

Bacteria are able to colonize and adapt to different environments, and the process of quorum sensing greatly enhances this ability. Quorum sensing is a mechanism of cell-to-cell communication mediated by exchange of small chemical

molecules, named autoinducers, that allows bacteria to monitor their population density and regulate gene expression accordingly (2,8,9). Behaviors regulated by this process include biofilm formation, virulence factor expression, antibiotic production, and bioluminescence (3,50). Autoinducer-2 (AI-2) is a signaling molecule produced and recognized by a wide variety of bacteria and thus facilitates interspecies communication (46,47,50). Since bacteria often produce species-specific signals as well, the relative proportion of AI-2 depends on the identity and number of bacterial species present in the environment; thus AI-2-sensitive bacteria can, in principle, regulate their behavior according to the species composition of the community (80).

To date, all characterized AI-2 receptors have high sequence similarity to, and complete binding site identity with, one of two receptors: LuxP or LsrB (46,47,50). Both receptors belong to the high affinity substrate-binding protein family. Importantly, the two receptors bind chemically distinct derivatives of the AI-2 precursor molecule DPD ((S)-4,5-dihydroxypentane-2,3-dione) (46,47). Whereas LuxP binds S-THMF-borate [(2S,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran-borate], LsrB binds R-THMF [(2R,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran], with discrimination between these AI-2 adducts being presumably due to differences in the amino acid composition of the binding sites of these receptors (46). Regulatory mechanisms also differ by receptor type. In response to AI-2 binding, LuxP modulates the activity of a membrane-spanning sensor histidine kinase protein, thus regulating a phosphorylation signal transduction cascade, whereas LsrB is part of an ATP binding cassette (ABC) transport system that internalizes AI-2 (12,63,132). LsrB was first identified in *Salmonella enterica* serovar Typhimurium but, unlike LuxP, which has only been found in the Vibrionales, it is also present in other enteric bacteria and members of the *Rhizobiaceae* and *Bacillaceae* families (46,50,63).

The LsrB-mediated transport system is comprised of two transmembrane proteins, LsrC and LsrD, and an ATPase, LsrA, that are encoded in the *lsr* (for

LuxS regulated) operon together with *LsrB*. As the density of AI-2-producing bacteria increases, AI-2 accumulates extracellularly. When the threshold concentration is reached, *LsrB*-bound AI-2 is internalized via this transport system (52,63). Internalized AI-2 is subsequently phosphorylated by the kinase *LsrK* (62,66,67), and phosphorylated AI-2 relieves the repression caused by the *lsr* regulator, *LsrR* (68,69). This derepression leads to *lsrABCDGF* expression and thus rapid depletion of AI-2 from the extracellular medium. In most bacteria that possess the *lsr* operon, the enzymes responsible for further AI-2 processing, *LsrG* and *LsrF*, are also encoded in the *lsr* operon (see Fig. 2.1 for an explanatory scheme) (70,71).

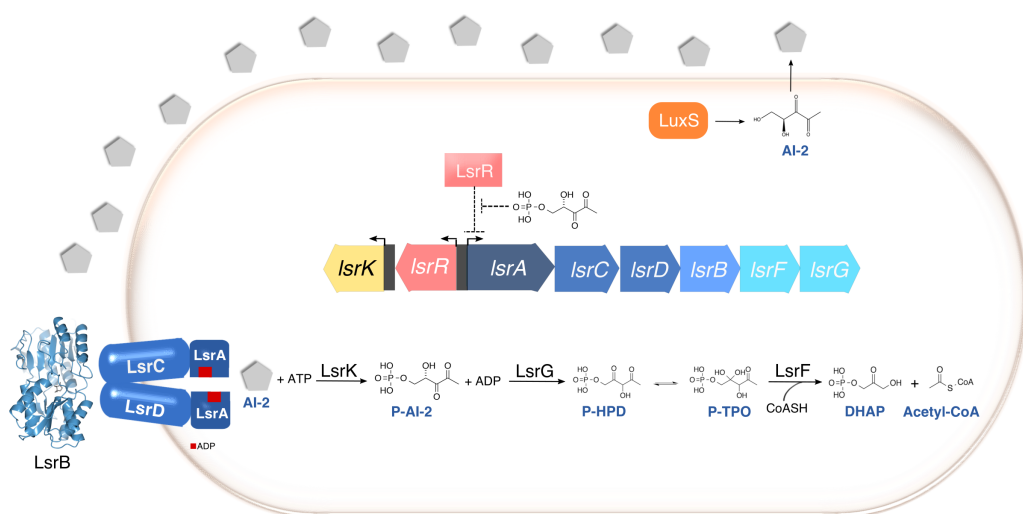


Figure 2.1 - Schematic representation of the AI-2 internalization process. *LuxS* produces AI-2 as (4S)-4,5-dihydroxy-2,3-pentanedione. After synthesis, AI-2 is exported to the extracellular milieu and accumulates in proportion to bacterial density. Once AI-2 accumulates to a threshold concentration, it is internalized and phosphorylated by *LsrK*. Phosphorylated AI-2 (P-AI-2) can then bind the transcriptional regulator *LsrR*, blocking repression and leading to the expression of the genes in the *lsr* operon. Subsequent expression of the *LsrB* receptor and of the associated ABC transporter (*lsrACD*) promotes AI-2 internalization and consequent depletion of extracellular AI-2. In addition, phosphorylation by *LsrK* sequesters the signal inside the cell. In *E. coli*, P-AI-2 is further processed by *LsrG* which catalyzes the isomerization of P-AI-2 to 3-hydroxy-2,4-pentadione-5-phosphate (P-HPD), an isomer that exists in equilibrium with its hydrated form 3,4,4-trihydroxy-2-pentanone-5-phosphate (P-TPO). *LsrF* then catalyzes the transfer of an acetyl group from P-HPD to coenzyme A,

Chapter 2

creating dihydroxyacetone phosphate (DHAP) and acetyl-CoA (key metabolites used by the cell in metabolic pathways like glycolysis and the citric acid cycle).

It has been shown that Lsr-mediated AI-2 internalization influences the expression of downstream genes involved in regulating aggregation, attachment, and biofilm formation in *Escherichia coli* (76,78,79). Moreover, in pathogenic enterohemorrhagic *E. coli*, extracellular AI-2 acts as a chemoattractant; this activity, mediated by the LsrB receptor, is required for cell aggregation and biofilm formation (75-77,79,133). Thus, AI-2 internalization presumably regulates these behaviors via the different responses exerted by intra and extracellular AI-2.

In previous work, we identified LsrB receptors in several species (including *Bacillus anthracis* and *Bacillus cereus*, the first Firmicutes to be shown to have functional LsrB receptors) through an approach that used bioinformatics to identify candidate receptors and biochemical and genetic studies to confirm their function (17). Based on this work, we proposed the following criteria to identify functional LsrB receptors: i) more than 60% sequence identity with the LsrB from *Salmonella* serovar Typhimurium, ii) conservation of all six-binding site amino acid residues that interact with AI-2 and iii) co-occurrence with orthologs to the other key transport proteins encoded by the *lsr* operon. With the constant increase in the number of sequenced genomes available, new strains of bacteria can now be studied. It became clear in the course of this work that, while these criteria do identify functional receptors, the identification of a more diverse range of AI-2 receptors requires the expansion of these criteria.

AI-2 receptors are relevant in complex communities such as the gut microbiota, a clinically significant community where species imbalance has been linked to bowel disease, obesity and susceptibility to pathogen colonization. However, little is known about the interspecies communication mechanisms typically employed in this environment, partly due to the inability to identify the AI-2 receptors of certain members of the microbiota (21,134). Recently, it was shown that AI-2 can influence the proportions of the major phyla in antibiotic-

treated mice gut microbiota (21). Specifically, the Firmicutes phylum was shown to respond positively to this signal. Interestingly, the majority of Firmicutes, including many Clostridia, have orthologs of the AI-2 synthase (LuxS) and are therefore putative AI-2 producers (21); however, no AI-2 receptors were known or readily identified by prior criteria in the Clostridiales. These findings, together with increasing understanding of the importance of Clostridia members as both commensals and pathogens in the gut microbiota, led us to search for AI-2 receptors in this class (57,135,136).

In this work we identify new LsrB-type receptors in *Clostridium saccharobutylicum* and *Clostridium autoethanogenum*, species shown to exist in the microbiota of mammals (137,138). We characterize the *C. saccharobutylicum* receptor and show that this novel receptor binds R-THMF like previously characterized LsrBs, despite the fact that two of the amino acids involved in AI-2 binding differ from those of the canonical LsrBs (17). This variation demonstrates that the plasticity of the LsrB binding site is higher than previously thought and expands our understanding of the diversity among AI-2 binding receptors. To the best of our knowledge, this is the first report of the identification and characterization of AI-2 receptors in Clostridia. Moreover, this knowledge allowed us to identify new noncanonical putative AI-2 receptors in other bacterial species belonging to the Firmicutes phylum, including bacterial species isolated from human microbiota. We also identified putative noncanonical LsrB receptors in organisms belonging to other phyla where AI-2 receptors had not previously been identified and where AI-2 mediated quorum sensing can now be studied. The identification of AI-2 receptors in these species represents a key step in understanding the mechanisms by which AI-2 regulates the levels of Firmicutes in the mouse gut and, more broadly, how AI-2-mediated quorum sensing influences the behavior of this and other complex communities.

2.3. Results

2.3.1 Identification of LsrB orthologs and in vitro binding to AI-2

To identify novel LsrB receptors, we searched the complete genomes deposited in the Kyoto Encyclopedia of Genes and Genomes (KEGG) for LsrB orthologs. We were seeking to identify functional LsrB receptors in the Firmicutes phylum, but we only obtained hits in members of the *Bacillus* genus using the previously established criteria (17). However, we were intrigued by the identification of LsrB homologous proteins in the Clostridia class. These hits had sequence identities lower than 60%, the minimum observed in previously identified LsrB functional receptors, but still higher than 36%, the maximum observed for the homologs unable to bind AI-2. Proteins identified from *C. saccharobutylicum* and *C. autoethanogenum* shared a sequence identity with the LsrB from *B. cereus* of 39.2 and 37.9%, respectively (Table S2.2). Sequences for these proteins were submitted to the Phyre2 fold-recognition server and the top two hits for both were the LsrBs from *B. anthracis* and *Salmonella* serovar Typhimurium, suggesting that these proteins share the same overall tertiary structure as previously characterized LsrB receptors. We aligned the resulting predicted structures to the LsrB from *B. anthracis* (Protein Data Bank (PDB) ID: 4PZ0) and found that the putative binding site in the Clostridia orthologs varied from canonical LsrBs in two of the six AI-2 binding residues (P223N, A225S; numbering based on that of the PDB structure of CsLsrB). We also analyzed the genome context of these proteins and found the genome location of the genes encoding the putative LsrB receptors of *C. saccharobutylicum* and *C. autoethanogenum* to be identical to that of bacteria with functionally characterized LsrB receptors. The genes are located in a putative operon with homologs for an ABC transport system (Fig. 2.2, A; Table S2.2). These findings, combined with the biological relevance of the Clostridia class, led us to test if these proteins were functional LsrB-like receptors despite their lower sequence identity and the absence of a fully

conserved AI-2 binding site.

To determine the functionality of these proteins, *in vitro* AI-2 binding studies using *Vibrio harveyi* strain MM32 (a reporter strain for AI-2 activity) were conducted (45). Candidate receptors from *C. saccharobutylicum* and *C. autoethanogenum* were expressed in AI-2-producing *E. coli* strains, purified, and heat-denatured to release the bound ligand. The ligands released from the putative receptors induced light production in *V. harveyi* at levels similar to that of our positive control, ligand released from *B. anthracis* LsrB (Fig. 2, B).

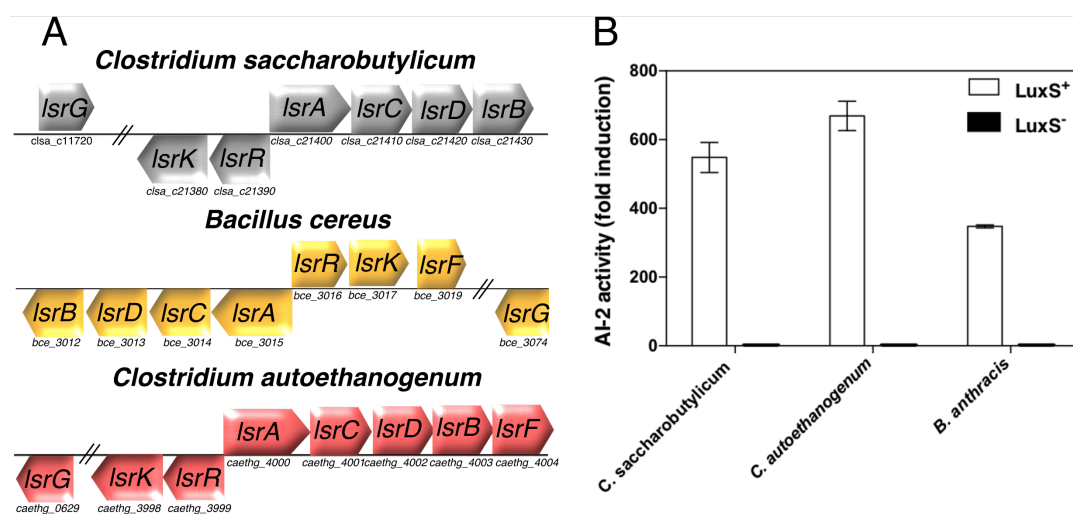


Figure 2.2 - LsrB-like receptors from *C. saccharobutylicum* and *C. autoethanogenum* bind AI-2. (A) Comparison of the putative AI-2 transport and processing orthologs in *C. saccharobutylicum* (top) and *C. autoethanogenum* (bottom) with *B. cereus* (middle), which has a functional Lsr transporter system. (B) AI-2 binding was assessed by measurement of light production of *V. harveyi* MM32 after addition of ligand released from the pure putative LsrBs from *C. saccharobutylicum* and *C. autoethanogenum* expressed in a LuxS⁺ *E. coli* BL21 strain (white bars). The LsrB receptor from *B. anthracis* was used as a positive control. As negative controls we tested the mentioned proteins expressed in a LuxS⁻ mutant (black bars). Results are shown as fold induction relative to the light production induced by the growth medium.

As a negative control, receptors were also expressed in a LuxS⁻ *E. coli* strain (i.e. non-AI-2-producing); as predicted these proteins elicited no response from *V. harveyi* showing that the response is AI-2-dependent (Fig 2.2, B). Given this demonstration of AI-2 binding ability by the candidate proteins, we concluded

that they are functional AI-2 receptors. We named these receptors CsLsrB for *C. saccharobutylicum* LsrB and CaLsrB for *C. autoethanogenum* LsrB.

2.3.2. AI-2 internalization in *C. saccharobutylicum*

As *C. saccharobutylicum* and *C. autoethanogenum* have orthologs for the Lsr transporter system, we tested whether accumulation and internalization of AI-2 was similar to previously studied bacteria with the *lsr* operon (17,52,62). Due to the similarity of the putative LsrBs identified in *C. saccharobutylicum* and *C. autoethanogenum*, we focused our studies on *C. saccharobutylicum* since the strain is more amenable to manipulation. Cell-free supernatants from *C. saccharobutylicum* cultures were collected at different time points post-inoculation and the activity of extracellular AI-2 was assayed via a *V. harveyi* MM32 luminescence assay (Fig. 2.3, A).

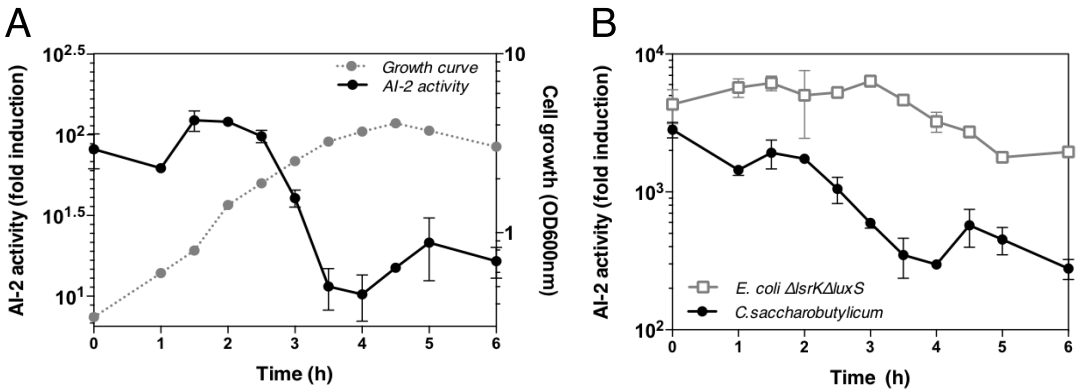


Figure 2.3 – *C. saccharobutylicum* LsrB receptor internalizes AI-2. (A) AI-2 internalization profile of *C. saccharobutylicum* (black circles) and respective growth curve (grey circles). (B) AI-2 internalization profile of *C. saccharobutylicum* (black circles) and *E. coli* Δ lsrK Δ luxS (empty squares) after addition of 40 μ M of AI-2 at time zero. Cell free supernatants were collected at the indicated time points and activity of extracellular AI-2 was measured by assessing light produced by *V. harveyi* MM32 in response to cell-free supernatants. AI-2 activity is reported as fold induction relative to light production induced by the growth medium. The internalization curves are representative of three independent experiments, and error bars represent the standard deviation of three technical replicates.

We observed that the levels of extracellular AI-2 increased until a certain

threshold level was reached, then rapidly decreased starting at mid-exponential growth. To further support the observation that *C. saccharobutylicum* is capable of internalizing AI-2, we added synthetic AI-2 to cultures of *C. saccharobutylicum* and of *E. coli* at time zero. We then compared the concentration of extracellular AI-2 present in *C. saccharobutylicum* supernatants to that in supernatants of the *E. coli* mutant that can neither produce or internalize AI-2 ($\Delta lsrK\Delta luxS$ mutant) (Fig. 2.3, B). As this *E. coli* mutant is not capable of internalizing AI-2, any decrease in AI-2 activity in the supernatants from these *E. coli* cultures would be due to degradation of AI-2 in these conditions. At 6h, supernatants of *C. saccharobutylicum* only stimulated approximately 280-fold induction, whereas supernatants of *E. coli* $\Delta lsrK\Delta luxS$ caused a 2000-fold induction, supporting the conclusion that *C. saccharobutylicum* depletes AI-2 from the culture by internalization. Moreover, AI-2 is not degraded extracellularly, as no significant degradation was observed in the course of 6h incubation of 40 μ M of AI-2 in cell-free supernatants (Fig. S2.1, A) collected from a late stationary phase culture of *C. saccharobutylicum* (culture shown in Fig. S2.1, B). Thus, we concluded that *C. saccharobutylicum* is able to internalize AI-2 as previously observed for other LsrB receptors.

2.3.3. *C. saccharobutylicum* LsrB binds R-THMF

To determine the form of AI-2 recognized by CsLsrB and CaLsrB and the identity of the AI-2 binding residues, we determined the crystal structure of CsLsrB. This receptor was expressed in a LuxS⁺ *E. coli* strain and crystallized; the structure was solved at 1.35Å resolution by molecular replacement omitting the ligand from the molecular replacement model (Table S2.3). The crystal structure shows that CsLsrB has a classic periplasmic binding protein fold like the canonical LsrB receptors from *Salmonella* serovar Typhimurium (46), *Salmonella enterica* serovar Typhi (PDB ID: 5GTA), *Sinorhizobium meliloti* (65), *Yersinia pestis* (64), and *B. anthracis* (PDB ID: 4PZ0), with two α/β domains connected through a three-

stranded hinge. The binding site is located near this hinge in the cleft between the two domains (Fig. 2.4, A). Superimposition of the CsLsrB structure with that of *B. anthracis* LsrB shows that the structures of the two proteins are very similar (RMSD 0.78 Å (Fig. 2.4, A)).

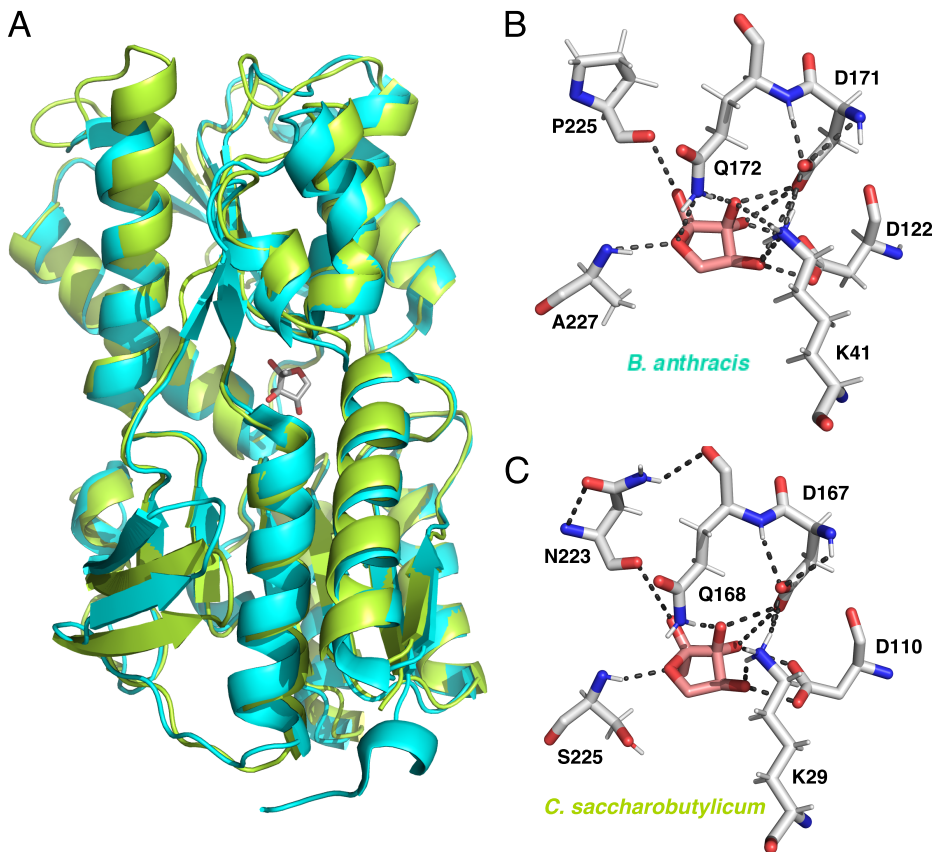


Figure 2.4 - Crystal structure of the LsrB receptor from *C. saccharobutylicum* in complex with AI-2. (A) Ribbon diagram of *C. saccharobutylicum* LsrB (green) superimposed to the LsrB of *B. anthracis* (PDB ID: 4PZ0; blue). The structures are very similar with a root mean square deviation of 0.78 Å, calculated by PyMOL. Bound AI-2 is shown as sticks. (B) Hydrogen bonds between AI-2 and LsrB from *B. anthracis* and (C) LsrB from *C. saccharobutylicum*.

After refinement of the structure, clear electron density showed the ligand to be R-THMF, the same form of AI-2 recognized by the canonical LsrB receptors. The crystal structure also revealed that, as predicted by Phyre2, the AI-2 binding residues in CsLsrB (and, presumably by extension, CaLsrB) differ from those of

the canonical LsrBs in two positions: a proline replaced with an asparagine and an alanine substituted with a serine (P225N, A227S numbering using *B. anthracis* LsrB as reference; compare Fig. 2.4, B with 2.4, C). The structure has been deposited in the PDB under the description “LsrB from *Clostridium saccharobutylicum* in complex with AI-2” with the PDB ID 6DSP.

2.3.4. The LsrB binding site is more plastic than previously thought

The observation that CsLsrB binds *R*-THMF with two of the six canonical AI-2 binding residues altered, led us to investigate whether other amino acid residues could be altered without loss of AI-2 binding ability. We performed alanine scanning mutagenesis on each of the six AI-2-interacting residues in the binding site (K29, D110, D167, Q168, N223, S225). To assess AI-2 binding of these mutants, each alanine mutant protein was expressed, purified, and tested for AI-2 binding ability via the *V. harveyi* MM32 assay. We observed that alanine substitutions in three out of the four amino acid residues that bind AI-2 through side chains (K29, D110 and Q168) caused loss of AI-2 binding ability and are thus necessary for binding (Fig. 2.5, A). The exception to this observation was D167, which is thought to make a hydrogen bond with AI-2 through the carboxylic group of the γ carbon and did not cause CsLsrB to lose its ability to bind AI-2 when substituted by an alanine.

A D167N substitution was also constructed since the asparagine mutation at this site was previously observed as one of two recurring amino acid substitutions in putative LsrB homologs that failed to bind AI-2 (17); AI-2 binding was abolished in this mutant (Fig. 2.5, A). The second substitution commonly observed in non-functional LsrB orthologues is a threonine at position 225. As CsLsrB has a serine in that position, we performed the single S225T substitution and observed that AI-2 binding was abolished (Fig. 2.5).

We also tested whether reverting the noncanonical binding site residues to the canonical ones present in previous characterized LsrBs would hinder AI-2

binding by replacing the serine at position 225 with an alanine (S225A) and the asparagine at position 223 with a proline (N223P). Both mutants were able to bind AI-2, showing that the binding site of the CsLsrB receptor is plastic enough to accept reversion to the canonical LsrB binding site composition. Altogether, these results allow us to conclude that the residues interacting with AI-2 in the LsrB binding site are not constrained to a single composition, as previously thought.

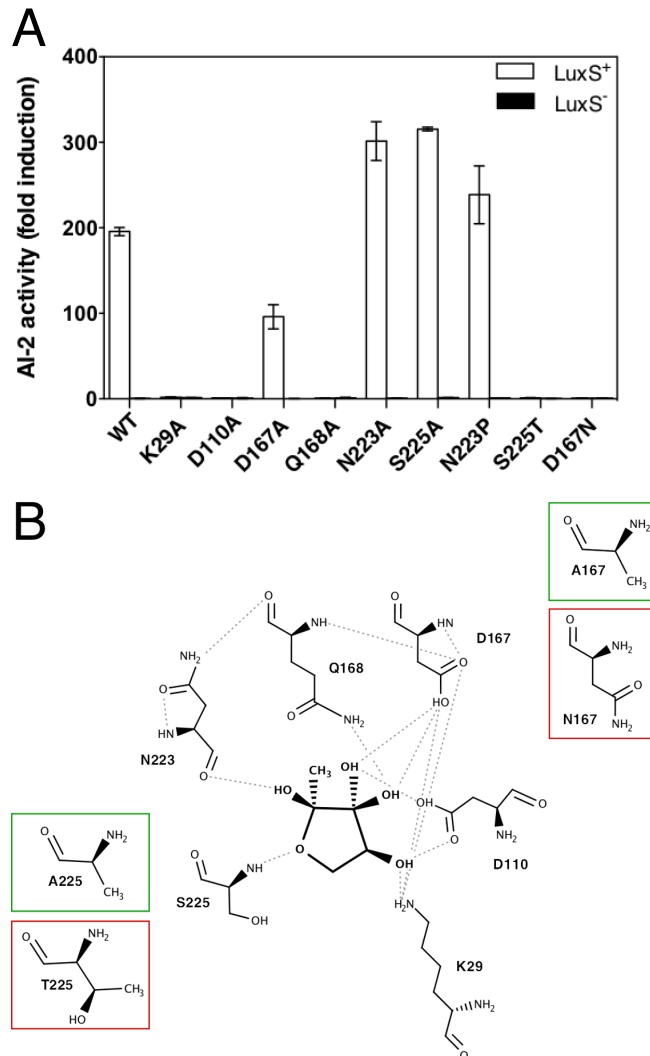


Figure 2.5 - Determination of the amino acid residues essential for binding of AI-2 by *C. saccharobutylicum* LsrB. (A) Binding of AI-2 to wild type and binding site mutants of CsLsrB. Binding assessed by measurement of light production of *V. harveyi* MM32 after addition of ligands released from purified proteins by thermal denaturation. Proteins were expressed either in LuxS⁺

(white bars) or LuxS⁻ (black bars) *E. coli* BL21 strains. Results are shown as fold induction relative to the light production induced by the growth medium. (B) Two-dimensional representation of the binding site of CsLsrB. Substitutions of aspartate 167 or serine 225 (which interfere with AI-2 binding) are boxed in red while substitutions that allow AI-2 binding are boxed in green.

2.3.5. LsrB receptors bind AI-2 with high affinity

Both the canonical LsrB proteins and the noncanonical CsLsrB studied here share the fold with the proteins of the large family of substrate binding proteins. Proteins of this family usually bind their ligands with high affinity; in fact, the affinity is sufficiently strong as to allow their purification with bound substrates (139). Dissociation constants in this family of proteins and their cognate ligands typically range between 0.1 and 1 μM for sugars and around 0.1 μM for amino acids (139). Thus, we would expect that LsrB receptors bind AI-2 with a dissociation constant in the submicromolar range. Supporting this hypothesis, CsLsrB purified and crystallized with AI-2 bound (above). To quantify this interaction, we used ITC to determine the affinity of CsLsrB for AI-2 and compared it with the affinities measured for the canonical LsrBs from *B. anthracis* and *E. coli*. As expected, we observed that all LsrB proteins tested were able to bind AI-2 with a dissociation constant (K_d) smaller than 1 μM (Fig. 2.6). We measured K_d of $0.81 \pm 0.13 \mu\text{M}$ for CsLsrB (Fig. 2.6, A), $0.20 \pm 0.04 \mu\text{M}$ for *B. anthracis* LsrB (Fig. 2.6, B) and $0.19 \pm 0.03 \mu\text{M}$ for *E. coli* LsrB (Fig. 2.6, C). As a negative control, we performed the ITC experiments with a substrate binding protein from *Rhizobium etli*, that is homologous to LsrB and thought to bind rhamnose (17,140). Both the raw data and the heat of reaction curve show that, even with AI-2 at a concentration nearly 4 times higher than in the other ITC assays, there is no specific binding, though nonspecific binding seems to occur without apparent saturation of the binding site (Fig. 2.6, D). Together, these results show that the AI-2 receptors tested bind AI-2 with a submicromolar affinity.

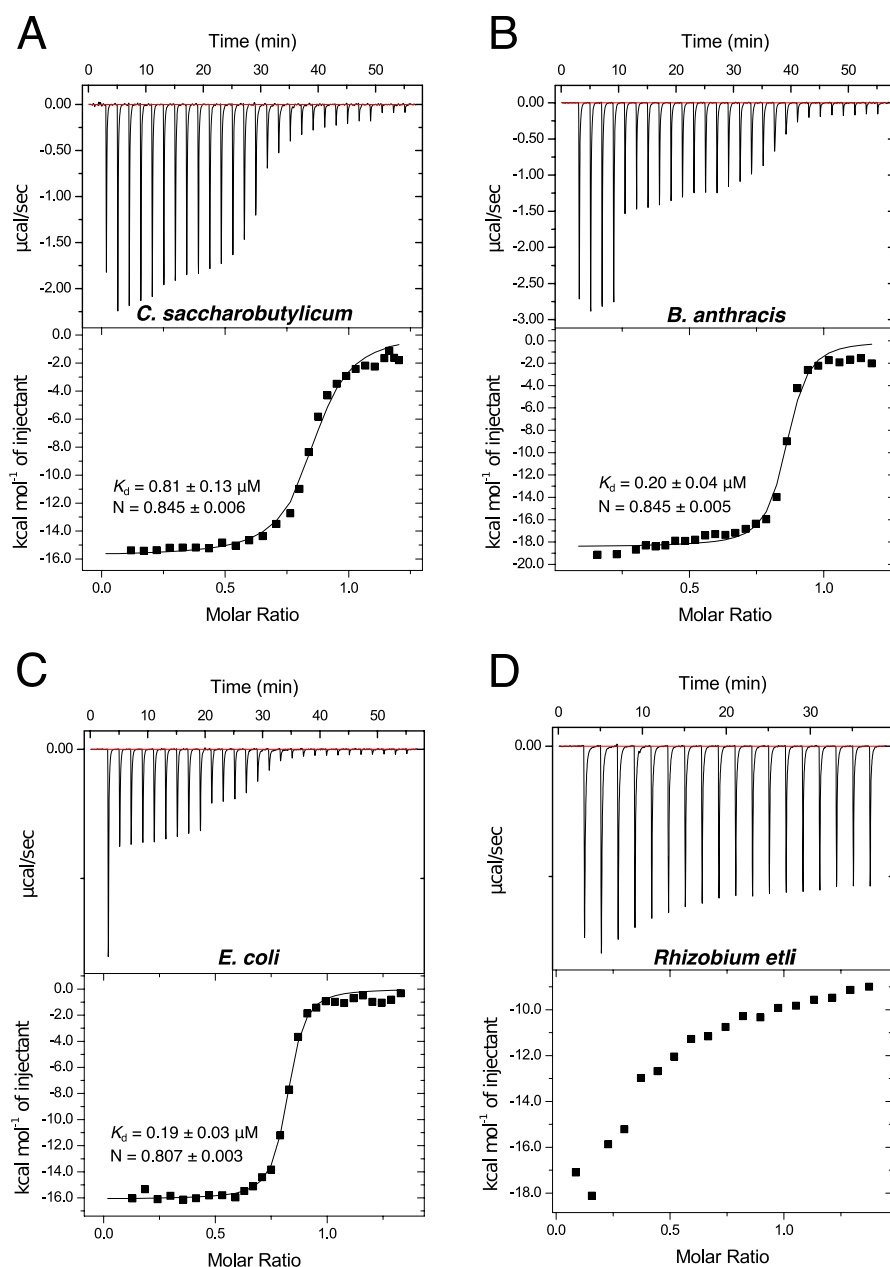


Figure 2.6 - LsrB receptors bind AI-2 with submicromolar affinity. Binding curves of AI-2 to LsrB proteins from (A) *C. saccharobutylicum*, (B) *B. anthracis*, (C) *E. coli*, and (D) from an LsrB ortholog from *Rhizobium etli* as negative control. Upper graphs show the raw ITC data and lower graphs show the integrated heat of reaction. Dissociation constants were obtained by fitting a one-site binding model to the displayed curves, which are representative of three independent experiments.

2.3.6. CsLsrB shares a common ancestor with canonical LsrB receptors

The identification of a functional LsrB receptor with naturally occurring variations in the amino acid residues of the binding site led us to inquire if CsLsrB arose from the same ancestor as the canonical LsrBs. To answer this question, a phylogenetic analysis was performed by querying the protein sequence of CsLsrB against the UniProt database of Reference Proteomes. Our analysis indicates that the large majority of LsrB proteins are distributed between two unrelated phyla, the Proteobacteria, particularly in Gammaproteobacteria and Alphaproteobacteria and the Firmicutes, from the orders Clostridiales and Bacillales (Fig. 2.7). The LsrB sequence of *C. saccharobutylicum* clusters in close proximity with other sequences from Clostridiales and in the same evolutionary lineage as the LsrB proteins from the *B. cereus* group previously shown to have functional AI-2 receptors (Fig. 2.7). There is clear evidence of lateral gene transfer events in the evolution of LsrB sequences, which explains the disjoint presence of LsrB orthologs within the two unrelated phyla (Proteobacteria and Firmicutes), as well as the presence of LsrB sequences from Actinobacteria and Spirochaetes clustering within a clade largely composed of Firmicutes species.

To obtain further support for the evolutionary position of CsLsrB, we performed a similar phylogenetic analysis focusing on the ATP-binding proteins (LsrA), a protein typically encoded by a gene co-transcribed with *lsrB* as part of the *lsr* operon. Given that the function of these two proteins is usually associated, we expected their evolutionary histories to be congruent. Importantly, the ATP-binding proteins from the ABC transporters are known to evolve more slowly than the substrate-binding proteins and are therefore better genetic markers for tracing the evolution of the genes in this operon (141,142). The evolution of LsrA sequences corroborated all major conclusions obtained with LsrB (Fig. S2.4). Although some local differences in topology and bootstrap support between the two trees do occur, these differences are likely due to the different rates of protein

evolution and do not cause conflict with the interpretation of the evolution of *C. saccharobutylicum* Lsr genes. The inferred evolution of LsrA sequences also corroborates the close relationship of the same species of Spirochaetes and Actinobacteria with the Clostridiales, which strongly suggests that events of operon lateral transfer are responsible for the observed taxonomic distribution of Lsr genes.

To determine which organisms have LsrB proteins with canonical (KDDQPA) or noncanonical (KDDQNS) predicted binding sites, we analyzed sequence conservation at the binding sites across species as shown in Fig. 2.7. Interestingly, some of the protein sequences in this phylogenetic tree have a serine instead of an alanine in position 225, as in CsLsrB, but retain the proline in position 223 observed for canonical receptors (KDDQPS) (Fig. 2.7). According to the results of our mutagenesis studies we infer that these noncanonical receptors with one substitution in the predicted AI-2 interacting residues will also be functional AI-2 receptors because the CsLsrB N223P mutant was able to bind AI-2 (Fig. 2.5, A).

Overall, the evolutionary history of the LsrB receptors supports the conclusion that lateral gene transfer is responsible for the appearance of canonical LsrB receptors in the Firmicutes, which subsequently evolved to noncanonical receptors in a stepwise manner accumulating one, then two, substitutions in the AI-2 binding residues. Most noncanonical receptors cluster together, but events of lateral gene transfer are again evident, as members from the Spirochaetes and Actinobacteria seem to have acquired noncanonical LsrB receptors with one or two substitutions.

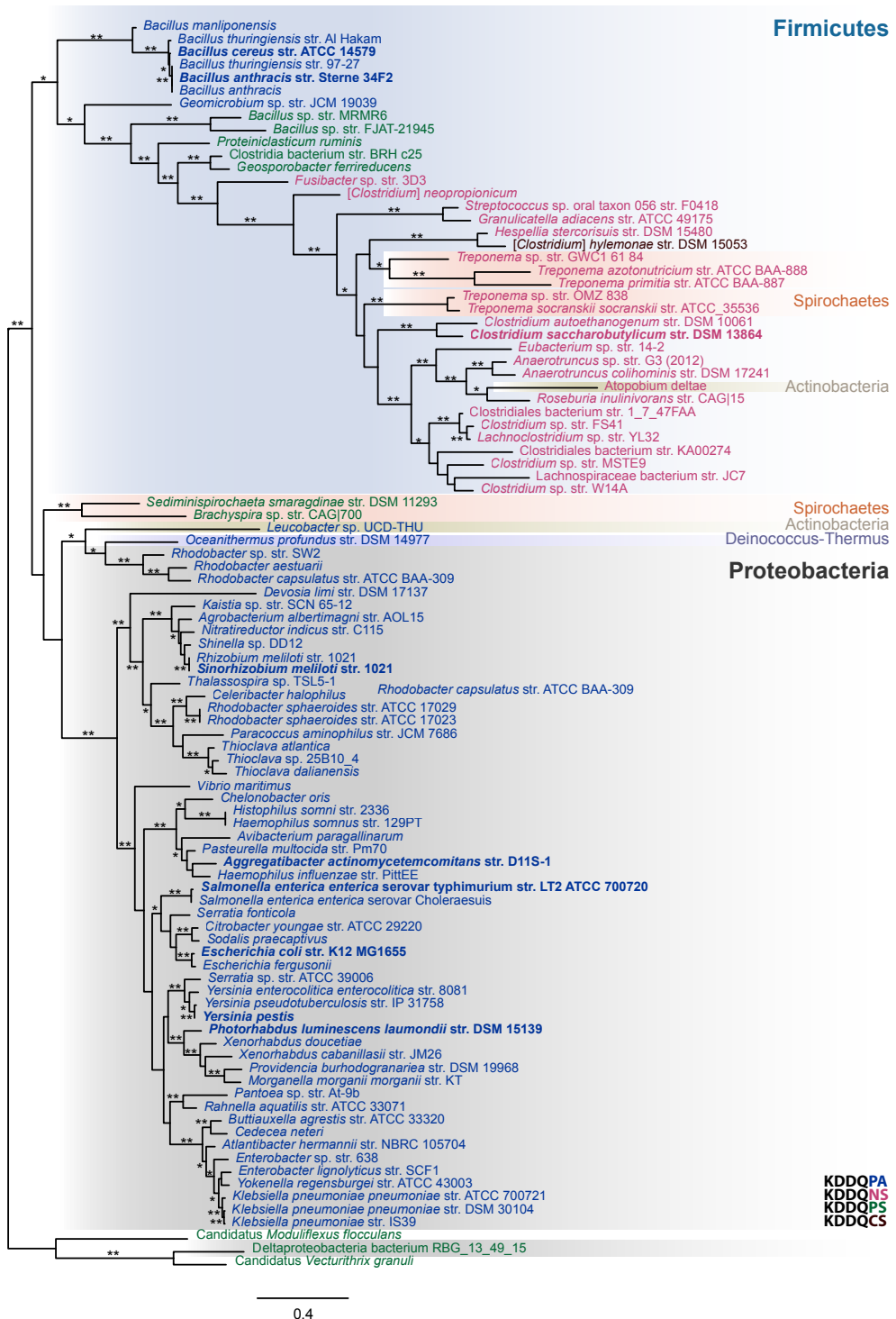


Figure 2.7 - Phylogenetic history of LsrB protein sequences. LsrB proteins with canonical residues (KDDQPA) are depicted in blue, non-canonical residues with one substitution (KDDQPS) in green

and non-canonical residues with two substitutions (KDDQNS) in red. In bold is our focal taxon *C. saccharobutylicum* and the organisms for which there are experimental studies supporting the functionality of the LsrB in at least one strain of these species (17,46,52,64,65,81,143). Taxonomic classifications are shown on the right and shades of the same color were used to group bacterial species belonging to the phylum following NCBI. Analyses were done with maximum-likelihood (LG+I+G model of protein evolution) and the tree was rooted following results from supplementary Fig. S2.5. Asterisk on the nodes represent bootstrap support estimated in RaxML using the autoFC option. * Values lower than 75%, ** values higher than 75%.

2.3.7. Identification of additional putative noncanonical LsrB receptors

Following the phylogenetic analysis of the LsrBs, it became clear that the noncanonical LsrB receptors were present in organisms other than *C. saccharobutylicum* and *C. autoethanogenum*. For this reason, we conducted a comprehensive bioinformatic analysis of proteomes currently available in the NCBI RefSeq non-redundant protein database using the sequence of CsLsrB to further identify organisms that possess noncanonical LsrB proteins with one (KDDQPS) or two substitutions (KDDQNS) in the putative binding site.

We identified 95 noncanonical putative LsrB receptors of which 26 have one substitution and 69 have two substitutions, as CsLsrB. The organisms encoding these proteins belong to the same phyla identified in our phylogenetic analysis with the large majority belonging to the Firmicutes phylum. To characterize these hits, all 95 protein sequences were submitted to the fold recognition server Phyre2. The predicted structures for all putative LsrB hits were consistent with the class I periplasmic binding protein fold shared by the canonical LsrBs and the noncanonical CsLsrB (Table S2.4a). We also determined whether these hits had Lsr transport proteins next to LsrB in their genomes (Table S2.4a). From the 95 hits, only the LsrB proteins from *Treponema primitia* and *Treponema azotonutricium* did not have Lsr transport proteins next to the LsrB protein (Table S2.4). In these genomes the LsrB homolog is located near a hybrid sensor histidine kinase/response regulator, although homologs for Lsr proteins were identified in

another location in the genome (Table S2.4). The presence of the enzymes involved in AI-2 metabolism was also assessed. Even though LsrF was present in 93 of the 95 hits, LsrG was only present (with a query cover higher than 75%) in 9 proteomes. Finally, we checked the proteomes for the existence of the AI-2 synthase LuxS. We did not identify LuxS homologs in all these organisms (79 LuxS homologs in 95 hits; Table S2.4). However, this is also true for *Sinorhizobium meliloti*, a species that carries a functional *lsr* operon but is unable to produce AI-2 (65). Altogether, these results suggest that functional, noncanonical LsrB receptors are present in many bacterial species, particularly in the Firmicutes phylum.

2.4. Discussion

LsrB AI-2 receptors share an identical fold to substrate-binding proteins that recognize a wide variety of substrates including 5-carbon sugars similar to AI-2. Thus, the identification of LsrB receptors has been possible due to the conservation of the six amino acid residues that bind AI-2, which allows the distinction between LsrB receptors and other substrate-binding proteins. In this study, we provide the first report of a functional LsrB receptor in Clostridia; this receptor has a previously unobserved binding site variation and gives insights into the plasticity of the binding site of LsrB receptors. The crystal structure of this receptor identified in *C. saccharobutylicum*, CsLsrB, showed that it binds the R-THMF form of AI-2 with two variations in the amino acid residues that bind AI-2: an asparagine instead of a proline (P223N) and a serine instead of an alanine (A225S) (Fig. 2.4 B and C). The fact that the substitution of the nonpolar proline and alanine for the polar asparagine and serine do not abolish AI-2 binding was surprising both due to the dissimilar nature of the amino acids and the fact that the canonical AI-2 binding residues were thought to be essential. However, examination of both canonical and noncanonical AI-2 binding pockets reveals that the amino acids in these positions bind AI-2 through the backbone. Moreover, the crystal structure of CsLsrB shows that the amino and carboxyl groups are still

positioned to form hydrogen bonds with similar geometry to canonical receptors; thus the local region must be plastic enough to accept side chain variation without significantly changing the overall structure of the protein. Still, not all amino acid residues can be accommodated in these positions as the S225T substitution caused loss of AI-2 binding. The side chain hydroxyl group of the threonine would possibly be shifted by the presence of the additional methyl group, causing a change in geometry that would likely impact the interaction with the ligand (Fig. 2.5). In addition, this shift could alter the interaction between S225T and N33, potentially causing a conformational change that impacts the surrounding amino acids, altering the conformation of the binding site further. As for D167N, which similarly to S225T, is a naturally occurring substitution in sugar binding proteins with homology to LsrB, the inability to bind AI-2 might be a result of the change from a presumably negatively-charged side chain to a side chain containing an amine. This could lead to unfavorable interactions with the ligand or the adjacent Q168, which also binds AI-2 through side chain interactions. Moreover, the introduction of the amine in this mutation could lead to more fundamental conformational changes in the binding site and surrounding environment than might be expected for a simple alanine mutation, explaining the loss of binding activity in the Asn but not the Ala mutant (Fig. 2.5). Accordingly, S225T and D167N substitutions also caused loss of AI-2 binding ability in *B. anthracis* LsrB (17), a canonical receptor, indicating that the presence of a threonine in position 225 and an asparagine in position 167 might be characteristic of sugar-binding proteins.

The two variations in AI-2 binding residues of CsLsrB might explain the slightly lower affinity of this receptor to AI-2 when compared with the canonical LsrBs tested. Nevertheless, CsLsrB bound AI-2 with high affinity (submicromolar K_d ; Fig. 2.6). Interestingly, in a previous study by Zhu and Pei a significantly higher K_d (160 μ M) was reported for the binding of the LsrB from *Salmonella* serovar Typhimurium to AI-2 (144). However, our results were obtained in the

absence of boron and, as was shown in Miller *et. al*, the removal of boron is essential to avoid shifting the equilibrium of AI-2 molecules toward the S-THMF-borate form (46). Hence, the higher apparent K_d value obtained by Zhu and Pei might result from the presence of boron. Moreover, the authors reported a K_d of 0.16 μM for LuxP from *V. harveyi* (144), a value closer to the ones we obtained for the LsrBs tested here and more consistent with the values expected for substrate binding proteins towards their cognate ligands (139).

Similarly to canonical LsrB receptors, CsLsrB was able to internalize AI-2 (Fig. 2.3), likely through interaction with the ABC transporter encoded in the same operon as *lsrB* (Fig. 2.2, A). Curiously though, AI-2 uptake started at mid-exponential phase (Fig. S2.1, B) earlier than what is usually observed for *E. coli* K12 MG1655 that starts internalization in the transition from exponential to stationary phase (Fig. S2.1, C and (17,52)). Additionally, *C. saccharobutylicum* supernatants induced 10-fold less light production, indicating that it accumulates lower concentrations of extracellular AI-2. The observed difference is likely because *C. saccharobutylicum* starts AI-2 uptake at lower extracellular AI-2 concentrations indicating that, at least under the laboratory conditions tested, *C. saccharobutylicum* has a lower concentration threshold than *E. coli*. This might indicate that the AI-2 concentration needed for induction of the *lsr* operon in *C. saccharobutylicum* is lower than in *E. coli* K-12 MG1655. As *C. saccharobutylicum* and *E. coli* LsrB receptors have high affinity to AI-2, it is unlikely that differences in affinities of the receptors cause the variation in concentrations required for the induction of the operons and, thus, in the start of internalization. Moreover, *E. coli* can accumulate extracellular AI-2 to concentrations up to 40 μM without starting to internalize AI-2. This indicates that, in *E. coli*, additional mechanisms are preventing this bacterium from starting internalization of AI-2 at lower concentrations (73). In fact, in *E. coli*, the Lsr system is regulated not only by LsrR but also by catabolite repression and the phosphoenolpyruvate phosphotransferase (PTS) system, such that the availability of PTS substrates

inhibits Lsr-mediated AI-2 internalization (72,73). Thus, it is possible that these additional regulatory mechanisms are either different or not present in *C. saccharobutylicum*, allowing internalization to start at lower concentrations of AI-2. Further work on the regulation of these systems is necessary to address this hypothesis. Moreover, during the testing of growth media for *C. saccharobutylicum* we observed that *E. coli* K-12 MG1655 does not internalize AI-2 under anaerobic conditions (Fig. S2.2). Additionally, a $\Delta lsrR$ mutant is able to internalize AI-2, suggesting that the inhibition caused by the absence of oxygen is regulated by LsrR (Fig. S2.2). However, the observation that *C. saccharobutylicum* is able to internalize AI-2 in anaerobic conditions suggests that this effect might be specific to bacteria that are facultative anaerobes. Further studies with other facultative anaerobe bacteria and/or commensal *E. coli* strains would be needed to understand if this effect is biologically significant. Nevertheless, it is worth to keep this observation in mind when *E. coli* K12 MG1655 is used to manipulate AI-2 levels in the mammalian gut.

The search for new noncanonical receptors revealed that in most organisms where putative noncanonical LsrB receptors were identified, the *lsrB* neighboring genes encoded homologs for the components of the Lsr transport system. In contrast, homologs of the AI-2 processing enzymes, LsrG and LsrF, identified in *E. coli* and *Salmonella* serovar Typhimurium, were not as widely conserved. For example, in *C. saccharobutylicum* no LsrF ortholog was identified and in both *C. saccharobutylicum* and *C. autoethanogenum*, the gene that encodes the LsrG orthologs is present in another region of the genome (Fig. 2.2, A). These results, together with previous reports of canonical LsrB receptors encoded in operons that also encode functional Lsr transport systems but have one or both LsrG and LsrF orthologs located in other regions of the genome (17,65) indicate that AI-2 metabolism might not be as conserved as the steps required for the uptake, phosphorylation and regulation of the *lsr* operon. Thus, the lack of orthologs for the proteins responsible for AI-2 metabolization (LsrG and LsrF) should not be an

excluding factor when searching for new LsrB receptors. Curiously, in *T. primitia* and *T. azotonutricium* we identified putative noncanonical LsrBs that were near a sensor kinase instead of an ABC transporter in the genome. LuxP, the other type of known AI-2 receptor known, interacts with a sensor kinase (LuxQ) that facilitates signal transduction through a phosphorylation cascade (2). However, when searching for homologs of the Lsr proteins, we found hits for all the proteins, except LsrR, located in a more distant region of the genome (Table S2.4). Hence, further studies are needed to understand if these putative LsrBs evolved to interact with the nearby sensor kinase or if they instead use an ABC transporter located in another region of the genome.

In summary, the characterization of the amino acid variation in the AI-2 binding site of CsLsrB fulfills our initial motivation as it allowed the identification of new putative LsrB receptors in microbes present in biologically relevant niches like *Roseburia inulinivorans* (145) and *Clostridium merdae* (146), Firmicutes isolated from human microbiota that appear to have a noncanonical receptor with two substitutions (like CsLsrB). This is a significant step towards understanding the molecular mechanisms involved in interspecies communication in these niches. In particular, we expect the identification of AI-2 receptors in Firmicutes to help in the characterization of the mechanism by which their colonization is favored in the presence of AI-2 in the mammalian gut microbiome, a clinically relevant niche where interspecies interactions are highly prevalent.

2.5. Experimental procedures

2.5.1. Identification of LsrB orthologs

The search for LsrB orthologs was performed as described previously (17). Briefly, we searched the complete genomes in the KEGG SSDB (Sequence Similarity Database, 3478 bacterial genomes in September 2015) for amino acid sequences with similarity to LsrB from *B. cereus*. All pairwise genome comparisons

were performed using SSEARCH program and best bidirectional hits with a Smith-Waterman score of at least 120 were selected. From the obtained hits with sequence identity between 30-60% we focused on the putative LsrBs of *C. saccharobutylicum* and *C. autoethanogenum*. The genome context of the CsLsrB and CaLsrB genes was assessed through analysis of the neighboring genes. To determine the presence/absence of Lsr orthologs in *C. saccharobutylicum* and *C. autoethanogenum*, we looked for best bidirectional hits using the Lsr proteins of *B. cereus* ATCC 10987 and *B. anthracis* Sterne as query. Simultaneously, we determined their similarity (Table S2.2). The amino acid sequences of the putative LsrBs from *C. saccharobutylicum* and *C. autoethanogenum* were submitted to the Phyre2 fold recognition server. The presence/absence of conserved amino acid residues in AI-2 binding was determined through structural alignment of the structures predicted by Phyre2 with the LsrB from *B. anthracis* (PDB ID: 4PZ0) using COOT (147).

2.5.2. AI-2 synthesis

The AI-2 precursor DPD was synthesized as previously described (51). Boron-free DPD was obtained by performing the synthesis with plastic material and using boron-free water (prepared by batch incubation with Amberlite IRA743 resin for 2h at room temperature, as described before (148)).

2.5.3. AI-2 internalization assays

Anaerobic frozen stocks of *C. saccharobutylicum* DSM13864 were revived on modified PY+X solid medium with galactose (see Table S2.1 for media composition). 13 mL of broth (modified PY+X with arabinose and 100 mM MOPS buffer, pH 7.0) was inoculated with 3-6 colonies of *C. saccharobutylicum* and grown overnight at 37 °C in an anaerobic chamber (PlasLabs, USA) under a gaseous mix of 80% nitrogen, 15% carbon dioxide and 5% hydrogen to an optical density at 600 nm (OD_{600nm}) lower than 4. A 20% (v/v) inoculum was added to fresh medium

and incubated at 37 °C without agitation in a 100 mL Hungate Schott flask. Culture suspensions were collected at the stated time points for optical density measurement at 600 nm (UV-visible spectrophotometer; Helios Delta, ThermoSpectronic, USA) and for detection of AI-2. *E. coli* $\Delta luxS \Delta lsrK$ (*E. coli* ARO093) (21) was revived directly from the frozen stock to liquid medium and a 8% inoculum was performed. Exogenous AI-2 was supplemented to a final concentration of 40 μ M in fresh media. For AI-2 activity measurements, culture suspensions were filtered using multiscreen filter plates (Millipore). The cell-free culture was frozen at -20°C overnight and AI-2 activity was accessed in triplicate using the *V. harveyi* MM32 bioluminescence reporter assay, as described previously (45,65). Light production was measured at 7h for *C. saccharobutylicum* cell-free culture fluids and at 5h for the cultures supplemented with 40 μ M AI-2. Luminescence was measured with a Glomax Explorer microplate luminometer (Promega, EUA). AI-2 activity is reported as the induction of light production compared with the background light obtained with the appropriate growth medium. Standard deviation was calculated from three technical replicates. Propagation of uncertainty was employed to calculate the standard deviation after normalization. This experiment was repeated on three different days with three independent cultures. For simplicity, a representative experiment of the three independent experiments is shown (Fig. 2.3). It was not possible to analyze the data from the three different experiments together due to variations of the growth curves, in particular with respect to differences in the lag phase of the cultures that varied from 30 min to 4 hours. The three experiments are shown in Fig. S2.3.

2.5.4. Protein expression and purification

The gene encoding LsrB ortholog in *C. saccharobutylicum* DSM13864 was amplified from genomic DNA (DSMZ) and cloned in the plasmid pDEST-527 (gifted by Dominic Esposito, Addgene plasmid #11518) via the pENTR/TEV/D-TOPO cloning kit (Thermo Fisher Scientific) for expression as a His6-tagged

protein. The signaling sequence for secretion, as determined by SignalP 4.1 (MKKKAAVALALIGAMIFTTLVGCG), was excluded from the construct. *E. coli* BL21 (DE3) LuxS⁺ or LuxS⁻ cells were transformed with the construct and grown in LB with 1 µg/mL ampicillin at 37 °C until the OD_{595 nm} was 0.3. At this point, temperature was decreased to 22 °C. At OD_{595 nm} = 0.9, 0.3 mM of IPTG was added and the cells were induced for 6h before being harvested by centrifugation. Cells were resuspended in 50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl, 10 mM imidazole, 2.5 µg/mL DNase and 2.5 µg/mL leupeptin and lysed using a M-110Y microfluidizer (Microfluidics, USA). The lysate was centrifuged and the tagged protein was purified from the clarified supernatants using Nickel-nitrilotriacetic (Ni²⁺-NTA) acid affinity chromatography (Qiagen). The protein was eluted from the column in 50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl and 250 mM imidazole and subsequently swapped into 25 mM Tris-HCl pH 8.0, 50 mM NaCl and 1 mM dithiothreitol (DTT) using Sephadex-G25 agarose. The His-6-tag was removed by cleavage with tobacco etch virus (TEV) protease at a proportion of 1 mg TEV protease per 250 mg of protein at 4 °C, overnight. A second round of Ni²⁺-NTA affinity chromatography was performed to remove the His-tagged TEV protease, the cleaved tag and any uncut fusion protein. The protein collected from the flow-through was buffer swapped into 25 mM Tris-HCl pH 8.0, 50 mM NaCl and 1 mM DTT via Sephadex-G25 agarose and further purified by anion exchanging chromatography (SourceQ column, GE Healthcare Life Sciences) using a NaCl gradient from 0 to 1 M. As a final purification step, the protein underwent a size exclusion chromatography on a Superdex75 (GE healthcare) column and was eluted in 25 mM Tris-HCl pH 8.0, 150 mM NaCl and 1 mM DTT.

2.5.5. Crystallization studies

Crystals of *C. saccharobutylicum* LsrB expressed in *E. coli* BL21 LuxS⁺ were grown via the sitting drop method with a well solution of 0.1 M citric acid pH 2.75 and 26% w/v PEG 3350 and developed in approximately one week at room

temperature. Crystals were frozen after a 30 second soak in a solution of 27% w/v PEG 3350 plus 15% v/v glycerol. Diffraction data was collected on beamline BL14-1 at the Stanford Synchrotron Radiation Light Source. Data was processed with CCP4 software (149). A molecular replacement solution was determined via PHENIX (150) using 1TJY as the search model. The initial model was built by PHENIX with subsequent manual building in Coot (147) and refinement in PHENIX. The resulting PDB file is used as a reference for amino acid numbering of CsLsrB throughout this manuscript. Thus based on the structure, the amino acids of the binding site are numbered K29/D110/D167/Q168/N223/S225, while the numbering based on the complete amino acid sequence of the protein is K48/D129/D186/Q187/N242/S244.

2.5.6. Site directed mutagenesis

The QuikChange Lightning Site Directed Mutagenesis kit (Agilent) was employed to make the single amino acid substitutions in pDEST527/*C. saccharobutylicum* LsrB constructs. Primers were designed using the QuikChange primer design program (Agilent) and sequences are given in Table S2.5. The mutant proteins were expressed as described above. The His₆-tagged mutants were purified through Ni²⁺-NTA affinity chromatography (Qiagen) as above and then buffer swapped into 25 mM Tris-HCl pH 8.0, 150 mM NaCl and 1 mM DTT using PD-10 desalting columns (GE Healthcare). Protein was concentrated in 10 kDa Centricons (Millipore) until a concentration of at least 10 mg/mL was achieved.

2.5.7. *In vitro* AI-2 binding assay

The *in vitro* AI-2 binding ability of the studied proteins was assessed as previously described (17). Briefly, LsrB proteins at a concentration of 10 mg/mL were heated at 70 °C for 10 min. The denatured protein was pelleted via centrifugation and 10 µL of the supernatant were added to 90 µL of a 1:5000

dilution of an overnight culture of *V. harveyi* MM32, as described above. Bioluminescence after 5h of incubation at 30 °C was measured. A 1420 Victor 2 plate reader (Perkin Elmer, USA) and a 1420 Victor 3 plate reader were used for data in Fig. 2.2 and 2.5, respectively. Binding data is representative of three independent experiments and the standard deviations are derived from three technical replicates. Propagation of uncertainty was used to calculate the standard deviation after normalizing the light produced by AI-2 by the light produced by buffer alone.

2.5.8. Isothermal titration calorimetry

ITC measurements were performed in a MicroCal iTC200 microcalorimeter (GE Healthcare Biosciences, Northampton, MA, USA) at 25 °C. Boron-free AI-2 was diluted in boron-free buffer containing 25 mM sodium phosphate buffer pH 8.0, 150 mM NaCl 1 mM β -mercaptoethanol. Amberlite IRA743 resin was employed to remove the boron from the water used to prepare buffer as previously described (148). To avoid contamination by boron silicates present in glass, only plastic material was used. AI-2 at 800 μ M was injected into 117.4 μ M and 108 μ M of LsrB protein from *C. saccharobutylicum* and *E. coli* K-12 MG1655, respectively. For *B. anthracis* Sterne 34F2 LsrB, a solution of 750 μ M AI-2 was added to 106.9 μ M of protein. As a negative control, 3 mM of AI-2 was injected into 114.7 μ M of the LsrB homolog from *Rhizobium etli* CFN42 (RHE-PE00289 in KEGG, annotated as substrate binding protein involved in the rhamnose transport system). Measurements were made with the reference power at 5 μ cal/s and a syringe stirring speed of 800 rpm. The heat of dilution for successive injections of AI-2 into buffer was included in the final analysis. The heat of reaction for each injection was calculated by integrating the area under each titration peak under the assumption of a one-site binding model using MicroCal/Origin 7.0 software (provided by the manufacturer). Injections were: *C. saccharobutylicum*: 1 $\times$ 0.5 μ L + 14 $\times$ 1.5 μ L + 10 $\times$ 1 μ L + 3 $\times$ 0.5 μ L; *B. anthracis*: 1 $\times$ 0.5 μ L + 4 $\times$ 2 μ L + 23 $\times$ 1 μ L; *E.*

coli: 1 x 0.5 μ L + 1 x 3 μ L + 8 x 1.5 μ L + 18 x 1 μ L; *R. etli*: 1 x 0.5 μ L + 18 x 2 μ L. Binding curves shown are representative of three different runs.

2.5.9. Phylogenetic analysis

A homology search was performed by querying the protein sequence of CsLsrB against the UniProt database of Reference Proteomes using phmmer (151). This database contains proteomes that have been selected either manually or algorithmically in order to provide a broad coverage of the tree of life and a balanced cross-section of the taxonomic diversity found within UniProtKB, thus removing over-representation of certain species of bacteria caused by over sampling of preferred niches. At the time of analysis (June 22, 2017) it contained a total of 10454 proteomes, 6469 of which were from Bacteria. To extract significant hits we applied different e-value thresholds of increasing stringiness starting at 10^{-4} and performed phylogenetic analyses on the aligned datasets to evaluate its evolutionary structure. The results presented here resulted from applying an e-value threshold of 10^{-30} , retrieving 167 protein sequences (Table 6, <http://www.jbc.org/content/early/2019/01/29/jbc.RA118.006938/suppl/DC1>). This threshold excluded distant homolog sequences while keeping all sequences in our focal group and in the sister clade identified as rhamnose-binding proteins. To this dataset we added other sequences that were previously shown to be either functional LsrB receptors (11 sequences) or rhamnose-binding proteins (2 sequences) (Table S7, <http://www.jbc.org/content/early/2019/01/29/jbc.RA118.006938/suppl/DC1>) (17). This dataset was aligned with Mafft (version 7.310) using the L-INS-i method and default parameter values (152). The aligned dataset was analyzed with Prottest version 3.42 (153) to estimate the most likely model of protein evolution and RaxML version 8.0.26 (154) to produce a maximum-likelihood inference of the phylogenetic history of these proteins. Nodal support in the phylogenetic analysis was estimated with nonparametric bootstrap using an automatic frequency-based criterion (autoFC option) to determine the number of

replicates. We identified two truncated proteins, one from *Clostridium magnum* str. DSM 2767 and the other from *Klebsiella pneumoniae* str. ISC21, that were eliminated from final analyses. The phylogenetic analysis of the 180 protein sequences homologous to CsLsrB indicated a clear separation between bona fide LsrB sequences and the paralogous sequences identified as rhamnose binding proteins (Fig. S2.5). This result was used to define a smaller dataset focused on exclusively LsrB sequences keeping all sequences within the most inclusive and highly supported clade that includes all confirmed functional LsrB sequences and our target LsrB sequence from *C. saccharobutylicum*. This final dataset of 97 protein sequences was realigned and reanalyzed as before. All sequences included have an *e*-value smaller than 10^{-63} and an average sequence identity of 43%. This smaller dataset produced a better alignment and a more accurate phylogenetic inference. For the phylogenetic study of LsrA, we performed a homology search by querying the ATP-binding protein of *C. saccharobutylicum* against the UniProt database of Reference Proteomes using *phmmer* algorithm. We extracted all significant hits using an *e*-value threshold of 10^{-95} , which retrieved 470 sequences. To this dataset we added all ATP-binding protein sequences that pertain to the same strain/operon as the functional LsrB (11 sequences), if not previously included, as well as the two ATP-binding proteins from the rhamnose-binding operon (Table S8, <http://www.jbc.org/content/early/2019/01/29/jbc.RA118.006938/suppl/DC1>). These sequences were analyzed as explained above for LsrB sequences to produce a phylogenetic analysis of *bona fide* LsrA sequences present in the Reference Proteome database (Fig. S2.4).

2.5.10. Identification of noncanonical LsrBs

The search for putative noncanonical LsrB receptors was accomplished using BLASTP against the NCBI RefSeq non-redundant protein database using CsLsrB as query sequence (in May 2018). RefSeq allows sequence submission and includes complete and incomplete genomes, thus having more genomes available than

KEGG Genome, which is a collection of organisms with complete genome sequences or than UniProt database of Reference Proteomes, which is a balanced database to inference evolutionary history (155,156). The alignments between the amino acid sequence of CsLsrB and the hit proteins were examined to determine the identity of the six amino acid residues involved in AI-2 binding. To simplify the analysis, the search was conducted for each one of the 35 phyla described in KEGG separately. From the functional mutations identified in this study, only putative LsrBs with N223P and N223P/S225A were found. The annotation of the proteins in the vicinity of these noncanonical hits was assessed. For the hits with nearby Lsr orthologs, the amino acid sequences of these proteins were submitted to a comparative BLASTP with the respective ortholog in *C. saccharobutylicum*. For the absent orthologs and for LuxS, a BLASTP directed to the proteome of the organisms using the matching ortholog in *C. saccharobutylicum* as query was performed. As no LsrF homolog was identified in *C. saccharobutylicum*, the amino acid sequence of LsrF from *E. coli* K12 MG1655 was employed. Additionally, very few hits for LsrG homologs were obtained using the *C. saccharobutylicum* LsrG as query, so we also checked for homologs using the biochemically characterized LsrG from *E. coli* K12 MG1655. All of the 95 putative LsrB sequences were submitted to the Phyre2 fold recognition server.

2.6. Acknowledgements

The work performed in the laboratory of Karina B. Xavier was supported by Portuguese National funds from Fundação para a Ciência e Tecnologia (FCT) PTDC/BIA-MIC/4188/14. Karina B. Xavier and Inês M. Torcato acknowledge FCT for individual grants IF/00831/2015 and PD/BD/105736/2014, being the latter within the scope of the PhD program Molecular Biosciences PD/00133/2012. The laboratory of Stephen T. Miller acknowledges Swarthmore College. Meghann R. Kasal acknowledges Mayer Davidson '57 and Frances Velay Womens Science research fellowships. The authors also thank Ana Rita Oliveira and Roberto

Balbontín for helpful suggestions, Osvaldo Ascenso, Vanessa Miranda and Rita Ventura for the boron-free DPD, Steven Chen and Kendall Yoshii for help with the *in vitro* AI-2 binding assay of CsLsrB and CaLsrB and Paula Chicau and Marco Domingues for technical support with ITC.

2.7. Supplementary information

2.7.1. Figures

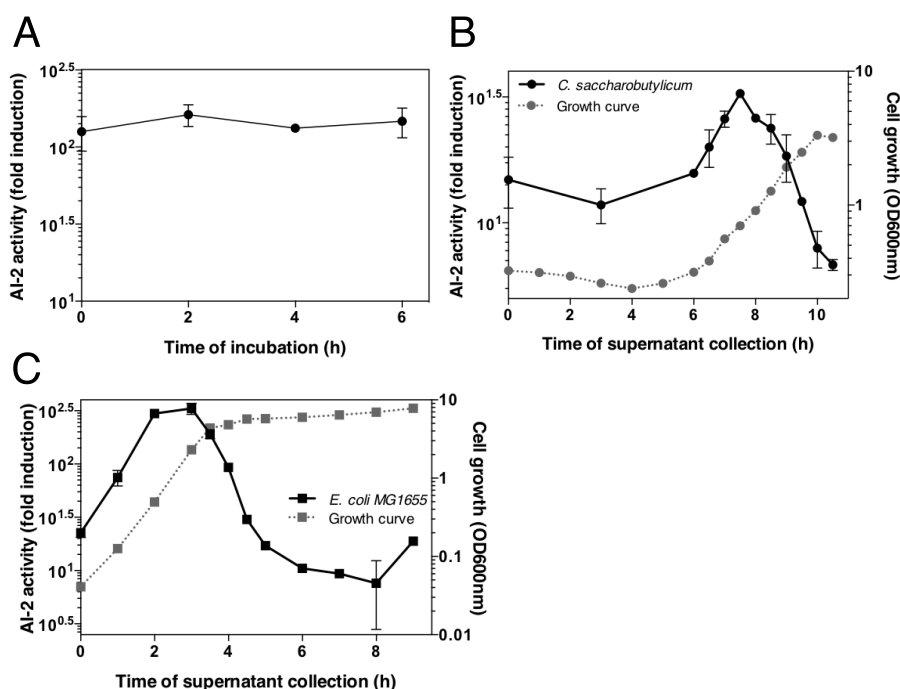


Figure S2.1 - AI-2 activity measured by *Vibrio harveyi* MM32. (A) activity of 40 μ M AI-2 when incubated with supernatants of *C. saccharobutylicum* at 9h30 of growth. (B) AI-2 internalization profile in *C. saccharobutylicum* (black circles) and respective growth curve (grey circles). (C) AI-2 internalization profile in *E. coli* (black squares) and respective growth curve (grey squares). AI-2 activity reported as fold induction relative to light production induced by the growth medium. Error bars represent the standard deviation of three technical replicates.

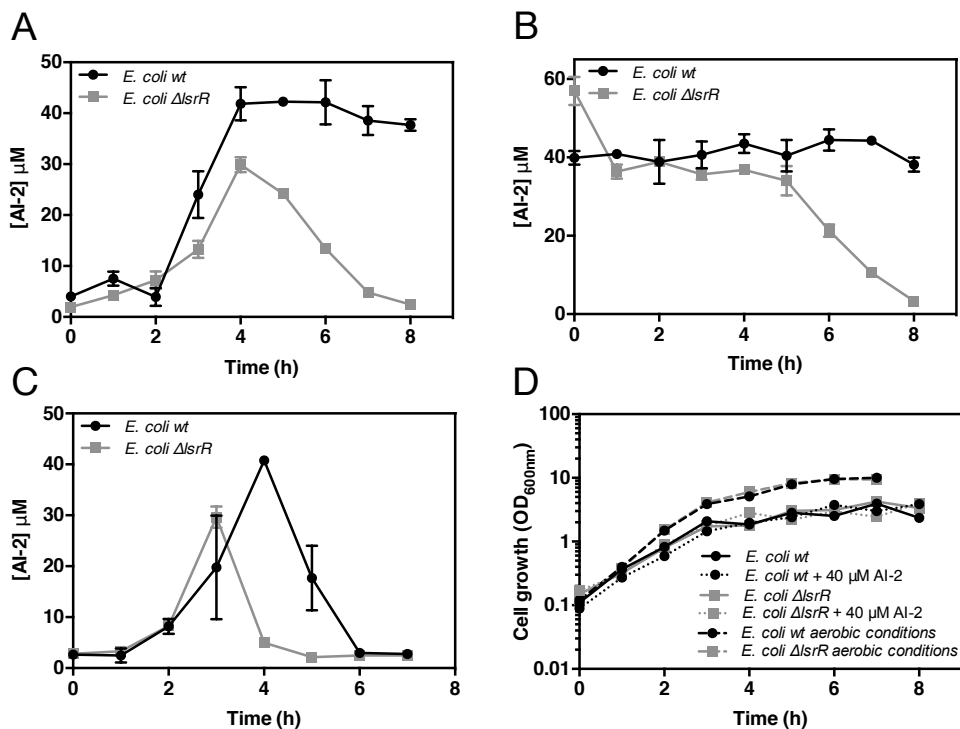


Figure S2.2 - AI-2 internalization by *E. coli* under aerobic and anaerobic conditions. AI-2 activity in cell-free fluids collected from a growing culture of *E. coli* K-12 MG1655 wild type and Δ *lsrR* mutant (A) under anaerobic conditions, (B) under anaerobic condition upon addition of 40 μM of AI-2 at time zero and (C) under aerobic conditions. Whereas the Δ *lsrR* mutant is able to internalize AI-2 in all conditions, wild type *E. coli* can only internalize AI-2 in the presence of oxygen. This inhibition is not reversed by an increase in the concentration of AI-2. Error bars represent the standard deviation of three technical replicates. (D) Growth curves of the studied strains under the different growth conditions.

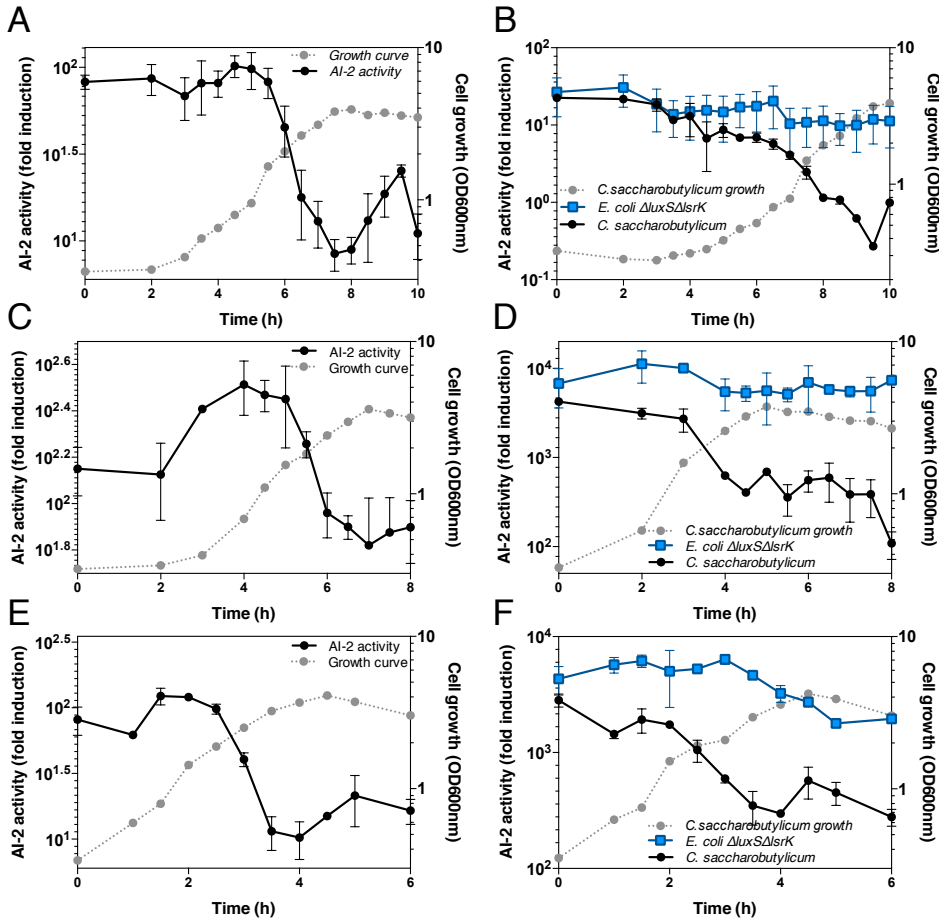


Figure S2.3 - AI-2 internalization by *C. saccharobutylicum*. (A), (C) and (E) internalization profile of *C. saccharobutylicum* (black circles) and respective growth curve (grey circles). (B), (D), (F) AI-2 internalization profile of *C. saccharobutylicum* (black circles) and *E. coli ΔlsrKΔluxS* (blue squares) after addition of 40 μ M AI-2 at time zero; *C. saccharobutylicum* respective growth curve is shown as grey circles. Cell free supernatants were collected at the indicated time points and activity of extracellular AI-2 was measured by assessing light produced by *V. harveyi* MM32 in response to cell-free supernatants. AI-2 activity is reported as fold induction relative to light production induced by the growth medium. The curves represent three independent experiments. Error bars represent the standard deviation of three technical replicates.

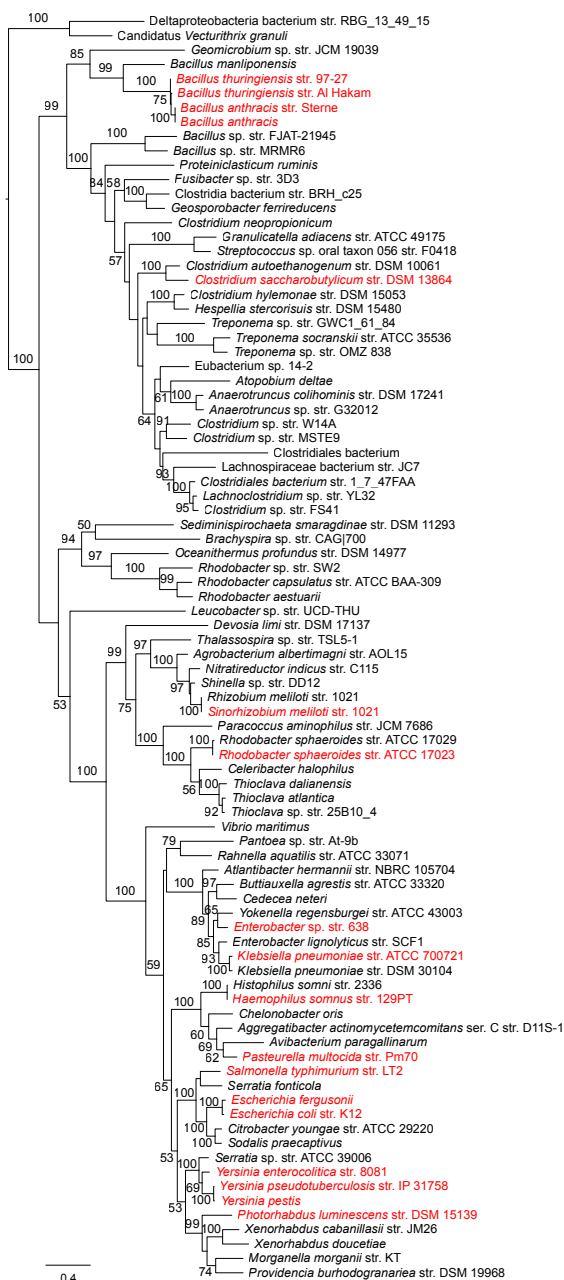
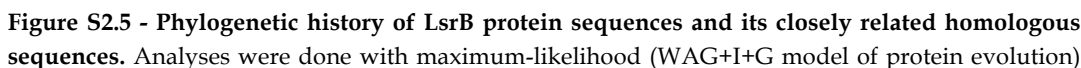


Figure S2.4 - Phylogenetic history of LsrA protein sequences. Analyses were done with maximum-likelihood (LG+I+G model of protein evolution) and the tree was rooted following previous results with a larger dataset that also includes the ATP-binding protein from the Rhamnose and the Ribose ABC-type transporters (data not shown). Numbers on the nodes are bootstrap support estimated in RaxML using the autoFC option. Only bootstrap support larger than 50% is shown. Scale bar indicates expected number of amino acids substitutions per site. In red we denote our focal taxon *C. saccharobutylicum*, taxa with LsrB sequences experimentally determined to be functional AI-2 receptors, and taxa from group I of (17).



in RaxML, and the tree was rooted by midpoint rooting. Numbers on the nodes are bootstrap support estimated with the autoFC option in RaxML. Only bootstrap support larger than 50% is shown. Scale bar indicates expected number of amino acids substitutions per site. In red we denote our focal taxon *C. saccharobutylicum*, taxa with LsrB sequences experimentally determined to be functional AI-2 receptors, and taxa from group I of (17). In green are taxa represented with Rhamnose-binding proteins.

2.7.2. Experimental procedures

2.7.2.1. Incubation of AI-2 in *C. saccharobutylicum* supernatants.

C. saccharobutylicum culture at 9h30 of growth was collected and filtered using a 0.2 µm filter (Acrodisc, Pall Life Sciences) originating cell-free supernatant that was frozen at -20°C, overnight. The supernatant was supplemented with synthetic AI-2 to a final concentration of 40 µM and incubated for 2h, 4h or 6h at 37 °C without agitation in a transparent, flat bottom 96 well plate (Costar 3596). After incubation, supernatants were frozen at -20 °C, overnight. AI-2 activity was measured using *V. harveyi* MM32 bioluminescence reporter assay as described for the internalization assays. Data presented corresponds to 5h of incubation with *V. harveyi* MM32 at 30 °C. Standard deviation was calculated from three technical replicates and propagation of uncertainty was used to calculate the standard deviation after normalization of the light produced by AI-2 by the light produced by the medium.

2.7.2.2. AI-2 internalization in *C. saccharobutylicum* and *E. coli*.

Determination of AI-2 activity in *C. saccharobutylicum* was performed as described in the experimental procedures of the main text in the section of AI-2 internalization assays. Determination of AI-2 activity in *E. coli* K12 MG1655 cell cultures during growth was performed as previously described (17). Briefly, *E. coli* was revived from the -80 °C frozen stock directly to 5 mL of broth (modified PY+X with arabinose and 100 mM MOPS buffer, pH 7.0; Table S2.1) and grown at 37 °C

with 200 rpm of agitation until saturation (under aerobic conditions). A dilution of approximately 100 times to 12.5 mL of fresh medium in a 125 mL erlenmeyer was performed and culture suspensions were collected at the indicated time points for optical density measurement at 600 nm in a UV-visible spectrophotometer (Helios delta, ThermoSpectronic, USA) and for detection of AI-2. For AI-2 activity assessment, the culture suspensions were filtrated originating cell-free fluids (Multiscreen filter plates, Millipore). The cell free supernatants were frozen at -20 °C overnight and AI-2 activity in these fluids was accessed in triplicate using *V. harveyi* MM32 bioluminescence reporter assay, as described above. Data presented corresponds to 6h of incubation with *V. harveyi* MM32, at 30 °C.

2.7.1.3. AI-2 internalization in *E. coli* under anaerobic and aerobic conditions

The protocol for cell free fluids collection was very similar as the one described before. Briefly, *E. coli* K-12 MG1655 (ARO 045) and *E. coli* Δ *lsrR* (JAT 156) (21) were revived from the -80 °C frozen stock directly to 5 mL of rich media (17.5 g/L meat extract, 10 g/L peptone, 5 g/L NaCl, 2.5 g/L Na₂HPO₄, 0.25 g/L cysteine-HCl, 1 g/L NaHCO₃ and 2 g/L arabinose). The strain ARO 045 was used as the wild type as it was the strain used to construct the LsrR mutant. For aerobic conditions, 100 µL of overnight cultures grown at 37 °C, with 240 rpm shaking, were added to 10 mL of rich media in 100 mL erlenmeyers and were further grown at 37 °C with shaking. For anaerobic conditions, 500 µL of overnight cultures grown at 37 °C in an anaerobic chamber (PlasLabs, USA) under a gaseous mix of 80% nitrogen, 15% carbon dioxide and 5% hydrogen, were added to 10 mL of rich media in 100 mL Hungate Schott flasks. To the cultures supplemented with a final concentration of 40 µM AI-2, synthetic AI-2 was added to fresh media, immediately before inoculation. At the mentioned time points, bacterial suspensions were collected for OD_{600nm} measurement and for filtration in multiscreen plates (Millipore) to originate cell free fluids. These supernatants were

stored at -20 °C, overnight. For AI-2 quantification, *in vitro* response of CLPY (LuxP fused to a cyan and to a yellow fluorescent protein) was measured as previously described (51,157) using a multilabel counter (1420 Victor 3, Perkin Elmer). 2.5 µL of the collected samples were added to 280 µL of 12.5 µg/mL of CLPY in 25 mM of sodium phosphate buffer pH 8.0, 35 mM NaCl and 1 mM boric acid. The Förster resonance energy transfer (FRET) ratio was calculated (535/440 nm). To determine the concentration of AI-2 in the cell free fluids, the obtained results were compared against a calibration curve obtained with different AI-2 concentrations. Error bars are derived from three technical replicates.

2.7.3. Tables

Table S2.1 - Composition of the liquid growth medium employed.

Modified PY+X with arabinose and 100 mM MOPS buffer, pH 7.0.

Component	Quantity
Trypticase peptone.....	5 g/L
Peptone from meat (pepsin digested).....	5 g/L
Yeast extract.....	10 g/L
Arabinose.....	5 g/L
L-Cysteine-HCl x H ₂ O.....	0.50 g/L
0.5 M MOPS buffer pH 7.0 ^a	200 mL/L
Salt solution.....	20 mL/L
CaCl ₂ x 2 H ₂ O.....	0.25 g/L
MgSO ₄ x 7 H ₂ O.....	0.50 g/L
K ₂ HPO ₄	1 g/L
KH ₂ PO ₄	1 g/L
NaHCO ₃	10 g/L
NaCl.....	2 g/L

^aMOPS buffer was filter sterilized and added to the autoclaved medium.

Agar plates (Modified PY+X with galactose)

Composition of the solid medium was the same as in table S2.1 with the addition of 15 g/L of agar and of 5 g/L of galactose instead of arabinose. No MOPS buffer was added to this solid medium.

Table S2.2 - Sequence identity of the Lsr orthologs and LuxS present in *C. saccharobutylicum* (*C. sacch*) and *C. autoethanogenum* (*C. auto*) when compared with the ones from previously studied organisms.

Organism	KEGG identifier	Protein	%identity <i>C. sacch</i>	KEGG identifier	%identity <i>C. auto</i>	KEGG identifier
<i>Bacillus cereus</i> (bca)	BCE_3012	LsrB	39.2	CLSA_c21430	37.9	CAETHG_4003
	BCE_3015	LsrA	46.3	CLSA_c21400	47.1	CAETHG_4000
	BCE_3014	LsrC	43.9	CLSA_c21410	43.1	CAETHG_4001
	BCE_3013	LsrD	40.9	CLSA_c21420	46.7	CAETHG_4002
	BCE_3016	LsrR	39.7	CLSA_c21390	40.7	CAETHG_3999
	BCE_3017	LsrK	51.7	CLSA_c21380	52.9	CAETHG_3998
	BCE_3019	LsrF	-	-	40.2	CAETHG_4004
	BCE_3074	LsrG	33.3	CLSA_c11720	31.9	CAETHG_0629
	BCE_4946	LuxS	51.8	CLSA_c30740	30.2	CAETHG_0412
<i>Bacillus anthracis</i> Sterne (151)	BAS2763	LsrB	39.2	CLSA_c21430	37.9	CAETHG_4003
	BAS2766	LsrA	46.0	CLSA_c21400	46.3	CAETHG_4000
	BAS2765	LsrC	42.4	CLSA_c21410	44.0	CAETHG_4001
	BAS2764	LsrD	43.4	CLSA_c21420	46.1	CAETHG_4002
	BAS2767	LsrR	41.1	CLSA_c21390	41.0	CAETHG_3999
	BAS2771	LsrF	-	-	40.5	CAETHG_4004
	BAS2825	LsrG	32.0	CLSA_c11720	29.0	CAETHG_0629
	BAS4687	LuxS	51.8	CLSA_c30740	30.2	CAETHG_0412
<i>Escherichia coli</i> K-12 MG1655 (eco)	b1516	LsrB	43.6	CLSA_c21430	40.2	CAETHG_4003
	b2687	LuxS	42.8	CLSA_c30740	32.5	CAETHG_0412
<i>Salmonella</i> serovar Typhimurium LT2 (stm)	STM4077	LsrB	40.1	CLSA_c21430	39.8	CAETHG_4003
	STM2817	LuxS	41.4	CLSA_c30740	33.1	CAETHG_0412

Table S2.3 - Data collection and refinement statistics.

Wavelength	1181
Resolution range	50.59 - 1.368 (1.417 - 1.368)
Space group	P 43 2 2
Unit cell	90.78 90.78 182.81 90 90 90
Total reflections	1150236 (54053)
Unique reflections	160866 (7881)
Multiplicity	7.1 (6.9)
Completeness (%)	100.0 (100.0)
Mean I/sigma(I)	14.4 (1.6)
R-merge	0.056 (0.819)
R-meas	0.065 (0.896)
R-pim	0.033 (0.509)
CC1/2	0.999 (0.731)
Reflections used in refinement	160771 (15892)
Reflections used for R-free	2015 (202)
R-work	0.1599 (0.2199)
R-free	0.1765 (0.2271)
CC(work)	0.879 (0.764)
CC(free)	0.889 (0.750)
Number of non-hydrogen atoms	5327
macromolecules	4948
ligands	20
solvent	359
Protein residues	652
RMS(bonds)	10
RMS(angles)	1.07
Ramachandran favored (%)	96.76
Ramachandran allowed (%)	2.93
Ramachandran outliers (%)	0.31
Rotamer outliers (%)	0.19
Average B-factor	19.77
macromolecules	19.50
ligands	14.42
solvent	23.79

Statistics for the highest-resolution shell are shown in parentheses.

Chapter 2

Table S2.4a. Identification of putative non-canonical LsrB receptors. Identity of Lsr homologs from the organisms with putative non-canonical LsrB receptors, corresponding sequence identity to *C. saccharobutylicum* (*C. sacch*) proteins and 3-dimensional fold prediction by Phyre2 are shown. For the sequences in bold no Lsr homologues were identified in the nearby proteome and a blast against the indicated proteome was performed to identify homologues.

Phylum	Organism	LsrB				LsrA				LsrC				LsrD				Phyre2 hits (pdb code)
		NCBI protein reference	Query cover	% ID C. sacch	Type of putative LsrB	% ID C. sacch	Query cover	NCBI protein reference	% ID C. sacch	Query cover	NCBI protein reference	% ID C. sacch	Query cover	NCBI protein reference	% ID C. sacch	Query cover		
Actinobacteria	<i>Olsenella</i> sp. KH3B4	WP_091004526.1	88%	67%	2 subs (KDDQNS)	57%	97%	WP_091004519.1	54%	96%	WP_091004521.1	54%	95%	WP_091004523.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Olsenella umbonata</i> KHGC19	WP_091008946.1	88%	67%	2 subs (KDDQNS)	56%	99%	WP_078687486.1	55%	99%	WP_091008949.1	49%	92%	WP_078687488.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Olsenella</i> sp. DNF00959	KXB61513.1	95%	64%	2 subs (KDDQNS)	59%	95%	KXB61510.1	56%	99%	KXB61511.1	49%	95%	KXB61512.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Olsenella urininfantis</i> Marseille-P3197	WP_077598408.1	95%	63%	2 subs (KDDQNS)	58%	95%	WP_077598790.1	55%	97%	WP_077598409.1	49%	96%	WP_097677390.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Atopobium</i> sp. ICM42b	WP_035428383.1	88%	67%	2 subs (KDDQNS)	59%	97%	WP_035428388.1	52%	97%	WP_035428387.1	52%	93%	WP_035428385.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Atopobium</i> sp. BS2	WP_035437357.1	95%	64%	2 subs (KDDQNS)	59%	97%	WP_084055005.1	53%	97%	WP_035437352.1	51%	94%	WP_035437355.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Atopobium</i> sp. oral taxon 199	WP_016477588.1	97%	63%	2 subs (KDDQNS)	59%	95%	WP_016477591.1	52%	100%	WP_016477590.1	51%	95%	WP_016477589.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Atopobium deltae</i> DNF00019	KXB35017.1	99%	60%	2 subs (KDDQNS)	56%	95%	KXB35020.1	54%	96%	KXB35019.1	51%	95%	KXB35018.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Libanicoccus massiliensis</i>	WP_083580771.1	88%	66%	2 subs (KDDQNS)	61%	98%	WP_083580772.1	55%	98%	WP_073296587.1	52%	98%	WP_073296584.1	1 st : 4pz0; 2 nd : 1ty			
Bacteroidetes ^a	<i>Bacteroides xyloxyticus</i> DSM 3808	PPK78747.1	88%	69%	2 subs (KDDQNS)	60%	96%	PPK78750.1	53%	97%	PPK78749.1	59%	90%	PPK78748.1	1 st : 4pz0; 2 nd : 1ty			
Firmicutes	<i>Clostridium uliginosum</i> DSM 12992	WP_090088678.1	95%	92%	2 subs (KDDQNS)	85%	100%	WP_090088683.1	91%	100%	WP_090088681.1	90%	98%	WP_090088680.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. BL-8	WP_077861112.1	100%	88%	2 subs (KDDQNS)	85%	100%	WP_077861115.1	88%	100%	WP_077861114.1	87%	98%	WP_077861113.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium butyricum</i> AGR2140	WP_027635215.1	100%	87%	2 subs (KDDQNS)	82%	100%	WP_035771860.1	85%	100%	WP_027635216.1	86%	98%	WP_051238861.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium ragsdalei</i> P11	WP_065077960.1	99%	80%	2 subs (KDDQNS)	78%	99%	WP_065078018.1	73%	99%	WP_065078017.1	76%	97%	WP_065077961.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridiales bacterium</i> Marseille-P2846	WP_102412252.1	96%	69%	2 subs (KDDQNS)	66%	96%	WP_102413356.1	62%	92%	WP_102413355.1	60%	95%	WP_102413354.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Marasmitruncus massiliensis</i> Marseille-P3646	WP_101911019.1	99%	64%	2 subs (KDDQNS)	69%	96%	WP_101911692.1	62%	97%	WP_101911021.1	64%	96%	WP_101911020.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Desulfotomaculum guttoideum</i> DSM 4024	WP_092240434.1	99%	64%	2 subs (KDDQNS)	63%	96%	WP_092240436.1	54%	97%	WP_025231381.1	60%	93%	WP_092240435.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. Marseille-P2415	WP_077612143.1	99%	62%	2 subs (KDDQNS)	62%	96%	WP_077612146.1	54%	97%	WP_077612145.1	60%	95%	WP_077612144.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Roseburia inulinivorans</i> (multispecies)	WP_021923372.1	100%	61%	2 subs (KDDQNS)	53%	96%	WP_055040579.1	49%	98%	WP_021923374.1	51%	97%	WP_021923373.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium celeraceres</i> 18A	WP_100306014.1	99%	63%	2 subs (KDDQNS)	63%	96%	WP_100306017.1	54%	97%	WP_100306016.1	60%	94%	WP_100306015.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium sphenoides</i> JCM 1415	WP_054792150.1	99%	63%	2 subs (KDDQNS)	63%	96%	WP_100043652.1	54%	97%	WP_100043653.1	60%	93%	WP_025231382.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. ASB410	WP_025231383.1	99%	63%	2 subs (KDDQNS)	63%	96%	WP_025231380.1	54%	97%	WP_025231381.1	60%	93%	WP_025231382.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridiales bacterium</i> 1_7_47FAA	EEQ58208.1	89%	69%	2 subs (KDDQNS)	63%	96%	EEQ58211.1	54%	98%	EEQ58210.1	60%	95%	EEQ58209.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium clostridioforme</i>	WP_057571018.1	89%	70%	2 subs (KDDQNS)	63%	98%	WP_057571015.1	54%	98%	WP_057571016.1	60%	91%	WP_057571017.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. W14A	WP_066644543.1	99%	63%	2 subs (KDDQNS)	65%	95%	WP_066644536.1	54%	98%	WP_066644539.1	58%	90%	WP_066644541.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium jeddahense</i> JCD	WP_042436285.1	99%	63%	2 subs (KDDQNS)	65%	96%	WP_009060578.1	54%	98%	WP_042436287.1	59%	87%	WP_009060602.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Lachnospirillum</i> YL32	WP_065551070.1	89%	69%	2 subs (KDDQNS)	63%	98%	WP_065551067.1	54%	98%	WP_065551068.1	61%	91%	WP_065551069.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. C105KSO14	WP_089776431.1	89%	69%	2 subs (KDDQNS)	63%	98%	WP_057571015.1	54%	98%	WP_089776433.1	61%	91%	WP_089776432.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium clostridioforme</i> AGR2157	WP_027640572.1	89%	69%	2 subs (KDDQNS)	63%	98%	WP_027640569.1	54%	98%	WP_027640570.1	61%	91%	WP_027640571.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. MSTE9	WP_009060716.1	99%	62%	2 subs (KDDQNS)	65%	96%	WP_009060578.1	54%	98%	WP_009060703.1	59%	89%	WP_009060602.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium amygdalinum</i>	WP_097004796.1	99%	62%	2 subs (KDDQNS)	61%	95%	WP_097004799.1	54%	97%	WP_097004798.1	57%	95%	WP_097004797.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. FS41	WP_045093313.1	88%	69%	2 subs (KDDQNS)	63%	96%	WP_007863866.1	54%	98%	WP_007863868.1	61%	91%	WP_007863869.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium citroniae</i> WAL-19142	WP_048930652.1	88%	69%	2 subs (KDDQNS)	63%	96%	WP_007863866.1	54%	98%	WP_007863868.1	61%	91%	WP_007863869.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Clostridium</i> sp. AT4	WP_081422794.1	89%	69%	2 subs (KDDQNS)	62%	96%	WP_066543004.1	54%	99%	WP_066543003.1	59%	95%	WP_066543002.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Lachnospiraceae bacterium</i> KA00044	WP_102905567.1	100%	59%	2 subs (KDDQNS)	60%	96%	WP_102905573.1	54%	97%	WP_102905571.1	61%	86%	WP_102905569.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Lachnospiraceae bacterium</i> Jc7	WP_044995901.1	100%	58%	2 subs (KDDQNS)	62%	95%	WP_044995906.1	51%	98%	WP_044995904.1	55%	95%	WP_044995903.1	1 st : 4pz0; 2 nd : 1ty			
	<i>Kallipyga massiliensis</i> ph2	WP_019133810.1	100%	62%	2 subs (KDDQNS)	59%	96%	WP_019133807.1	52%	96%	WP_019133808.1	52%	98%	WP_019133809.1	1 st : 4pz0; 2 nd : 1ty			
<i>Oribacterium</i> sp. C9	WP_078038379.1	89%	67%	2 subs (KDDQNS)	62%	94%	WP_078038382.1	51%	98%	WP_078038381.1	55%	95%	WP_078038380.1	1 st : 4pz0; 2 nd : 1ty				
<i>Oribacterium</i> sp. oral taxon 108	WP_009429067.1	99%	54%	2 subs (KDDQNS)	59%	96%	WP_083816317.1	46%	97%	WP_009429048.1	50%	91%	WP_009429585.1	1 st : 4pz0; 2 nd : 1ty				
<i>Oribacterium asaccharolyticum</i> ACB7	WP_009537119.1	99%	54%	2 subs (KDDQNS)	59%	96%	WP_009537122.1	46%	97%	WP_009537121.1	50%	91%	WP_009537120.1	1 st : 4pz0; 2 nd : 1ty				
<i>Anaerocolumna aminovaleric</i> a DSM1283	WP_091683830.1	100%	62%	2 subs (KDDQNS)	63%	98%	WP_091684120.1	59%	96%	WP_091683834.1	62%	97%	WP_091683832.1	1 st : 4pz0; 2 nd : 1ty				
<i>Clostridium minihomine</i> Marseille-P4642	WP_101697108.1	99%	63%	2 subs (KDDQNS)	65%	97%	WP_101697111.1	53%	98%	WP_101697110.1	57%	98%	WP_101697109.1	1 st : 4pz0; 2 nd : 1ty				
<i>Cellulosilyticum</i> sp. I15G1012	WP_069998611.1	100%	61%	2 subs (KDDQNS)	64%	95%	WP_083258605.1	61%	96%	WP_069998613.1	62%	94%	WP_069998612.1	1 st : 4pz0; 2 nd : 1ty				
<i>Clostridiales bacterium</i> KA00274	WP_066712194.1	88%	66%	2 subs (KDDQNS)	52%	94%	WP_082714243.1	50%	90%	WP_082714244.1	58%	85%	WP_066712192.1	1 st : 4pz0; 2 nd : 1ty				
<i>Pygmiaobacter massiliensis</i> Marseille-P3336	WP_102050943.1	88%	67%	2 subs (KDDQNS)	63%	95%	WP_102050940.1	56%	98%	WP_102050941.1	52%	94%	WP_102050942.1	1 st : 4pz0; 2 nd : 1ty				
<i>Oscillibacter ruminantium</i> GH1	WP_040659548.1	88%	67%	2 subs (KDDQNS)	62%	96%	WP_040659551.1	53%	99%	WP_040659552.1	56%	87%	WP_040659550.1	1 st : 4pz0; 2 nd : 1ty				
<i>Anaerotruncus</i> sp. G3	WP_016314928.1	88%	66%	2 subs (KDDQNS)	62%	96%	WP_022788191.1	54%	96%	WP_016314926.1	57%	87%	WP_016314927.1	1 st : 4pz0; 2 nd : 1ty				

LsrB-like receptors in Clostridia

	<i>Clostridium sporosphaeroides</i> DSM 1294	WP_020074139.1	99%	62%	2 subs (KDDQNS)	64%	96%	WP_020074136.1	53%	98%	WP_020074137.1	60%	84%	WP_020074138.1	1 st ; 4pz0; 2 nd 1ty
	<i>Clostridiales bacterium</i> VE202-13	WP_024729909.1	88%	65%	2 subs (KDDQNS)	64%	91%	WP_081726327.1	54%	96%	WP_006872978.1	54%	99%	WP_006872977.1	1 st ; 4pz0; 2 nd 1ty
	<i>Anaerotruncus colihominis</i> DSM 17241	WP_006872976.1	88%	65%	2 subs (KDDQNS)	64%	91%	WP_081726327.1	54%	96%	WP_006872978.1	54%	99%	WP_006872977.1	1 st ; 4pz0; 2 nd 1ty
	<i>Firmicutes bacterium</i> HGW-Firmicutes-7	PKM52094.1	100%	60%	2 subs (KDDQNS)	64%	97%	PKM52153.1	61%	96%	PKM52092.1	64%	93%	PKM52093.1	1 st ; 4pz0; 2 nd 1ty
	<i>Clostridium merdae</i> Marseille-P2953	WP_085834219.1	99%	61%	2 subs (KDDQNS)	62%	97%	WP_085834222.1	53%	98%	WP_085834221.1	56%	96%	WP_085834220.1	1 st ; 4pz0; 2 nd 1ty
	<i>Eubacterium</i> sp. 14-2	WP_031200912.1	92%	64%	2 subs (KDDQNS)	65%	96%	WP_022750948.1	55%	98%	WP_016213796.1	61%	83%	WP_016213797.1	1 st ; 4pz0; 2 nd 1ty
	<i>Hespelia stercoris</i> DSM 15480	WP_073104636.1	100%	59%	2 subs (KDDQNS)	63%	96%	WP_073104627.1	58%	97%	WP_073104630.1	60%	93%	WP_073104633.1	1 st ; 4pz0; 2 nd 1ty
	<i>Streptococcus fluxtorum</i> DSM 22810	WP_037622845.1	100%	59%	2 subs (KDDQNS)	56%	96%	WP_027972160.1	51%	97%	WP_027972159.1	57%	87%	WP_027972158.1	1 st ; 4pz0; 2 nd 1ty
	<i>Streptococcus</i> sp. oral taxon 056	WP_009755053.1	88%	64%	2 subs (KDDQNS)	57%	95%	WP_009443569.1	46%	98%	WP_009443426.1	58%	89%	WP_009754658.1	1 st ; 4pz0; 2 nd 1ty
	<i>Hydrogenoanaerobacterium saccharovorans</i> CGMCC	WP_092753797.1	89%	67%	2 subs (KDDQNS)	61%	96%	WP_092753803.1	53%	98%	WP_092753801.1	55%	94%	WP_092753799.1	1 st ; 4pz0; 2 nd 1ty
	<i>Granulicatella</i> (multispecies)	WP_005605806.1	97%	60%	2 subs (KDDQNS)	56%	96%	WP_005605810.1	48%	99%	WP_005605809.1	55%	87%	WP_005605808.1	1 st ; 4pz0; 2 nd 1ty
	<i>Anaerotruncus neopropionum</i> DSM-3847 ^b	WP_066083873.1	100%	59%	2 subs (KDDQNS)	59%	97%	WP_066083867.1	54%	99%	WP_066083869.1	54%	97%	WP_066083870.1	1 st ; 4pz0; 2 nd 1ty
	<i>Acetanaerobacterium elongatum</i> CGMCC	WP_092638660.1	99%	58%	2 subs (KDDQNS)	61%	96%	WP_092642787.1	54%	96%	WP_092642790.1	55%	94%	WP_092642793.1	1 st ; 4pz0; 2 nd 1ty
	<i>Firmicutes bacterium</i> HGW-Firmicutes-16	PKM72027.1	89%	62%	2 subs (KDDQNS)	60%	96%	PKM72024.1	49%	95%	PKM72025.1	52%	97%	PKM72026.1	1 st ; 4pz0; 2 nd 1ty
	<i>Firmicutes bacterium</i> HGW-Firmicutes-9	PKM39249.1	89%	60%	2 subs (KDDQNS)	59%	99%	PKM39252.1	47%	97%	PKM39251.1	54%	90%	PKM39250.1	1 st ; 4pz0; 2 nd 1ty
	<i>Anaerovirgula multivorans</i> SCA	WP_089285157.1	100%	52%	1 subs (KDDQPS)	57%	97%	WP_089285160.1	48%	97%	WP_089285159.1	49%	95%	WP_089285158.1	1 st ; 4pz0; 2 nd 1ty
	<i>Fusibacter</i> sp. 3D3	WP_069871715.1	100%	49%	1 subs (KDDQPS)	58%	95%	WP_069871711.1	48%	97%	WP_069871713.1	50%	93%	WP_069872049.1	1 st ; 4pz0; 2 nd 1ty
Firmicutes	<i>Bacillus mesonae</i> FJAT-13985	WP_084798056.1	98%	50%	1 subs (KDDQPS)	52%	94%	WP_084798058.1	48%	84%	WP_084798057.1	43%	93%	WP_066388750.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus</i> sp. AFS073361	WP_098572761.1	98%	48%	1 subs (KDDQPS)	51%	96%	WP_098572749.1	47%	95%	WP_098572748.1	48%	89%	WP_098572747.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus</i> sp. OG2	WP_094768697.1	96%	51%	1 subs (KDDQPS)	51%	94%	WP_094768654.1	43%	95%	WP_094768653.1	45%	94%	WP_094768698.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus</i> sp. AFS037270	WP_098996177.1	96%	49%	1 subs (KDDQPS)	52%	95%	WP_098996050.1	47%	95%	WP_098996051.1	47%	95%	WP_098996052.1	1 st ; 4pz0; 2 nd 1ty
	<i>Salirhabdus</i> sp. Marseille-P4669	WP_102026542.1	100%	47%	1 subs (KDDQPS)	52%	98%	WP_102026540.1	46%	83%	WP_102026715.1	49%	93%	WP_102026541.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus</i> sp. AFS076308	WP_098532648.1	96%	49%	1 subs (KDDQPS)	52%	95%	WP_098532581.1	46%	95%	WP_098532582.1	50%	86%	WP_098532583.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus</i> sp. UNC41MF55	WP_081825476.1	99%	48%	1 subs (KDDQPS)	52%	95%	WP_026565767.1	47%	95%	WP_051633641.1	48%	95%	WP_026565778.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus onubensis</i>	WP_099361000.1	100%	48%	1 subs (KDDQPS)	54%	99%	WP_099360999.1	46%	97%	WP_099361167.1	51%	86%	WP_099361168.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus drentensis</i> NBRC 102427	WP_066250480.1	99%	48%	1 subs (KDDQPS)	52%	96%	WP_066250473.1	47%	94%	WP_066250475.1	50%	86%	WP_066250477.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus</i> sp. MRMR6	WP_075689723.1	99%	46%	1 subs (KDDQPS)	52%	97%	WP_075689726.1	47%	95%	WP_075689725.1	50%	88%	WP_075689724.1	1 st ; 4pz0; 2 nd 1ty
	<i>Geosporobacter ferrireducens</i> IRF9	WP_069979555.1	100%	47%	1 subs (KDDQPS)	57%	97%	WP_069979552.1	47%	96%	WP_069981526.1	51%	94%	WP_069979554.1	1 st ; 4pz0; 2 nd 1ty
	<i>Oceanobacillus damuensis</i> PT-20	WP_067729926.1	100%	47%	1 subs (KDDQPS)	52%	95%	WP_084268847.1	43%	95%	WP_084268845.1	48%	93%	WP_067729984.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus novalis</i> NBRC 102450	WP_066087728.1	99%	48%	1 subs (KDDQPS)	52%	96%	WP_066087742.1	47%	95%	WP_066087737.1	48%	94%	WP_066087732.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus massiliogabonensis</i> Marseille P2639	WP_102273483.1	100%	49%	1 subs (KDDQPS)	53%	98%	WP_102273480.1	47%	95%	WP_102273481.1	49%	86%	WP_102273482.1	1 st ; 4pz0; 2 nd 1ty
	<i>Bacillus cihuenis</i> FJAT-14515	WP_028391996.1	100%	48%	1 subs (KDDQPS)	53%	97%	WP_034301160.1	46%	97%	WP_034301157.1	47%	93%	WP_028391997.1	1 st ; 4pz0; 2 nd 1ty
<i>Paenibacillus</i> sp. 7884-2	WP_095313828.1	95%	48%	1 subs (KDDQPS)	51%	96%	WP_095313824.1	44%	95%	WP_095313825.1	48%	89%	WP_095313826.1	1 st ; 4pz0; 2 nd 1ty	
<i>Bacillus praedii</i> FJAT-25547	WP_057759881.1	100%	47%	1 subs (KDDQPS)	52%	98%	WP_057759872.1	48%	89%	WP_095396672.1	50%	86%	WP_057759878.1	1 st ; 4pz0; 2 nd 1ty	
<i>Bacillus</i> sp. FJAT-21945	WP_053478418.1	100%	48%	1 subs (KDDQPS)	53%	96%	WP_053478415.1	47%	95%	WP_053478416.1	50%	86%	WP_053478417.1	1 st ; 4pz0; 2 nd 1ty	
<i>Virgibacillus indicus</i> P2-C2	WP_094887270.1	100%	48%	1 subs (KDDQPS)	51%	98%	WP_094887267.1	44%	97%	WP_094887268.1	51%	85%	WP_094887269.1	1 st ; 4pz0; 2 nd 1ty	
<i>Proteinilactium ruminis</i> ML2	WP_074912245.1	99%	45%	1 subs (KDDQPS)	52%	94%	WP_083422751.1	47%	96%	WP_074912286.1	48%	93%	WP_074912246.1	1 st ; 4pz0; 2 nd 1ty	
Proteobacteria	<i>Succinobacillus dextrinosolvens</i> 22B	WP_074840785.1	100%	59%	2 subs (KDDQNS)	56%	98%	WP_074840782.1	53%	94%	WP_074840783.1	54%	95%	WP_074840784.1	1 st ; 4pz0; 2 nd 1ty
	<i>Deltaproteobacteria bacterium</i> RBG_13_49_15	OGP59397.1	99%	41%	1 subs (KDDQPS)	40%	96%	OGP59394.1	37%	96%	OGP59395.1	38%	90%	OGP59396.1	1 st ; 4pz0; 2 nd 1ty
	<i>Treponema socranskii</i> ATCC 35536	WP_021330428.1	97%	65%	2 subs (KDDQNS)	57%	96%	WP_021330444.1	54%	94%	WP_021330436.1	61%	90%	WP_021330450.1	1 st ; 4pz0; 2 nd 1ty
	<i>Treponema</i> sp. OMZ 838	WP_044014762.1	97%	64%	2 subs (KDDQNS)	58%	97%	WP_044014768.1	54%	96%	WP_044014766.1	61%	88%	WP_044014764.1	1 st ; 4pz0; 2 nd 1ty
	<i>Treponema medium</i> ATCC 700293	WP_016522510.1	97%	64%	2 subs (KDDQNS)	57%	97%	WP_016522507.1	53%	96%	WP_016522508.1	61%	88%	WP_016522509.1	1 st ; 4pz0; 2 nd 1ty
	<i>Treponema</i> sp. GWA1_62_8	OHE58488.1	99%	62%	2 subs (KDDQNS)	61%	95%	OHE58487.1	similar to WP_008719214.1; too many ambiguous residues	60%	94%	OHE58489.1	1 st ; 4pz0; 2 nd 1ty		
Spirochaetes	<i>Treponema primitia</i> ZAS-1	WP_010258002.1	100%	56%	2 subs (KDDQNS)	36%	94%	WP_010254545.1	32%	90%	WP_010263554.1	32%	86%	WP_010259556.1	1 st ; 4pz0; 2 nd 1ty
	<i>Treponema azotonutricium</i> ZAS-9	WP_015711767.1	100%	55%	2 subs (KDDQNS)	37%	94%	WP_015710032.1	32%	93%	WP_015709735.1	33%	87%	WP_015711890.1	1 st ; 4pz0; 2 nd 1ty
	<i>Brachyspira</i> sp. CAG:700	CCY74820.1	98%	42%	1 subs (KDDQPS)	46%	94%	CCY74825.1	33%	95%	CCY74824.1	41%	86%	CCY74822.1	1 st ; 4pz0; 2 nd 1ty
	<i>Sediminispirochaeta bajacaliforniensis</i> DSM16054	WP_013256124.1	90%	44%	1 subs (KDDQPS)	45%	95%	WP_020610755.1	37%	89%	WP_020610756.1	36%	96%	WP_020610757.1	1 st ; 4pz0; 2 nd 1ty
	<i>Sediminispirochaeta smaragdinae</i> DSM11293	WP_013256124.1	90%	44%	1 subs (KDDQPS)	45%	95%	WP_013256127.1	37%	89%	WP_013256126.1	36%	96%	WP_013256125.1	1 st ; 4pz0; 2 nd 1ty

^a This bacterial species is classified as a Bacteroidetes, but recent information indicates that it might be a Firmicute.

^b Previously annotated as *Clostridium neopropionum*

Chapter 2

Table S2.4b. Identification of putative non-canonical LsrB receptors. Identity of Lsr homologs from the organisms with putative non-canonical LsrB receptors and corresponding sequence identity to *C. saccharobutylicum* (*C. sach*) proteins. For the sequences in bold no Lsr homologues were identified in the nearby proteome and a blast against the indicated proteome was performed to identify homologues.

Phylum	Organism	LsrK			LsrK			LsrF ^a			LsrG ^d			LsrG ^d			LuxS ^d		
		% ID <i>C. sach</i>	Query cover	NCBI protein reference	% ID <i>C. sach</i>	Query cover	NCBI protein reference	% ID <i>E. coli</i> ^a	Query cover	NCBI protein reference	% ID <i>C. sach</i>	Query cover	NCBI protein reference	% ID <i>E. coli</i>	Query cover	NCBI protein reference	% ID <i>C. sach</i>	Query cover	NCBI protein reference
Actinobacteria	<i>Olsenella</i> sp. KH3B4	47%	96%	WP_091004518.1	68%	99%	WP_091004516.1	51%	94%	WP_091004528.1	No significant similarity found	No significant similarity found	No significant similarity found	39%	90%	WP_091003972.1			
	<i>Olsenella umbonata</i> KHGC19	47%	95%	WP_078687485.1	68%	100%	WP_091009024.1	52%	97%	WP_091009024.1	No significant similarity found	No significant similarity found	No significant similarity found	38%	92%	WP_09084785.1			
	<i>Olsenella</i> sp. DNF0099	45%	95%	KXB61509.1	65%	100%	KXB61508.1	53%	98%	KXB61514.1	No significant similarity found	29%	32%	WP_062531789.1	39%	91%	WP_00927862.1		
	<i>Olsenella urinifantis</i> Marseille-P3197	44%	94%	WP_077598410.1	65%	100%	WP_077598407.1	53%	98%	WP_077598407.1	No significant similarity found	No significant similarity found	No significant similarity found	40%	90%	WP_077597510.1			
	<i>Atopobium</i> sp. ICM42b	48%	93%	WP_035428390.1	65%	100%	WP_035428392.1	51%	98%	WP_035428381.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found			No significant similarity found		
	<i>Atopobium</i> sp. B52	48%	93%	WP_035437349.1	65%	100%	WP_035437344.1	51%	98%	WP_035428381.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found			No significant similarity found		
	<i>Atopobium</i> sp. oral taxon 199	49%	93%	WP_016477592.1	66%	100%	WP_016477593.1	53%	98%	WP_016477587.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	43%	92%	WP_016477437.1		
	<i>Atopobium deltae</i> DNF00019	45%	96%	KXB35022.1	62%	99%	KXB35023.1	50%	98%	KXB35016.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	35%	91%	KXB32942.1		
	<i>Libanococcus massiliensis</i>	47%	98%	WP_073296590.1	68%	100%	WP_073296593.1	50%	98%	WP_073296582.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	37%	94%	WP_073295068.1		
	Bacteroidetes <i>Bacteroides xyloxyticus</i> DSM 3808	58%	95%	PPK78751.1	67%	99%	PPK78752.1	51%	98%	PPK78746.1	42%	100%	PPK81823.1	No significant similarity found	37%	90%	WP_104433461.1		
Firmicutes	<i>Clostridium uliginosum</i> DSM 12992	84%	100%	WP_090088684.1	88%	100%	WP_090088686.1	52%	98%	WP_090088676.1	57%	100%	WP_090088139.1	No significant similarity found	89%	98%	WP_090090700.1		
	<i>Clostridium</i> sp. BL-8	84%	99%	WP_077861116.1	83%	100%	WP_077861117.1	53%	98%	WP_077861111.1	69%	100%	WP_077859727.1	26%	77%	WP_077859727.1	38%	90%	WP_009169589.1
	<i>Clostridium butyricum</i> AGR2140	87%	96%	WP_027635218.1	80%	100%	WP_027635219.1	52%	98%	WP_027635214.1	No significant similarity found	40%	26%	WP_071982258.1	38%	89%	WP_002581777.1		
	<i>Clostridium ragsdalei</i> P11	74%	96%	WP_065077962.1	81%	100%	WP_065077963.1	51%	98%	WP_065077959.1	No significant similarity found	33%	31%	WP_065077455.1	37%	90%	WP_063556345.1		
	<i>Clostridiales bacterium</i> Marseille-P2846	58%	98%	WP_102412253.1	68%	99%	WP_102412254.1	52%	98%	WP_102412251.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	38%	83%	WP_102411859.1		
	<i>Marasmitruncus massiliensis</i> Marseille-P3646	63%	98%	WP_101911022.1	75%	99%	WP_101911013.1	52%	98%	WP_101911018.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	34%	86%	WP_101910466.1		
	<i>Desulfotomaculum guttoideum</i> DSM 4024	55%	97%	WP_092240437.1	68%	99%	WP_092240438.1	52%	98%	WP_092240433.1	49%	100%	WP_025233884.1	No significant similarity found	37%	92%	WP_092246276.1		
	<i>Clostridium</i> sp. Marseille-P2415	55%	97%	WP_077612147.1	68%	99%	WP_077612148.1	52%	98%	WP_077612142.1	40%	94%	WP_077611464.1	No significant similarity found	37%	92%	WP_077613479.1		
	<i>Roseburia inulinivorans</i> (multispecies)	49%	99%	WP_021923376.1	69%	100%	WP_055040580.1	51%	98%	WP_021923371.1	No significant similarity found	24%	34%	WP_021923823.1	40%	90%	WP_007884105.1		
	<i>Clostridium celerescens</i> 18A	55%	97%	WP_100306018.1	68%	99%	WP_100306019.1	52%	98%	WP_100306013.1	48%	100%	WP_100303915.1	41%	17%	WP_100305011.1	37%	92%	WP_100306838.1
	<i>Clostridium sphenoides</i> JCM 1415	55%	97%	WP_054792151.1	68%	99%	WP_054792152.1	52%	98%	WP_092240433.1	48%	100%	WP_100042414.1	No significant similarity found	37%	92%	WP_092246276.1		
	<i>Clostridium</i> sp. ASB6410	56%	99%	WP_025231379.1	68%	99%	WP_025231378.1	52%	98%	WP_025231384.1	49%	100%	WP_025233884.1	No significant similarity found	37%	92%	WP_025230413.1		
	<i>Clostridiales bacterium</i> 1_7_47FAA	58%	95%	EEQ58212.1	70%	99%	EEQ58213.1	53%	98%	EEQ58207.1	29%	68%	WP_008716261.1	42%	12%	EEQ61402.1	No significant similarity found		
	<i>Clostridium clostridioforme</i>	58%	95%	WP_027640568.1	67%	99%	WP_027640567.1	52%	98%	WP_027640573.1	30%	38%	WP_002590493.1	No significant similarity found	40%	98%	WP_027641561.1		
	<i>Clostridium</i> sp. W14A	60%	95%	WP_066644531.1	71%	99%	WP_066644530.1	55%	98%	WP_066644560.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	35%	90%	WP_066644817.1		
	<i>Clostridium jeddahense</i> JCD	58%	96%	WP_009060540.1	71%	99%	WP_042436289.1	52%	98%	WP_027640573.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	53%	95%	WP_042433997.1		
	<i>Lachnospirillum</i> YL32	58%	95%	WP_027640568.1	68%	99%	WP_065551066.1	52%	98%	WP_027640573.1	31%	36%	WP_065551012.1	No significant similarity found	No significant similarity found	No significant similarity found			
	<i>Clostridium</i> sp. C105KSO14	58%	95%	WP_027640568.1	67%	99%	WP_027640567.1	52%	98%	WP_089776430.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	40%	98%	WP_022203069.1		
	<i>Clostridium clostridioforme</i> AGR2157	58%	95%	WP_027640568.1	67%	99%	WP_027640567.1	52%	98%	WP_027640573.1	33%	28%	WP_027640305.1	No significant similarity found	40%	98%	WP_027641561.1		
	<i>Clostridium</i> sp. MSTE9	58%	96%	WP_009060540.1	70%	99%	WP_009060607.1	50%	98%	WP_035316427.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	54%	95%	WP_009060984.1		
	<i>Clostridium amygdalinum</i>	59%	95%	WP_097004800.1	67%	99%	WP_097004801.1	51%	98%	WP_097004795.1	49%	100%	WP_097003004.1	No significant similarity found	36%	91%	WP_097005421.1		
	<i>Clostridium</i> sp. FS41	57%	95%	WP_045093314.1	70%	99%	WP_045093315.1	53%	98%	WP_045093312.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	38%	96%	WP_045091419.1		
	<i>Clostridium citroniae</i> WAI-19142	57%	95%	WP_007863861.1	69%	99%	WP_048930651.1	53%	98%	WP_045093312.1	31%	40%	WP_002572755.1	No significant similarity found	39%	96%	WP_007862026.1		
	<i>Clostridium</i> sp. AT4	56%	94%	WP_066543005.1	70%	100%	WP_066543007.1	49%	98%	WP_066542998.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	38%	98%	WP_066544852.1		
	<i>Lachnospiraceae bacterium</i> KA00044	50%	99%	WP_102905575.1	64%	100%	WP_102905577.1	52%	99%	WP_102905565.1	34%	68%	WP_094176435.1	32%	50%	WP_094175715.1	39%	92%	WP_099447707.1
	<i>Lachnospiraceae bacterium</i> JH2	54%	94%	WP_049932516.1	68%	100%	WP_049932510.1	50%	98%	WP_044995899.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	39%	88%	WP_044995953.1		
	<i>Kallipyga massiliensis</i> jC7	51%	94%	WP_034579389.1	61%	100%	WP_019133805.1	50%	98%	WP_019133811.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	59%	11%	WP_019133900.1		
	<i>Oribacterium</i> sp. C9	54%	94%	WP_078038383.1	67%	100%	WP_078038384.1	50%	98%	WP_044995899.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	40%	88%	WP_078037321.1		
	<i>Oribacterium</i> sp. oral taxon 108	59%	95%	WP_009427752.1	68%	99%	WP_009429936.1	53%	98%	WP_009429618.1	No significant similarity found	28%	41%	WP_036598124.1	43%	89%	WP_009428268.1		
	<i>Oribacterium asacharolyticum</i> ACB7	59%	95%	WP_009427752.1	68%	99%	WP_009537123.1	53%	98%	WP_009537118.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	42%	89%	WP_009536370.1		
	<i>Anaerocolumna aminovaleric</i> DSM1283	58%	96%	WP_091684121.1	65%	99%	WP_091683826.1	55%	100%	WP_091683826.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	39%	92%	WP_091688942.1		
	<i>Clostridium minihomine</i> Marseille-P4642	59%	95%	WP_101697112.1	70%	99%	WP_101697113.1	51%	98%	WP_101697107.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	35%	89%	WP_101697867.1		
	<i>Cellulosilyticum</i> sp. I15G1012	56%	97%	WP_069998763.1	66%	99%	WP_069998615.1	55%	100%	WP_069998610.1	No significant similarity found	24%	38%	WP_069997310.1	38%	91%	WP_070000766.1		
	<i>Clostridiales bacterium</i> KA00274	54%	95%	WP_066712236.1	51%	99%	WP_066712185.1	51%	99%	WP_066712196.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		No significant similarity found		
	<i>Pygmabacter massiliensis</i> Marseille-P3336	66%	95%	WP_102050939.1	72%	99%	WP_102050944.1	51%	100%	WP_102050944.1	35%	27%	WP_102048535.1	No significant similarity found	37%	93%	WP_102048746.1		
	<i>Oscillibacter ruminantium</i> GH1	54%	93%	WP_040659881.1	72%	100%	WP_040659556.1	51%	98%	WP_040659546.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		No significant similarity found		
	<i>Anaerotruncus</i> sp. G3	54%	97%	WP_016314924.1	69%	100%	WP_016314923.1	52%	98%	WP_016314929.1	No significant similarity found	25%	37%	WP_016317061.1	37%	86%	WP_016316489.1		

LsrB-like receptors in Clostridia

	<i>Clostridium sporosphaeroides</i> DSM 1294	57%	96%	WP_020074135.1	71%	99%	WP_020074134.1	51%	98%	WP_020074140.1	No significant similarity found	No significant similarity found	53%	95%	WP_020071981.1		
	<i>Clostridiales bacterium</i> VE202-13	53%	96%	WP_006872982.1	69%	100%	WP_006872983.1	53%	98%	WP_007863873.1	No significant similarity found	No significant similarity found	35%	89%	WP_024731321.1		
	<i>Anaerotruncus colihominis</i> DSM 17241	53%	96%	WP_006872982.1	69%	100%	WP_006872983.1	52%	98%	WP_006872974.1	No significant similarity found	No significant similarity found	35%	89%	WP_006875429.1		
	<i>Firmicutes bacterium</i> HGW-Firmicutes-7	56%	96%	PKM52091.1	66%	99%	PKM52090.1	54%	100%	PKM52095.1	No significant similarity found	No significant similarity found	39%	90%	PKM50341.1		
	<i>Clostridium merdae</i> Marseille-P2953	59%	96%	WP_085834223.1	71%	99%	WP_085834224.1	50%	99%	WP_085834218.1	42%	100%	WP_085832362.1	28%	73%	WP_085833776.1	
	<i>Eubacterium</i> sp. 14-2	54%	96%	WP_016213794.1	67%	100%	WP_016213793.1	52%	100%	WP_016213799.1	No significant similarity found	No significant similarity found	40%	90%	WP_016213847.1		
	<i>Hespelia stercorisius</i> DSM 15480	48%	96%	WP_084533887.1	66%	100%	WP_073104624.1	52%	98%	WP_073104639.1	No significant similarity found	No significant similarity found	39%	90%	WP_073107248.1		
	<i>Streptococcus pluxetorum</i> DSM 22810	42%	96%	WP_027972155.1	52%	99%	WP_027972154.1	47%	98%	WP_027972156.1	38%	42%	WP_027971411.1	No significant similarity found	48%	92%	WP_027972625.1
	<i>Streptococcus</i> sp. oral taxon 056	40%	95%	WP_009443257.1	51%	99%	WP_009754270.1	48%	98%	WP_009443857.1	No significant similarity found	No significant similarity found	46%	92%	WP_009754741.1		
	<i>Hydrogenoanaerobacterium saccharovorans</i> CGMCC	53%	95%	WP_092753805.1	67%	100%	WP_092753807.1	51%	99%	WP_092753795.1	No significant similarity found	55%	18%	WP_092753459.1	No significant similarity found		
	<i>Granulicatella</i> (multispecies)	38%	95%	WP_005605804.1	50%	99%	WP_005605802.1	49%	98%	WP_005605805.1	No significant similarity found	36%	34%	WP_065427332.1	52%	95%	WP_039848953.1
	<i>Anaerostignum neopropionicum</i> DSM-3847 ^b	46%	96%	WP_066083866.1	65%	100%	WP_066083864.1	51%	100%	WP_066083875.1	61%	100%	WP_066085454.1	26%	84%	WP_083531940.1	
	<i>Acetanaerobacterium elongatum</i> CGMCC	52%	95%	WP_092642784.1	61%	100%	WP_092642781.1	55%	99%	WP_092638662.1	No significant similarity found	No significant similarity found	36%	88%	WP_092641115.1		
	<i>Firmicutes bacterium</i> HGW-Firmicutes-16	50%	93%	PKM72083.1	67%	99%	PKM72023.1	52%	98%	PKM72028.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
	<i>Firmicutes bacterium</i> HGW-Firmicutes-9	47%	96%	PKM39253.1	57%	100%	PKM39254.1	56%	100%	PKM39256.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
	<i>Anaerovirgula multivorans</i> SCA	48%	96%	WP_089285161.1	51%	100%	WP_089285162.1	51%	98%	WP_089285161.1	29%	51%	SNS04453.1	No significant similarity found	58%	90%	WP_089282808.1
	<i>Fusibacter</i> sp. 3D3	48%	95%	WP_069871710.1	51%	100%	WP_069871708.1	51%	99%	WP_069871705.1	50%	14%	WP_069872658.1	No significant similarity found	34%	91%	WP_069874748.1
	<i>Bacillus mesoae</i> FJAT-13985	39%	95%	WP_084797764.1	57%	100%	WP_066388762.1	51%	98%	WP_066388746.1	49%	98%	WP_066383389.1	28%	40%	WP_066389153.1	
	<i>Bacillus</i> sp. AFS073361	38%	95%	WP_098572750.1	55%	100%	WP_098572751.1	51%	98%	WP_098572746.1	34%	58%	WP_098575183.1	32%	79%	WP_098576287.1	
	<i>Bacillus</i> sp. OG2	39%	96%	WP_094768655.1	54%	100%	WP_094768656.1	50%	98%	WP_094768652.1	29%	82%	WP_009793160.1	29%	40%	WP_094768370.1	
	<i>Bacillus</i> sp. AFS037270	38%	95%	WP_098996049.1	54%	100%	WP_098996048.1	51%	98%	WP_098996053.1	28%	78%	WP_098996466.1	30%	79%	WP_098996466.1	
	<i>Salinrhadius</i> sp. Marseille-P4669	42%	97%	WP_102026539.1	56%	100%	WP_102026538.1	53%	98%	WP_102026543.1	No significant similarity found	67%	15%	WP_102027633.1	49%	94%	WP_102027433.1
	<i>Bacillus</i> sp. AFS076308	38%	95%	WP_098532580.1	55%	100%	WP_098532579.1	51%	98%	WP_098532584.1	29%	78%	WP_098528900.1	32%	79%	WP_098530977.1	
	<i>Bacillus</i> sp. UNC41MF55	38%	95%	WP_051633640.1	55%	100%	WP_026565774.1	52%	98%	WP_026565780.1	31%	78%	WP_026565806.1	31%	67%	WP_026565967.1	
	<i>Bacillus ombensis</i>	42%	95%	WP_099360998.1	57%	99%	WP_099360997.1	51%	98%	WP_099361001.1	No significant similarity found	No significant similarity found	53%	88%	WP_099351901.1		
	<i>Bacillus drentensis</i> NBRC 102427	39%	95%	WP_066363083.1	53%	100%	WP_066250468.1	52%	98%	WP_066250483.1	54%	98%	WP_066249865.1	No significant similarity found	49%	92%	WP_066258319.1
	<i>Bacillus</i> sp. MRMR6	39%	95%	WP_075689727.1	57%	100%	WP_075689728.1	52%	98%	WP_075689722.1	56%	98%	WP_075687884.1	60%	15%	WP_075690505.1	
	<i>Geopobacter ferrireducens</i> IRF9	48%	95%	WP_083273656.1	62%	100%	WP_069979544.1	54%	99%	WP_069979540.1	No significant similarity found	No significant similarity found	57%	94%	WP_069978378.1		
	<i>Oscinobacillus damuensis</i> PT-20	40%	96%	WP_067729920.1	56%	100%	WP_067729980.1	52%	98%	WP_067729928.1	No significant similarity found	No significant similarity found	51%	90%	WP_067724955.1		
	<i>Bacillus novalis</i> NBRC 102450	38%	95%	WP_066147899.1	55%	100%	WP_066087747.1	50%	98%	WP_066087725.1	43%	24%	WP_066086681.1	No significant similarity found	49%	92%	WP_024030136.1
	<i>Bacillus massiliogabonensis</i> Marseille P2639	41%	96%	WP_102273479.1	57%	100%	WP_102273478.1	50%	100%	WP_102273484.1	No significant similarity found	No significant similarity found	50%	92%	WP_102274162.1		
	<i>Bacillus cihuenis</i> FJAT-14515	41%	94%	WP_028392000.1	55%	100%	WP_034301163.1	52%	96%	WP_028391995.1	49%	98%	WP_028391214.1	No significant similarity found	54%	86%	WP_028392646.1
	<i>Paenibacillus</i> sp. 7884-2	40%	95%	WP_095313822.1	54%	100%	WP_095313821.1	52%	98%	WP_095313829.1	No significant similarity found	33%	28%	WP_095307015.1	52%	91%	WP_095309744.1
	<i>Bacillus praedii</i> FJAT-25547	41%	94%	WP_057759869.1	56%	100%	WP_057759866.1	51%	98%	WP_057759884.1	No significant similarity found	No significant similarity found	49%	92%	WP_057761154.1		
	<i>Bacillus</i> sp. FJAT-21945	41%	96%	WP_082446766.1	56%	100%	WP_053478413.1	52%	98%	WP_053478419.1	35%	24%	WP_053478317.1	No significant similarity found	49%	92%	WP_053477421.1
	<i>Virgibacillus indicus</i> P2-C2	38%	95%	WP_094887266.1	56%	100%	WP_094887265.1	52%	100%	WP_094887271.1	No significant similarity found	42%	19%	WP_094886613.1	49%	92%	WP_094886575.1
	<i>Proteinilactium ruminis</i> ML2	44%	95%	WP_083422745.1	56%	99%	WP_074912249.1	52%	98%	WP_074912251.1	32%	43%	WP_051651683.1	28%	37%	WP_074909636.1	
	<i>Succinivibrio dectrimosolvens</i> 22B	44%	93%	WP_074840781.1	64%	100%	WP_074840780.1	51%	98%	WP_074840786.1	No significant similarity found	No significant similarity found	38%	92%	WP_074841749.1		
Proteobacteria	<i>Deltaproteobacteria bacterium</i> RBG_13_49_15	40%	96%	OGF59392.1	50%	99%	OGF59393.1	65%	100%	OGF59398.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
	<i>Treponema socranskii</i> ATCC 35536	44%	96%	WP_021330437.1	60%	99%	WP_021330465.1	52%	99%	WP_021330480.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
	<i>Treponema</i> sp. OMZ 838	47%	97%	WP_044014769.1	59%	99%	WP_044014771.1	53%	99%	WP_044014760.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
	<i>Treponema medium</i> ATCC 700293	47%	97%	WP_016522506.1	59%	99%	WP_016522505.1	53%	99%	WP_016522511.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
	<i>Treponema</i> sp. GWA1_62_8	67%	99%	OHES58486.1	67%	99%	OHES58486.1	52%	98%	OHES58489.1	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found	No significant similarity found		
Spirochaetes	<i>Treponema primitia</i> ZAS-1	No protein with significant similarity	24%	95%	010257110.1	No protein with significant similarity	41%	94%	WP_010263218.1	46%	23%	WP_010259931.1	35%	88%	WP_010257819.1		
	<i>Treponema azotonutricium</i> ZAS-9	No protein with significant similarity	25%	95%	WP_015710990.1	No protein with significant similarity	37%	94%	WP_015710185.1	26%	34%	WP_015710350.1	37%	89%	WP_015713045.1		
	<i>Brachyspira</i> sp. CAG:700	27%	92%	CCY74827.1	47%	99%	CCY74818.1	43%	86%	CCY74826.1	No significant similarity found	32%	55%	CCY74993.1	No significant similarity found		
	<i>Sedimentispirochaeta bajacaliforniensis</i> DSM16054	26%	95%	WP_020610753.1	46%	99%	WP_020610759.1	44%	86%	WP_020610754.1	No significant similarity found	41%	17%	WP_020613546.1	No significant similarity found		
	<i>Sedimentispirochaeta smaragdinae</i> DSM11293	26%	95%	WP_013256129.1	46%	99%	WP_013256123.1	44%	86%	WP_013256128.1	25%	29%	WP_013252902.1	41%	17%	WP_013253311.1	

^a This bacterial species is classified as a Bacteroidetes, but recent information indicates that is might be a Firmicute.

^b Previously annotated as *Clostridium neopropionicum*

^c A1-2 metabolism enzymes from *E. coli* K-12 MG1655 (LsrF KEGG ID: b1517; LsrG KEGG ID: b1518)

^d Hits coloured in red were not considered LsrB or LsrG proteins as the query cover and sequence similarity were very low (query cover below 75%)

Table S2.5 - Primers used in this study.

Construct	5' (sense) sequence	3' (antisense) sequence
<i>C. saccharobutylicum</i> WT LsrB	CAGGGCTCCTTCACCTCAAGC	CGCCCACCCCTTTTAGAAATC
<i>C. saccharobutylicum</i> LsrB K48A	GGAAATGTTACTGTAACATTTATTCCCGCACTTACAGG TAATGCATTTTTTGAATC	GATTCAAAAAATGCATTACCTGTAAGTGCGGGAATA AATGTTACAGTAACATTTC
<i>C. saccharobutylicum</i> LsrB D129A	GTTACAGTTACTACATGGGCTTCTGACGTAGATCCATC A	TGATGGATCTACGTCAGAAGCCCATGTAGTAACTGTA AC
<i>C. saccharobutylicum</i> LsrB D186A	TTATTCAAATGCAACAGTTACTGCTCAAAATTCATGGC AGGTAGAAG	CTTCTACCTGCCATGAATTTTGAGCAGTAACTGTTGC ATTTGAATAA
<i>C. saccharobutylicum</i> LsrB Q187A	TTCAAATGCAACAGTTACTGATGCAAATTCATGGCAG GTAGAAGGG	CCCTTCTACCTGCCATGAATTTGCATCAGTAACTGTTG CATTTGAA
<i>C. saccharobutylicum</i> LsrB N242A	CACATTCTGATATTGACTTAATTATATGTGCTGATTCT ACAGCACTTCCTGGA	TCCAGGAAGTGCTGTAGAATCAGCACATATAATTAA GTCAATATCAGAATGTG
<i>C. saccharobutylicum</i> LsrB S244A	TCTGATATTGACTTAATTATATGTAATGATGCTACAGC ACTTCCTGGAC	GTCCAGGAAGTGCTGTAGCATCATTACATATAATTAA GTCAATATCAGA
<i>C. saccharobutylicum</i> LsrB D186N	CATTATTCAAATGCAACAGTTACTAATCAAAATTCAT GGCAGGTAGAAG	CTTCTACCTGCCATGAATTTTGATTAGTAACTGTTGCA TTTGAATAATG
<i>C. saccharobutylicum</i> LsrB S244T	GATATTGACTTAATTATATGTAATGATACTACAGCACT TCCTGGACAAG	CTTGTCCAGGAAGTGCTGTAGTATCATTACATATAAT TAAGTCAATATC
<i>C. saccharobutylicum</i> LsrB N242P	TGTCTGCACATTCTGATATTGACTTAATTATATGTCCTG ATTCTACAGCACTTC	GGAAGTGCTGTAGAATCAGGACATATAATTAAGTCA ATATCAGAATGTGCAGACA

Chapter



**Development of a chemical probe
to identify AI-2 receptors**

Unpublished data and data published in:

Miranda, V., **Torcato, I. M.**, Xavier, K. B. and Ventura, M. R. (2019) Synthesis of D-desthiobiotin-AI-2 as a novel chemical probe for autoinducer-2 quorum sensing receptors. *Bioorg. Chem.* **92**. (Electronically published ahead of print) DOI: 10.1016/j.bioorg.2019.103200

Author contributions:

Inês M. Torcato, Vanessa Miranda, Karina B. Xavier and M. Rita Ventura conceptualized the studies involving chemical probes. Inês M. Torcato and Karina B. Xavier conceptualized the studies involving *Clostridioides difficile*. Vanessa Miranda and M. Rita Ventura designed the synthetic route. Vanessa Miranda synthesized the chemical probes tested and Inês M. Torcato did the remaining experiments.

3.1. Abstract

Quorum sensing is a cell-to-cell communication mechanism that enables bacteria to sense their numbers and regulate group behaviors accordingly. This process is mediated by small chemical molecules named autoinducers. Autoinducer-2 (AI-2) is an autoinducer involved in interspecies communication. Although the synthesis pathway of AI-2 is widely conserved, the same cannot be said about the detection mechanisms. Only two types of AI-2 receptors have been identified: LsrB and LuxP. The identification of novel AI-2 receptors poses a challenge since the known receptors do not have a specific fold and share little sequence homology, rendering useless the common strategies of fold and sequence comparison. Moreover, bacteria like *Clostridioides difficile* have phenotypes as biofilm formation and toxin production regulated by AI-2, but do not possess proteins with similarity with the known AI-2 receptors. Evidence such as this, suggests that other AI-2 receptors could exist. Thus, there is a need for the development of new strategies for identification of AI-2 receptors. Here, we propose a new chemical probe that consists in AI-2 bound to D-desthiobiotin, a derivative of biotin routinely used for pull down of biological molecules. We show that this probe is able to bind both known receptors and that, through immunoprecipitation, we are able to pull down LuxP from a partially purified protein solution. Additionally, we show the production profile of AI-2 in *C. difficile* and preliminary attempts of identifying a new AI-2 receptor in this bacterium. The development and further optimization of this strategy will allow the understanding and, possibly, the modulation of the molecular mechanisms in clinically relevant bacteria and in complex multispecies environments.

3.2. Introduction

Bacteria are able to sense their environment and regulate their behavior accordingly. To achieve this, they employ quorum sensing, a cell-to-cell

communication mechanism, mediated by the production and detection of small chemical molecules, named autoinducers (2,8,9). This mechanism allows bacteria to monitor their population density and regulate gene expression accordingly (2,3). The behaviors regulated by this mechanism are the ones that benefit from group action, like biofilm formation and virulence factor production, which would be ineffective when done by a low number of cells (2,3). Autoinducer-2 is an autoinducer involved in interspecies communication (50). Is produced and recognized by both Gram-negative and -positive bacteria and is synthesized as the linear precursor (S)-4,5-dihydroxypentane-2,3-dione (DPD) in a reaction catalyzed by LuxS. Although all bacteria that have AI-2 mediated quorum sensing synthesize AI-2 in its linear form, through LuxS, this molecule is rapidly interconvertible in other cyclic forms in aqueous solution. Thus, AI-2 comprises a family of molecules that exist in equilibrium in aqueous solution. The two identified signaling forms are (2R,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran (R-THMF) and (2S,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran-borate (S-THMF-borate) that bind the two types of AI-2 receptors known, LsrB and LuxP, respectively (46-48,50). LsrB is a substrate binding protein that interacts with an ABC transporter to internalize AI-2 whereas LuxP, also a substrate binding protein, interacts with a histidine kinase two component system, which regulates a phosphorylation signal transduction cascade (46,47,59,60,62,63).

There is evidence for the existence of other AI-2 receptors as bacteria that do not possess LuxP nor LsrB receptors have phenotypes regulated by AI-2. For example, biofilm formation and toxin production are suggested to be regulated by AI-2 in *Clostridioides difficile*, a toxin-producing, multidrug-resistant bacterium that is the major cause of nosocomial diarrhea in adults with a dysbiotic gut microbiota due to treatments with broad-spectrum antibiotics (57,158,159). Nevertheless, no protein with similarity to LsrB or LuxP receptors was identified in this organism. The difficulty in identifying new classes of AI-2 receptors resides in the fact that,

although LsrB and LuxP share the same fold, so do other substrate binding proteins (46,47,65,160). Thus, the fold is not specific enough for identification of novel receptors. Moreover, LsrB and LuxP share a very small amino acid sequence identity (approximately 11%) indicating that other receptors with low homology to the known ones might exist. However, if that is indeed the case, the usual sequence comparison approaches are ineffective for identification of these novel AI-2 receptors (46). So, the development of novel approaches to identify new classes of AI-2 receptors is an important step to understand the mechanistic processes that underlie AI-2 mediated quorum sensing.

An AI-2 chemical probe based on the structure of rhodamine has been developed (161). However, this probe could only bind LsrB receptors (161). This selectivity can be explored for some purposes, but for identification of novel AI-2 receptors is a disadvantage, as AI-2 is not able to acquire the needed conformation to bind LuxP. Since the active form of AI-2 that will bind new AI-2 receptors is not known, a probe that allows AI-2 to cyclize into different forms is essential to identify novel AI-2 receptors.

Biotin and streptavidin form one of the strongest noncovalent biological interactions with a constant of affinity of approximately 10^{-15} M (162). Additionally, the fact that their binding is highly specific and resistant to changes in temperature or pH, has made their utilization to study biological systems highly attractive. However, their bond can also withstand the presence of organic solvents and denaturing solvents. Thus, harsh conditions have to be employed to promote their release (163). To overcome this step, derivatives of streptavidin and biotin that have high binding affinity but lower than the one observed for the pair biotin/streptavidin have been developed (163). D-desthiobiotin is a derivative of biotin that lacks the sulfur atom and has lower binding affinity for streptavidin (approximately 10^{-13} M) (164,165). The pair D-desthiobiotin/streptavidin has been previously used for identifying biological molecules like lentivirus, DNA-protein conjugates and metalloproteins, and has shown potential to be used as a system

for improved drug screening efficiency and disease mechanism elucidation (162,165-167). The advantage of this system is the ability to obtain the molecules of interest by adding a buffered solution of biotin, which binds streptavidin with higher affinity and thus dislocates desthiobiotin. Moreover, antibodies able to recognize biotin, also recognize desthiobiotin making a variety of immunoassays available (165-168).

In this work, we describe the first AI-2 chemical probe able to bind both LuxP and LsrB receptors. This probe consists in AI-2 bound to D-desthiobiotin (DB-AI-2) and thus, many previously described strategies to pull down and identify novel AI-2 receptors are available. The binding of this probe to streptavidin was not possible due to the small size of AI-2 and the large size of the receptors, which seem to block the binding sites for streptavidin in desthiobiotin. Still, an immunoprecipitation protocol for pull down of partially purified receptors was developed and its applicability for novel AI-2 receptors in cell extracts of *C. difficile* is being optimized. The identification of novel AI-2 receptors will increase the understanding of the molecular mechanisms of AI-2 mediated quorum sensing in clinically relevant bacteria and in complex multispecies environments. Furthermore, this knowledge can afterwards be employed in the design of quorum sensing molecules able to modulate the quorum sensing process.

3.3. Results

3.3.1. Solvents used for synthesis of TentaGel based chemical probes denature the AI-2 receptor

To develop a chemical probe able to pull down AI-2 receptors, we tried a strategy that consisted in binding AI-2 to an insoluble chemical matrix. However, this matrix had to swell in aqueous solution, so that AI-2 would be available to bind the receptors. Thus, our collaborators from Rita Ventura's laboratory linked

the linear form of AI-2, DPD, to a TentaGel resin (TG-AI-2; Fig. 3.1). This type of resin results from a process where polyethylene glycol units of approximately 3 kDa are attached to low-cross-linked polystyrene (169,170). The final resin is a relatively hydrophilic support, which swells both in water and in most polar organic solvents. Moreover, direct evidence that large biomolecules can penetrate into TentaGel (TG) beads and that the environment within the beads is similar to that in solution, exist (170). The presence of a linker between the resin and AI-2 would provide the necessary space for AI-2 to cyclize into the needed conformation to bind AI-2 receptors (Fig. 3.1).

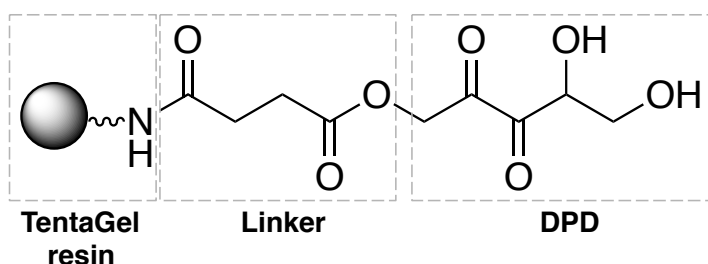


Figure 3.1 - TG-AI-2: TentaGel based chemical probe. The solid matrix is functionalized with an amine group that reacts with the linker bound DPD.

To test the binding of LuxP to this probe, TG-AI-2 was incubated with CLPY that consists in a modified LuxP from *V. harveyi*. This altered receptor has a cyan and a yellow fluorescent protein bound to each of its domains (171). Upon binding of AI-2, LuxP will suffer a conformational change that will result in a decrease in Förster resonance energy transfer (FRET) response (171). Thus, by using this altered LuxP we have the possibility to assess changes on the viability of the receptor during binding tests. After incubation, a batch purification process that includes a denaturation step to elute bound protein was performed. To evaluate the pull-down of LuxP, the samples were loaded onto a SDS-PAGE gel (Fig. 3.2).

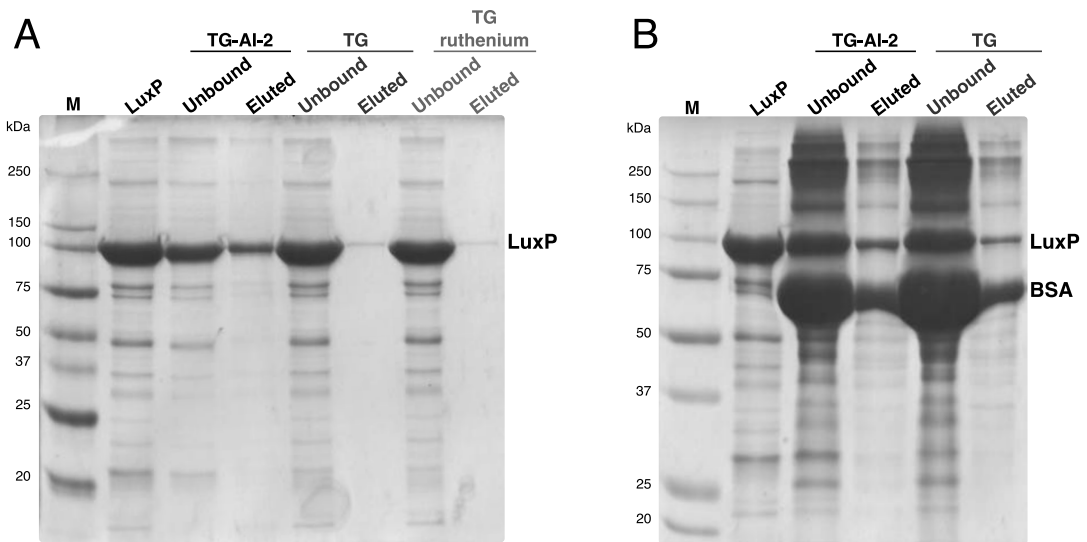


Figure 3.2 - TentaGel based probe does not specifically bind LuxP. (A) SDS-page gel of the unbound sample and eluted proteins after incubation with the TentaGel resin with AI-2 bound (TG-AI-2), TentaGel resin alone (TG) and TentaGel resin with ruthenium (TG+ruthenium). (B) SDS-page gel of the unbound sample and eluted proteins after incubation with the TG-AI-2 and TG. The resins were pre-treated with bovine serum albumin (BSA) to eliminate unspecific interactions. LuxP was loaded into the gel as control. M: protein standards.

No specific binding of LuxP to TG-AI-2 was observed. In the first attempt there was a difference when comparing the elutions of LuxP incubated with TG-AI-2 (Fig. 3.2, A, well 3) with the ones of LuxP incubated with the controls: TG and TG with ruthenium (Fig. 3.2, A, wells 5 and 7). However, that was most probably due to the fact that for TG-AI-2, 24 times more resin was used. Importantly, ruthenium did not seem to affect the results. A new test using the same amount of TG-AI-2 and TG resins was performed. To try to decrease the interference of unspecific binding an extra step was added and TG-AI-2 and TG were incubated with 3% bovine serum albumin (BSA) before being incubated with LuxP. Nevertheless, no significant difference between the assays performed with TG-AI-2 and TG could be observed (Fig. 3.2, B, wells 3 and 5, respectively).

To evaluate the viability of LuxP after incubation with the resins, we assayed the recovered unbound protein with multiple concentrations of synthetic AI-2 (Fig. 3.3, A).

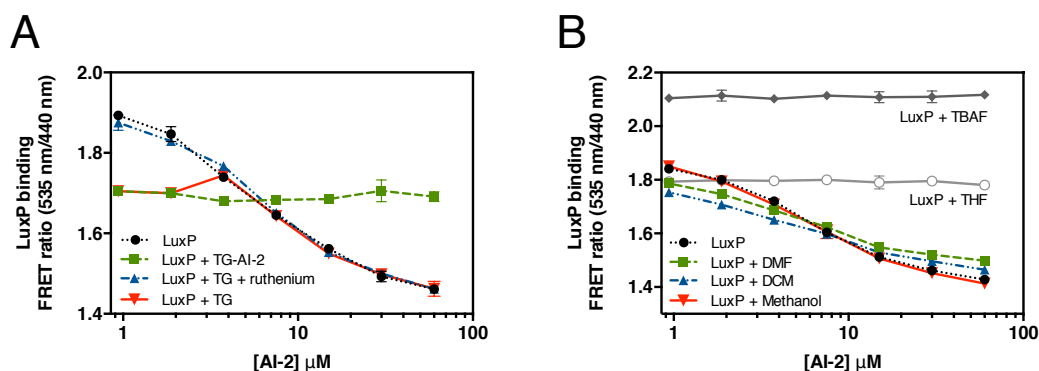


Figure 3.3 - Solvents used in the synthesis of the TentaGel-based probe cause denaturation of LuxP. (A) FRET assay using a modified LuxP protein (CLPY) shows that only the complex AI-2-TentaGel resin (TG-AI-2) causes denaturation of the protein as no binding to LuxP is observed. TentaGel resin alone (TG) or supplemented with ruthenium (TG + ruthenium), a contaminant from the synthesis process, did not affect the AI-2 binding ability of LuxP. (B) Incubation of LuxP with different solvents used for the synthesis process shows that whereas dimethylformamide (DMF), dichloromethane (DCM) and methanol do not affect the binding of AI-2 to LuxP, tetrabutylammonium fluoride (TBAF) and tetrahydrofuran (THF) alter LuxP making it unable to bind AI-2.

Only the receptor incubated with TG-AI-2 lost the ability to bind AI-2. This suggested the possibility that some solvent that remained from the synthesis process was denaturing LuxP. Thus, we incubated LuxP with the solvents used in the synthesis of TG-AI-2 and performed a new FRET assay. From these results we concluded that both tetrabutylammonium fluoride (TBAF) and tetrahydrofuran (THF) induced the denaturation of the receptor (Fig. 3.3, B). All together, these results indicate that although TG-AI-2 seemed like a promising probe, different solvents would have to be used in its synthesis for it to be viable.

3.3.2. D-Desthiobiotin-AI-2 binds both LuxP and LsrB receptors

Since TG-AI-2 was not a viable option, we decided to adapt a system previously used for pull-down of biological molecules and, thus, compatible with biological systems. Our collaborators at Rita Ventura's laboratory attached AI-2 to D-desthiobiotin (DB-AI-2; Fig. 3.4).

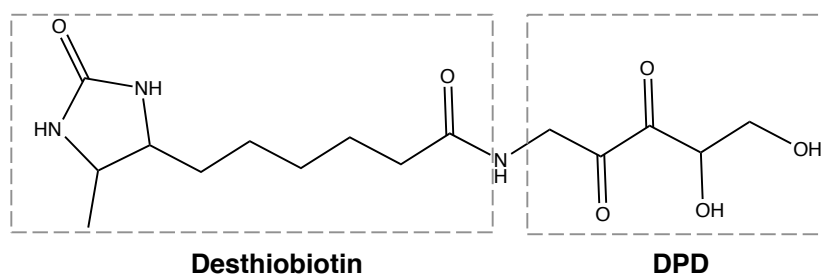


Figure 3.4 - D-desthiobiotin based probe. The linear form of AI-2, DPD, was bound to D-desthiobiotin. The hydroxyl group on D-desthiobiotin was substituted by an amide during the synthetic process of DPD binding.

To determine the functionality of this chemical probe we tested binding of DB-AI-2 to known AI-2 receptors, including the well-characterized LuxP from *V. harveyi* (CLPY) and the canonical LsrB receptor from *Escherichia coli* (EcLsrB). Recently, an LsrB receptor with a different binding site was identified in *Clostridium saccharobutylicum*, and we also tested this non-canonical LsrB receptor (CsLsrB) (Chapter 2; (160)). After affinity purification, these proteins were incubated with DB-AI-2. To test for specificity, receptor proteins were also incubated with non-conjugated D-desthiobiotin and with buffer alone. After incubation, these samples were run in a native gel to prevent denaturation of the receptors. The native gel was used for immunoblotting using an anti-biotin antibody able to recognize desthiobiotin (Fig. 3.5, A). A native gel run in parallel with the same samples was stained with Coomassie to visualize the bands (Fig. 3.5, B). In the wells of LsrB proteins incubated with DB-AI-2 (Fig. 3.5 B) some bands are fainter. This might suggest that upon binding to DB-AI-2 multimeric interactions might have influenced the migration of the protein in the native gel, as an SDS-PAGE gel, ran in denaturing conditions with the same samples (Fig. S3.1), showed that all the samples had the same protein concentrations. Signal was only detected in the western blot when the receptor proteins were incubated with DB-AI-2, showing that the antibody recognized the DB-AI-2 bound to both LuxP and LsrB receptors (Fig. 3.5, A). No bands were detected in the negative control wells loaded with the receptors incubated with non-conjugated D-desthiobiotin, or buffer. Thus, it is possible to conclude that there is no unspecific binding of

desthiobiotin (unbound desthiobiotin should run off the gel) and of the antibody to the receptor proteins.

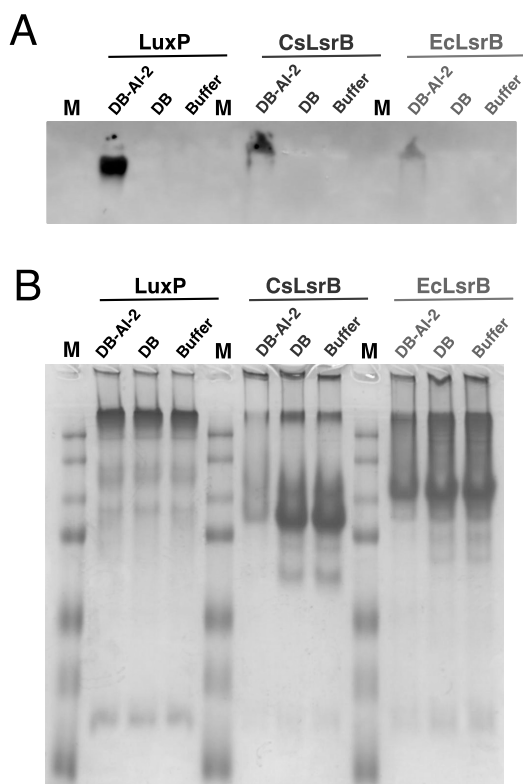


Figure 3.5 - D-desthiobiotin bound AI-2 binds both LuxP and LsrB receptors. (A) Western-blot membrane incubated with an anti-biotin antibody able to recognize desthiobiotin. M: protein standards; CsLsrB: non-canonical LsrB from *C. saccharobutylicum*; EcLsrB: canonical LsrB receptor from *E. coli*; DB-AI-2: receptor proteins incubated with D-desthiobiotin bound to AI-2; DB: receptor proteins incubated with D-desthiobiotin and Buffer: receptor proteins incubated with buffer. (B) Native PAGE gel run in parallel to the one used for western blot detection, stained with Coomassie Blue.

3.3.3. D-Desthiobiotin-AI-2 can pull-down LuxP

Having established the functionality of the chemical probe, a strategy for pulling down the AI-2 receptors was needed. We first explored the high affinity binding of streptavidin to desthiobiotin. We observed that when using either streptavidin agarose beads (Fig. S3.2, B) or streptavidin magnetic beads (Fig. S3.2, C), LuxP bound to desthiobiotin was always present in the flow-through and

washes, fractions where the unbound sample is removed. Moreover, binding of commercial D-desthiobiotin to streptavidin magnetic beads followed by incubation with an antibody able to recognize desthiobiotin (Fig. S3.2, D) showed that either desthiobiotin was not binding streptavidin, what seems unlikely, or that when streptavidin binds desthiobiotin, its binding points became occupied and it is not able to bind another protein. Thus, we hypothesized that when DB-AI-2 is bound to LuxP, desthiobiotin is not “free” to bind streptavidin. Since we were using CLPY that is even larger than common AI-2 receptors we tried inverting the order of binding using CsLsrB. So, streptavidin magnetic particles were incubated with DB-AI-2 and CsLsrB was only added afterwards (Fig. S3.2, E). Nevertheless, the protein was again recovered in the flow-through and wash fractions showing that if streptavidin was able to bind DB-AI-2, no contact points were available to bind CsLsrB.

However, we showed that an anti-biotin antibody is able to recognize the DB-AI-2-receptor complex. So, we decided to explore an immunoprecipitation approach. To optimize this procedure, we employed magnetic beads conjugated to protein G that has the ability to immobilize antibodies (Fig 3.6, A).

Using this approach, we were able to pull-down LuxP through one-step purification (Fig. 3.6, B). Moreover, magnetic separation improved the process as washes and elution steps are faster and consume fewer reagents. The denaturing elution used caused the antibody to elute with LuxP.

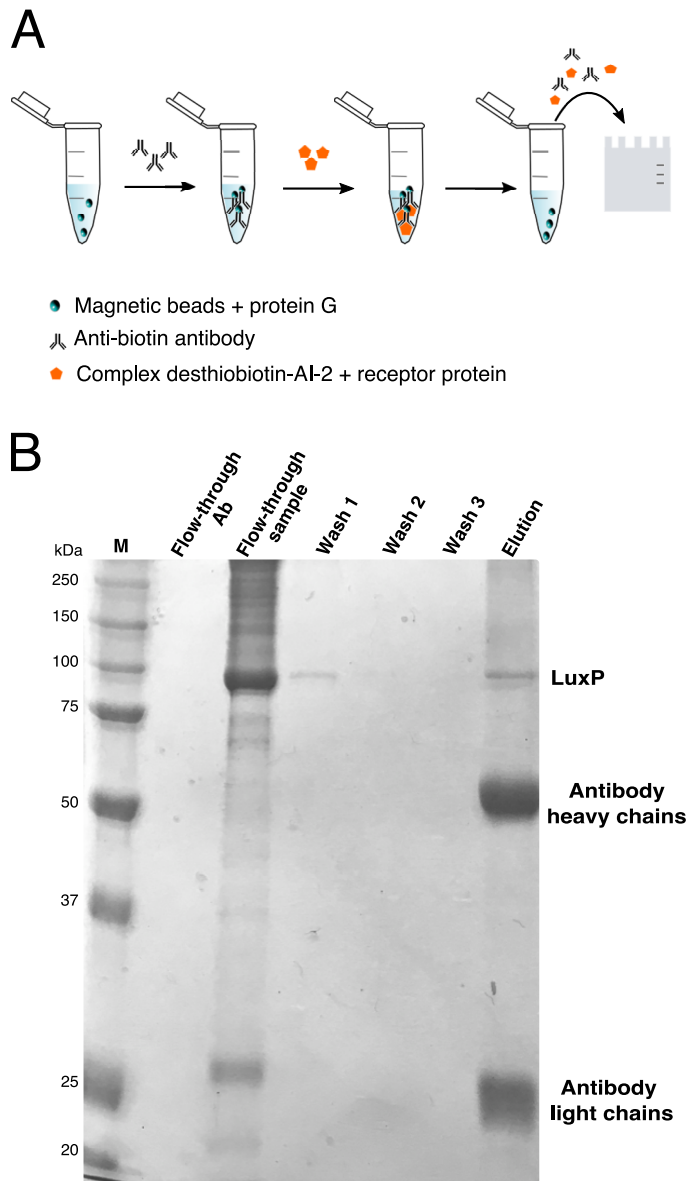


Figure 3.6 – Immunoprecipitation of LuxP using DB-AI-2. (A) Scheme of experimental procedure. Magnetic beads conjugated with protein G are incubated with an antibody that is able to bind the complex DB-AI-2-LuxP through desthiobiotin recognition. After incubation with DB-AI-2 to promote the binding of the antibody to the sample, three washes to remove unbound sample are performed. Afterwards, elution in denaturing conditions is performed and samples are subjected to SDS PAGE electrophoresis. (B) SDS PAGE gel of the washes and elution steps of the procedure described above. M: protein standards; Flow-through Ab: unbound antibody, Flow-through sample: sample that is not retained by the complex beads/antibody; Wash 1- 3: washes to eliminate unbound sample; Elution: elution of the receptor protein. Since denaturing conditions were used the anti-desthiobiotin antibody was also eluted.

3.3.4. AI-2 production profile in *C. difficile*

It has been reported that biofilm formation and expression of toxins A and B are regulated by LuxS, the AI-2 synthase or AI-2 in *C. difficile* (57,158,159). As many strains of this bacterium possess LuxS, we evaluated the profile of AI-2 production in *C. difficile* 630 Δ erm. Because we do not know how a putative AI-2 receptor is regulated in *C. difficile*, we choose a growth medium without added sugars. This would prevent catabolite repression of a system like the Lsr system that mediates the internalization of AI-2 (52,62). Moreover, we measured growth and AI-2 production at different pH values (Fig. S3.3). *C. difficile* is usually grown at pH 7.4 but at that pH, AI-2 is rapidly degraded (Fig. S3.3, A). Given the unstable nature of DPD, we wanted to choose the growth conditions where the chemical degradation of AI-2 would be as minimal as possible. Since AI-2 is stable at pH values equal or lower than 7 we incubated AI-2 at pH 6.8 and observed lower degradation (Fig. S3.3, B). Thus, we evaluated the AI-2 production profile of this bacterium in tryptone-yeast extract (TY) medium at pH 6.8 (Fig. 3.7).

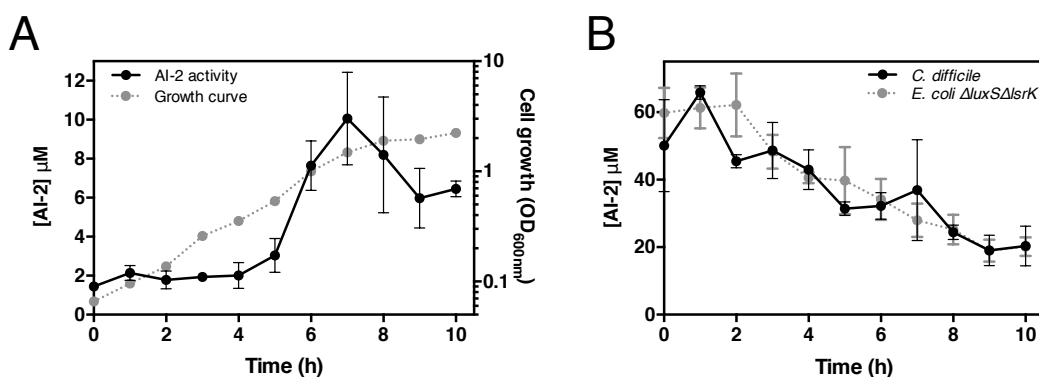


Figure 3.7 – AI-2 production profile of *C. difficile*. (A) Extracellular AI-2 detected in *C. difficile* (black circles) during bacterial growth (grey circles). (B) Concentration of AI-2 measured over time, after addition of 40 μ M of AI-2 at time zero to cultures of *C. difficile* and *E. coli* Δ luxS Δ lsrK. Cell-free fluids were collected at the indicated time points and AI-2 concentration was assayed using a FRET assay. The curves are representative of three independent experiments. Error bars represent the standard deviation of three technical replicates.

Cell free supernatants of *C. difficile* were collected at the indicated time points post-inoculation and the concentration of extracellular AI-2 was assayed through a

FRET assay (Fig. 3.7, A). It is possible to observe that AI-2 accumulated in the extracellular milieu after 5h of growth. There was a decrease in AI-2 concentration at 7h of growth. However, this decrease could be due to chemical degradation of AI-2, as the abrupt decrease of extracellular AI-2 concentration usually observed in bacteria with LsrB receptors was not observed. To confirm this observation, we added 40 μ M of AI-2 at the time of inoculation of *C. difficile* and of an *E. coli* mutated in the synthase, LuxS, and kinase, LsrK, of the Lsr system and thus, unable to produce and internalize AI-2. By comparing the levels of extracellular AI-2 in cell free fluids collected during the growth of these strains we can evaluate if the decrease in AI-2 concentration is due to chemical degradation or due to a transport system present in *C. difficile*. As the levels of AI-2 are very similar between the two conditions, we can infer that the decrease observed should be due to chemical degradation of the molecule and that *C. difficile* seems to produce and accumulate AI-2 extracellularly.

3.3.5. Strategy to pull-down the AI-2 receptor in *C. difficile*

The observation that *C. difficile* has phenotypes regulated by AI-2 suggests the presence of an AI-2 receptor. However, no protein with amino acid sequence homology with the known receptors was identified. Thus, we decided to use DB-AI-2 to try and pull-down the unknown AI-2 receptor in this bacterium. By performing the same protocol as the one employed to pull-down LuxP from partially purified protein (Fig. 3.6) we obtained many interfering protein bands (Fig. 3.8, A). Because all known AI-2 receptors have a molecular weight of approximately 40 kDa we were specially interested in protein bands in between the heavy (50 kDa) and light (25 kDa) chains of the antibody. Looking at the elution well, we observe that the bands obtained are in the same position as the ones observed in more quantity in the unbound sample (Fig. 3.8, A; denaturing elution vs flow-through sample). Thus, we need to pre-treat the lysates to improve the specificity of the pull-down. As we employed an anti-biotin antibody to

recognize desthiobiotin, to obtain the specific band for the new receptor, we had to eliminate the biotinylated proteins that would also be recognized by the antibody.

To eliminate these biotinylated proteins, we incubated the sample (cell extracts previously incubated with DB-AI-2) with streptavidin magnetic particles (Fig. 3.8, B), which are able to bind biotinylated proteins but cannot recognize or bind DB-AI-2 bound to AI-2 receptors (Fig. S3.2, C-E). We collected the unbound sample for further treatment but performed an elution to check if we had eliminated non-specific bands. Again, the bands with higher intensity were the ones observed both in the sample added to the beads and in the unbound proteins (Fig. 3.8, B denaturing elution vs cell lysate/flow-through sample). Thus, it was not possible to conclude if this step was contributing to remove unspecific protein bands. The unbound sample (Fig. 3.8, B flow-through sample) was then incubated with anti-biotin antibody bound magnetic beads. Besides, the usual denaturing elution, we tested an affinity elution where synthetic AI-2 was added to the sample. Free AI-2 should have more affinity than DB-AI-2 and thus, addition of synthetic AI-2 in excess should dislocate the bound protein. Nevertheless, direct addition of AI-2 caused denaturation of the sample and, consequently, co-elution of the antibody due to the low pH of the AI-2 sample, typically used to keep the molecule stable at high concentration. Thus, elution by molecular replacement and by denaturing conditions gave the same results (Fig. 3.8, B elution AI-2 and denaturing elution). For both procedures, no bands besides the bands from the denatured antibody were visible indicating that we might need to increase detection sensitivity to identify potential AI-2 binding proteins.

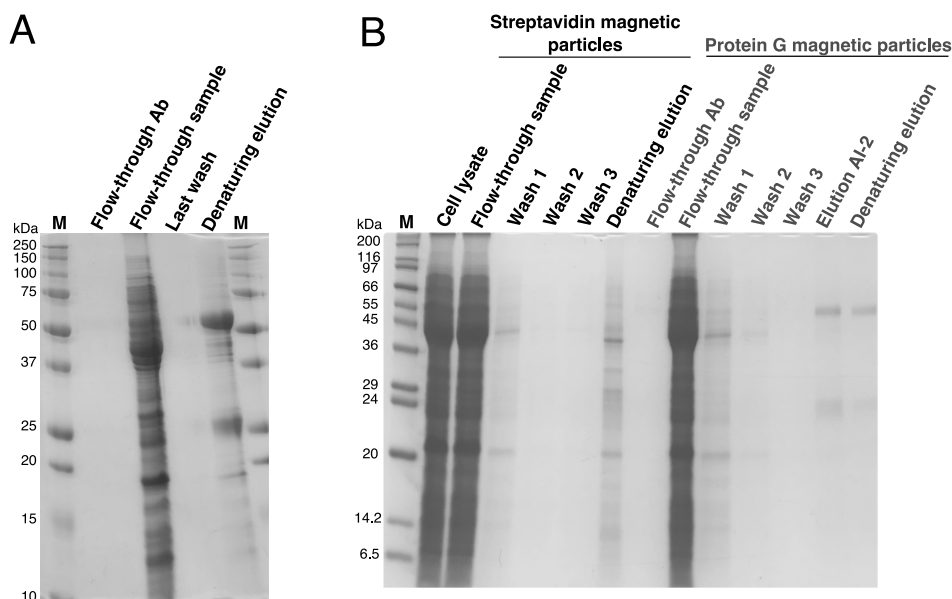


Figure 3.8 – Attempts to pull down the unknown AI-2 receptor of *C. difficile*. (A) SDS-page of the fractions obtained by performing the DB-AI-2 pull-down protocol without pre-treating the sample (B) SDS-PAGE gel from the steps to remove contaminant biotinylated proteins (using streptavidin magnetic particles) and to pull down *C. difficile* AI-2 receptor (using protein G magnetic particles bound to anti-desthiobiotin antibody). Flow-through Ab: unbound antibody; Flow-through sample: unbound sample; Wash: washes to remove unbound sample; Cell lysate: cell lysates incubated with DB-AI-2; Elution AI-2: sample dislocated by synthetic AI-2; Denaturing elution: sample removed from beads through denaturing buffers and heating; M: protein standards.

3.4. Discussion

The design of a chemical probe to identify novel AI-2 receptors has been challenging both due to the multimeric and instable nature of AI-2 and the difficulty in finding compatible reagents for the chemical and biological aspects. To make sure that AI-2 can acquire different conformations in aqueous solution, a linker that allows the linear precursor to cyclize is needed. Moreover, due to the instability of AI-2 a synthetic procedure that can be carried out without complete dryness, in aqueous solution, and at neutral/low pH values is desirable (51,172). Thus, the TentaGel based probe seemed to be a good solution, as this solid matrix could swell in water, being compatible with the biological assays (170). However, remains of the chemical solvents used in the synthesis process were too aggressive

for LuxP causing its denaturation (Fig. 3.3). The adjustment of the desthiobiotin/streptavidin binding pair for pull down of AI-2 receptors, similar to what has been done for other biological molecules, could overcome the problems of chemical and biological compatibility. Indeed, we showed that DB-AI-2 is able to bind all the known AI-2 receptors (Fig. 3.5). This is the first report of a chemical probe, able to bind both LuxP and LsrB receptors, as the only other probe known was able to bind only LsrB receptors (161). The inability to bind LuxP receptors might be due to structural constraints that prevent the conjugation of DPD with boron to form *S*-THMF-borate, which is larger than *R*-THMF, recognized by LsrB receptors. Moreover, the previously designed probe was modified from rhodamine allowing the visualization of the probe inside the cells (161). Nevertheless, it did not have immediate pull down potential. On the contrary, DB-AI-2 has potential for pull down and consequent identification of the proteins, as shown by the pull down of LuxP from the partially purified protein solution (Fig. 3.6). Additionally, streptavidin bound fluorophores that might allow the visualization of bacteria with novel AI-2 receptors also exist (173). These fluorophores could be used for direct visualization through microscopy or for identification of bacterial species with new AI-2 receptors through cell sorting. It has been shown that streptavidin has a β -barrel structure and that the biotin binding site is situated in the interior (163). So, to be able to employ this visual approach, a longer linker between AI-2 and DB would have to be designed to allow streptavidin recognition of DB, even when the AI-2 receptors are bound to AI-2.

The attempts to identify an AI-2 receptor in bacterial strains with AI-2 regulated phenotypes, using DB-AI-2, are still preliminary. *C. difficile* is an appealing bacterium to study the presence of a novel AI-2 receptor due to its infectious nature, its multiple resistances to antibiotics and the observation of existence of phenotypes regulated by AI-2 (57,158,159). *C. difficile* does not have proteins with significant similarity with LuxP or LsrB receptors and its AI-2

production profile suggests that no AI-2 internalization occurs, contrarily to what has been observed for other members of the *Clostridium* genus (Chapter 2; (160)). This observation is in line with a previous report for AI-2 production in *C. difficile* 630 grown in BHI medium (174). Both in our results and in the results described by Carter et al, a decrease in AI-2 extracellular concentration was observed in the late exponential/stationary phase of growth. However, our control experiment where 40 μ M of AI-2 was added both to *C. difficile* and to an *E. coli* unable to produce and internalize AI-2 showed that no AI-2 internalization system seemed to exist in *C. difficile* (Fig. 3.7, B). Thus, this decrease might be due to chemical degradation of the molecule as chemical degradation of AI-2 is significant, even in simple growth medium, from 6h onward (Fig. S3.3, B).

However, no new AI-2 receptor was identified yet. Although promising, the pull down process using DB-AI-2 with cell extracts of *C. difficile* needs optimization. A greater amount of bacterial suspension to prepare the cell extracts and silver staining of the polyacrylamide gels are procedures that can be employed to enhance sensitivity. To ameliorate the pre-treatment of the cell lysates, the lysates can be incubated with the antibody-bound magnetic beads. This treatment should be more efficient than the one with streptavidin magnetic particles, as it can remove the unspecific binding of the antibody used for pull-down for receptors. Also, by performing this purification step before incubating the cell lysates with DB-AI-2, we should not be eliminating the target proteins. Moreover, dilution of AI-2 in concentrated buffer with stronger ionic force should maintain the pH near 7, allowing the elution of the protein of interest without the co-elution of all others due to denaturation. As all known AI-2 receptors have around 30-40 kDa, other receptors are expected to have a similar molecular weight (46,47). This means that bands of interest should appear in between the heavy and light chains of the antibody. Nevertheless, if this presents a problem going forward, covalent linking of the antibody to the beads can readily be accomplished.

All in all, this work provides the first probe able to bind both LuxP and LsrB receptors. This probe has great potential for identification of novel AI-2 receptors as well as identification of bacterial species that possess these new receptors. The identification of the *C. difficile* AI-2 receptor can also bring important developments for dealing with this pathogenic bacteria as biofilm formation and toxin production are suggested to be regulated by AI-2.

3.5. Experimental procedures

3.5.1. Chemical synthesis

(4S)-DPD was synthesized as described in (51), TentaGel bound to AI-2 was synthesized as described in (175) and D-desthiobiotin bound to AI-2 was synthesized as described in (175).

3.5.2. Protein expression and purification

Escherichia coli BL21 $\Delta luxS$ transformed with the plasmid pQE30 (carrying LuxP protein from *Vibrio harveyi* fused to cyan (CFP) and yellow (YFP) fluorescent proteins) – pQE30-CLPY- was kindly provided by Richard Sayre (The Ohio State University, Ohio, USA). Expression and purification of this protein was carried out as previously described (171). Briefly, protein expression was induced at an optical density (OD₆₀₀) of 0.6 with 0.3 mM of isopropyl β -D-thiogalactopyranoside (IPTG) and grown for 6h, at 28 °C before cell harvesting by centrifugation. The resulting pellet was resuspended in 25 mM sodium phosphate buffer pH 8.0, 35 mM NaCl, 1:100 protease inhibitor cocktail (Roche) and cells were lysed by passing them three times through a French pressure cell. Cell debris were removed by centrifugation and LuxP was purified from clarified supernatants through Nickel-nitrilotriacetic (Ni²⁺-NTA) acid affinity chromatography (Qiagen). The protein was eluted in 25 mM sodium phosphate buffer pH 8.0, 35 mM NaCl,

40% glycerol, 250 mM imidazole, 1:100 protease inhibitor cocktail. The expression and purification of *Clostridium saccharobutylicum* DSM13864 and *E. coli* K-12 MG1655 LsrB proteins (CsLsrB and EcLsr, respectively) were carried out as previously described (17,160). Concisely, *E. coli* BL21 $\Delta luxS$ cells transformed with the constructs pPro-EcLsrB and pDEST-CsLsrB were grown in LB with 1 $\mu\text{g}/\text{mL}$ ampicillin at 37 °C until the OD₆₀₀ was 0.3. At this point, temperature was decreased to 22 °C. At OD₆₀₀ = 0.9, protein expression was induced with 0.1 mM or 0.3 mM of IPTG for the expression of EcLsr and CsLsrB, respectively. Cells grew for 6h after induction and were harvested by centrifugation. The resulting pellets were resuspended in 50 mM NaH₂PO₄ (pH 8.0), 300 mM NaCl, 10 mM imidazole, 2.5 $\mu\text{g}/\text{mL}$ DNase and 2.5 $\mu\text{g}/\text{mL}$ leupeptin and cells were lysed by passing thrice through a French pressure cell. Centrifugation was employed to remove cell debris and the tagged-proteins were purified by Ni²⁺-NTA affinity chromatography. EcLsrB and CsLsrB were eluted in 50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl, 250 mM imidazole and subsequently dialyzed (Slide-A-Lyzer, ThermoFisher Scientific) into 25 mM Tris buffer pH 8.0, 150 mM NaCl, 1 mM DTT. CLPY-LuxP, EcLsrB and CsLsrB have a molecular weight of approximately 98, 40 and 37.5 kDa, respectively. The higher molecular weight of CPLY-LuxP when compared with the other receptors is due to its fusion to CFP and YFP.

3.5.3. LuxP pull-down using the TentaGel based probe

Approximately 120 mg of TentaGel resin bound AI-2 (TG-AI-2; approximately 25 μmol of AI-2) was incubated with 2 mg of CLPY (~20.4 nmol), overnight at 4 °C with gentle shaking, followed by a 30 min incubation at room temperature. A centrifugation at 4000 rpm for 10 min was performed to remove the unbound sample. Next, three washes to remove the remaining unbound sample and non specifically bound proteins were performed with 1 mL of FRET buffer (25 mM sodium phosphate buffer pH 8.0, 35 mM NaCl and 1 mM boric acid) and centrifugations at 4000 rpm for 10 min. To collect the bound protein,

denaturing conditions were used. 500 μ L of loading buffer (240 mM Tris-HCl pH 6.8, 40% glycerol, 8% SDS, 0.04% bromophenol blue and 5% β -mercaptoethanol) were added to the sample and the resulting mixture was briefly vortexed and incubated at room temperature for 10 min. A final centrifugation at 4000 rpm, 10 min was performed to collect the specifically bound protein. As a control, the same procedure was followed for 2 mg of CLPY incubated with 5 mg of TentaGel resin (TG; Sigma-aldrich) and 5 mg of TG with ruthenium, a final contaminant of the synthesis process. To 10 μ L of the unbound sample collected before the washes and to 10 μ L of CLPY 2 mg/mL, 5 μ L of loading buffer and 5 μ L of MilliQ H₂O were added. These samples and 20 μ L of the elutions were heated at 99 °C for 5 min and were added to a 10%acrylamide SDS-page gel. 5 μ L of (precision plus protein all blue standards, Bio-Rad) were used as protein standards. The gel was run at 180 V for approximately 70 min at room temperature. In an effort to determine if non-specific interactions were responsible for the low binding of CLPY to TG-AI-2, the protocol described above was repeated but an extra step was added. 120 mg of TG-AI-2 and of TG alone were each incubated in 1 mL of FRET buffer supplemented with 3% bovine serum fraction V (Calbiochem, Merck) for 1h at room temperature with gentle shaking to block unspecific binding sites. To remove the blocking solution a centrifugation was made and 1 mL of FRET buffer with 2 mg of CLPY was added to each condition for overnight incubation.

3.5.4. FRET assays to determine LuxP viability

To determine the viability of LuxP, a FRET assay was employed, similarly to what has been done previously (157,172). Briefly, the protein that did not bind the TentaGel resins was quantified and diluted to the working concentration of 12.5 μ g/mL in FRET buffer. As a control, CLPY that had not been incubated was also tested. 280 μ L of these proteins were added to 2.5 μ L of different AI-2 concentrations, the resultant fluorescence was measured in a multilabel counter (1420 Victor 3, Perkin Elmer) and FRET ratio was calculated (535/440 nm). To test

the effect of the solvents used in CLPY, 100 μ L of 2 mg/mL CLPY was incubated with 10 μ L of the solvents tested (1:10, approximately the same proportion used during the synthesis), overnight at 4 °C with gentle shaking. After incubation, the incubated CLPY was diluted to 12.5 μ g/mL in FRET buffer and the procedure described above was followed.

3.5.4. Native gel and western-blot

Approximately 125 μ g of partially purified CPLY (CFP-LuxP-YFP; $\sim$ 1 nmol), CsLsrB ($\sim$ 3 nmol), and EcLsrB ($\sim$ 3 nmol) were incubated with approximately 59 μ g of D-desthiobiotin-AI-2 ($\sim$ 172 nmol) overnight under gentle agitation at 4°C. These quantities ensured the presence of at least 50 times more probe than receptors. As a control 125 μ g of the protein samples were incubated with D-desthiobiotin alone (59 μ g) or with buffer in the same conditions (FRET buffer for CPLY and 25 mM Tris buffer pH 8.0, 150 mM NaCl, 1 mM DTT for CsLsrB and EcLsrB). 20 μ L of all the samples were mixed with 5 μ L of 5X sample buffer (0.31 M Tris-HCl buffer pH 6.8, 0.05% bromophenol blue and 50% glycerol) and 11 μ L of these mixes were loaded into each of two 10% polyacrylamide non-denaturing gels. 5 μ L of protein standards (precision plus protein all blue standards, Bio-Rad) were also added. The gel was run in Tris-glycine buffer (25 mM Tris and 192 mM glycine, pH 8.3), at 160 V, 4°C for approximately 1h30. One of the gels was stained with Coomassie blue and the presence of the bands assessed. The second gel was used for a western-blot. Transfer was performed at 90 mA, 4°C, overnight in 25 mM Tris, 192 mM glycine and 0.05% SDS to a PVDF membrane. The membrane was incubated in a 1:5000 dilution of an anti-biotin, HRP conjugated antibody (PA1-30595, Invitrogen) for 1h at room temperature. An ECL chemiluminescent substrate was added (SuperSignal West Pico, Pierce) and signal detection was performed in a blot imager (Amersham Imager 680).

3.5.5. LuxP pull-down using the DB-AI-2 probe

Magnetic Dynabeads with protein G (Invitrogen, 10003D) were conjugated with the anti-biotin antibody (#31852, Pierce) as per the manufacturers instructions. Briefly, 50 μL of Dynabeads (1.5 mg) were added to a 1.5 mL eppendorf and the storage solution was removed through magnetic separation. 10 μg of antibody diluted in 200 μL of wash buffer (PBS buffer pH 7.4), 0.02% Tween-20) was added to the beads and the mixture was incubated for 10 min, at room temperature with agitation. The unbound antibody was then removed by magnetic separation and the beads were washed with 200 μL of wash buffer. Next, the sample was added to the anti-biotin antibody conjugated magnetic beads. The sample used for the pull down of CLPY was incubated following the conditions mentioned above and included approximately 76,5 μg of DB-AI-2 and 265 μg of partially purified CLPY. The mixture sample/antibody-conjugated beads was incubated for 1h at room temperature with agitation followed by three washes with 200 μL of wash buffer to remove unbound protein. CPLY, complexed with the AI-2 probe and retained by the antibody conjugated magnetic beads, was eluted by denaturation. The final volume was heated for 10 min at 70 °C in 20 μL of 50 mM glycine pH 2.8 and 10 μL of SDS loading buffer (240 mM Tris-HCl pH 6.8, 40% glycerol, 8% SDS, 0.04% bromophenol blue and 5% β -mercaptoethanol). 10 μL of the wash and elution steps were mixed with 5 μL of SDS loading buffer and added to a denaturing gel composed of 15% acrylamide in the separating gel and 5% in the stacking gel. As protein standards, 5 μL of precision plus protein all blue standards (Bio-Rad) were loaded. The gel was run in Tris-glycine buffer supplemented with 0.2% (w/v) SDS at room temperature, 180 V for approximately 70 min and stained with Comassie blue.

3.5.6. Assessment of AI-2 production

Frozen stocks of *C. difficile* 630 Δ erm (this strain was kindly provided by Adriano O. Henriques) were revived in BHI plates supplemented with 25 μ g/mL cefoxitin. Bacterial growth was performed in TY medium (30 g/L tryptone and 20 g/L yeast extract) supplemented with MOPS buffer (pH 7.0) and 1 g/L of cysteine-HCl. The pH of the medium was adjusted to 6.8 before filter sterilization. 5 mL of TY MOPS (pH 6.8) was inoculated with one *C. difficile* colony and grown overnight at 37 °C in an anaerobic chamber (PlasLabs, USA) under a gaseous mix of 80% nitrogen, 15% carbon dioxide and 5% hydrogen. The calculated volume of overnight inoculum necessary to achieve approximately OD_{600nm} of 0.05 was added to 25 mL of fresh medium and incubated at 37 °C without agitation in a 100 mL Hungate Schott flask. Culture suspensions were collected at the stated time points for OD_{600nm} measurement (UV-visible spectrophotometer; Helios Delta, ThermoSpectronic) and for detection of AI-2. *E. coli* Δ luxS Δ lsrK (ARO 93) was revived directly from the frozen stock to TY MOPS pH 6.8 and allowed to grow overnight at 37 °C without agitation in the anaerobic chamber. Similarly to what was performed for *C. difficile*, the necessary volume of overnight inoculum to achieve OD_{600nm} of approximately 0.05 was added to 25 mL of fresh medium and incubated at 37 °C without agitation in a 100 mL Hungate Schott flask. Exogenous AI-2 was supplemented to a final concentration of 40 μ M. For AI-2 measurements, culture suspensions were filtered using multiscreen filter plates (Millipore) and the cell-free cultures were frozen at -20 °C, overnight. AI-2 activity assessment was performed in triplicate using a FRET assay, similarly to what has been done previously (157,172). Briefly, CLPY was diluted to the working concentration of 12.5 μ g/mL in FRET buffer. Next, 280 μ L of 12.5 μ g/mL CLPY were added to 2.5 μ L of the cell-free supernatants. The resultant fluorescence was measured in a multilabel counter (1420 Victor 3, Perkin Elmer) and the FRET ratio was calculated (535/440 nm). AI-2 concentration in the cell-free supernatants collected was determined by comparing the results against a calibration curve obtained with

different DPD concentrations. This experiment was repeated on three different days with three different cultures.

3.5.7. Preparation of *C. difficile* cell lysates

C. difficile was grown as described above. To collect a higher amount of cells, besides growing cells in 25 mL of TY MOPS pH 6.8 in 100 mL Hungate Schotts, *C. difficile* was also grown in 62.5 mL of liquid medium in 250 mL Hungate Schotts. The cultures were allowed to grow until an OD_{600nm} of approximately 0.5 that is usually achieved around 5h of growth (mid-exponential phase). To stimulate the expression of AI-2 receptors, growth of *C. difficile* in the presence of 40 µM of AI-2 was also performed. The cells were harvested by centrifugation at 4000 rpm, 4 °C for 5 min, in 50 mL falcon tubes. The pellets were washed, twice, with 50 mL of PBS buffer and upon removal of buffer, stored at -80 °C. To prepare the cell lysates, the desired number of pellets (considering that each one corresponds to 50 mL of bacterial growth) was solubilized in 1 mL of 10 mM Tris-HCl pH 6.8, 10 mM MgCl₂, 0.5 mM EDTA, 0.2 mM NaCl, 1 mM DTT, 10% glycerol and 1 mM PMSF. The obtained suspension was passed three times through a French pressure cell and the lysed cells were subjected to a 15 min centrifugation, at 14 000 rpm, 4 °C to remove cell debris. If the cell lysates were not used immediately, they were stored at -20 °C.

3.5.8. Attempts to pull down *C. difficile* AI-2 receptor

Cell lysates of 50 mL of *C. difficile* were incubated with 60 µg of DB-AI-2 (~175 nmol) for 2 h at 4 °C with agitation. This sample was then added to 50 µL of protein G magnetic beads (Invitrogen, 10003D) previously conjugated with 15 µg of anti-biotin antibody (#31852, Pierce). The remaining pull-down protocol was performed as described in the section 3.5.5. 10 µL of all the samples were incubated with 5 µL of SDS loading buffer and ran in a 15% acrylamide denaturing gel for approximately 70 min, at 180 V, room temperature. 5 µL of

precision plus protein all blue standards (Bio-Rad) were used as protein standards.

Cell lysates of 100 mL of *C. difficile* grown in the presence of AI-2 were incubated with 60 µg of DB-AI-2 (~175 nmol) for 2 h at 4 °C with agitation. This sample was then added to 50 µL of streptavidin magnetic particles (Roche) and incubated for 30 min at room temperature with agitation. The unbound sample that should correspond to our sample without biotinylated proteins was collected. This unbound sample was then incubated with the antibody-conjugated protein G magnetic beads like described above in section 3.5.5. An extra step was added before the denaturing elution, in which 998 nmol of AI-2 were incubated with the complex magnetic beads and bound protein for 30 min at room temperature with agitation. 10 µL of all the samples were incubated with 5 µL of SDS loading buffer and ran in a 15% acrylamide denaturing gel for approximately 70 min, at 180 V, room temperature. 5 µL of SigmaMarker wide range was used as protein standards.

3.6. Acknowledgments

The work performed in Rita Ventura's laboratory was supported by Fundação para a Ciência e Tecnologia (FCT) and MostMicro Research Unit, financially supported by LISBOA-01-0145-FEDER-007660 funded by FEDER funds through COMPETE2020 (POCI) and by national funds through FCT. Vanessa Miranda acknowledges FCT for individual PhD grant SFRH/BD/102159/2014. The work performed in the laboratory of Karina B. Xavier was supported by Portuguese National funds from FCT, project PTDC/BIA-MIC/4188/14. Karina B. Xavier and Inês M. Torcato acknowledge FCT for individual grants IF/00831/2015 and PD/BD/105736/2014, being the latter within the scope of the PhD program Molecular Biosciences PD/00133/2012. We thank Paulo Duarte for helpful tips regarding the immunoprecipitation protocol, Mónica Dias for generously providing the magnetic Dynabeads, Adriano O. Henriques and Stephen Miller for

helpful discussions. Moreover, the authors thank Adriano O. Henriques for kindly providing the *C. difficile* 630 Δ erm strain used in this study and Sara Ramalhete, Aristides Mendes and Bruno Dupuy for helpful tips regarding the storage and growth procedures for *C. difficile*.

3.7. Supplementary information

3.7.1. Figures

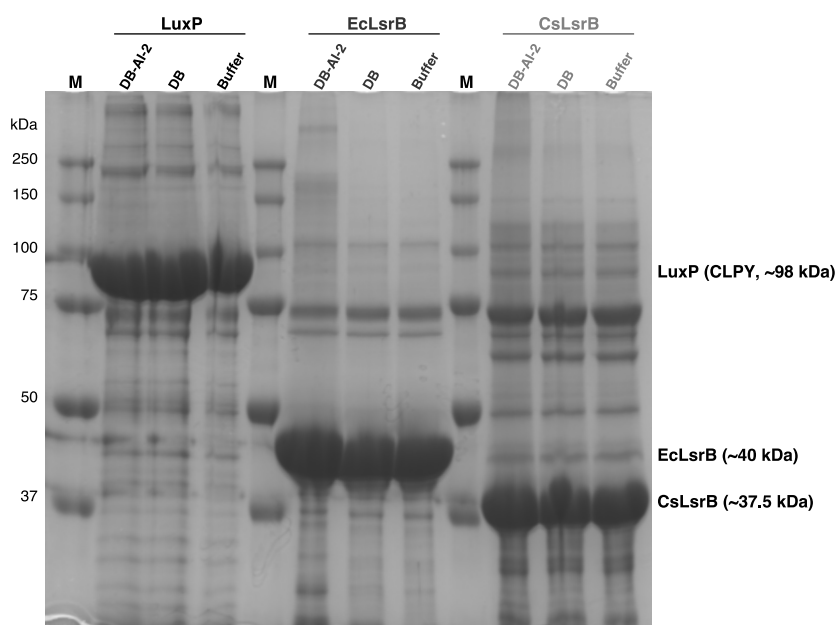


Figure S3.1 - Loading control of the samples ran in the native gel. The samples were heated at 95°C for 5 min and loaded into a denaturing gel composed of 15% acrylamide in the separating gel and 5% in the stacking gel. 5 μ L of precision plus protein all blue standards (Bio-Rad) were loaded as protein standards. The gel was run in SDS running buffer at room temperature, 180 V for approximately 70 min and stained with Coomassie blue. M: protein standards; DB-AI-2: receptor proteins incubated with D-desthiobiotin bound AI-2; DB: receptor proteins incubated with D-desthiobiotin; Buffer: receptor proteins incubated with buffer; CsLsrB: non-canonical LsrB from *C. saccharobutylicum* and EcLsrB: canonical LsrB receptor from *E. coli*.

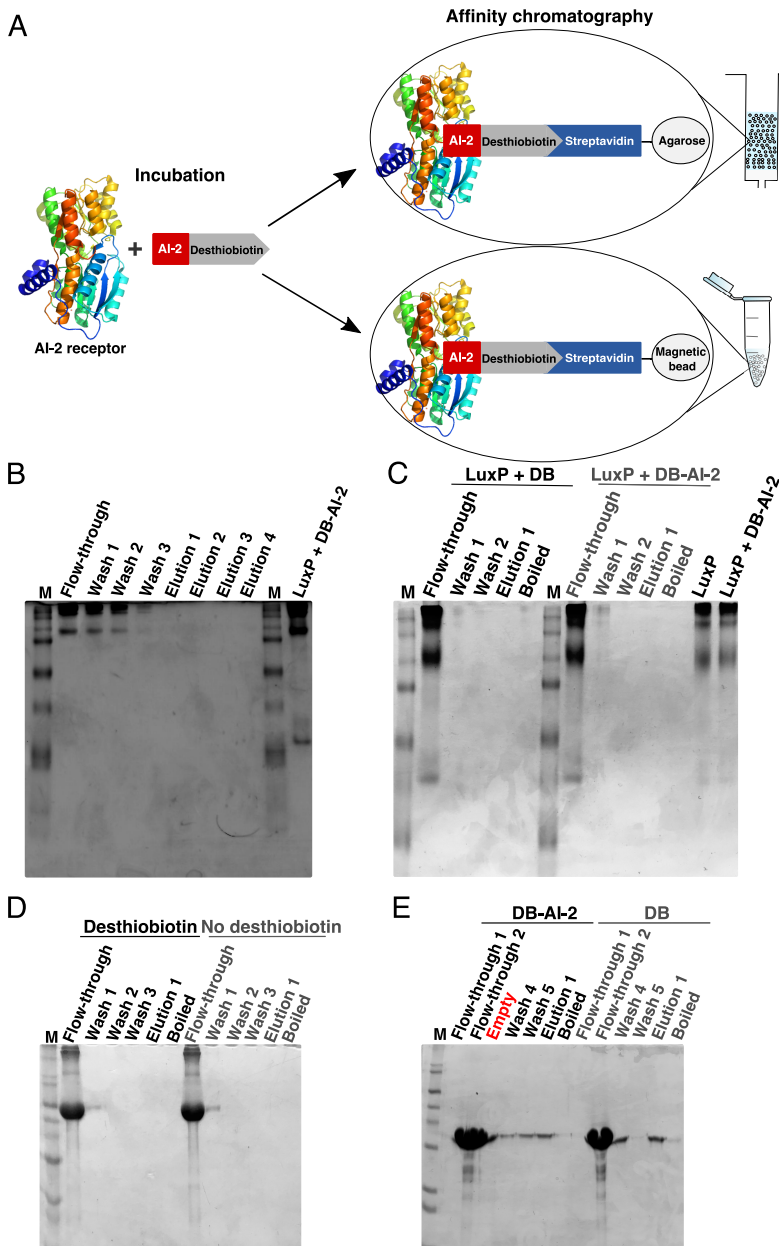


Figure S3.2 - Attempts of pulling down AI-2 receptors using streptavidin. (A) Scheme of the approaches tested using streptavidin bound to either agarose or magnetic beads. (B) Native PAGE gel from the purification of LuxP using streptavidin-agarose. (C) Native PAGE gel from the purification of LuxP using streptavidin-magnetic beads. (D) SDS-PAGE gel of the purification steps of anti-desthiobiotin antibody after sequential incubation of commercial D-desthiobiotin and anti-desthiobiotin antibody with streptavidin magnetic beads. As a control, buffer was added instead of D-desthiobiotin in one of the samples. (E) SDS-PAGE gel of the purification steps of CsLsrB after sequential incubation with DB-AI-2 or DB (control) and CsLrsB. Flow-through: unbound sample; Wash: washes to remove unbound sample; Elution: sample dislocated by biotin; Boiled: sample

dislocated by biotin and heated at 99 °C for 5 min. In E, flow-through 1 is the unbound sample after incubation with the D-desthiobiotin bound AI-2 (DB-AI-2) or D-desthiobiotin (DB) and flow-through 2 is the unbound sample after incubation with CsLsrB.

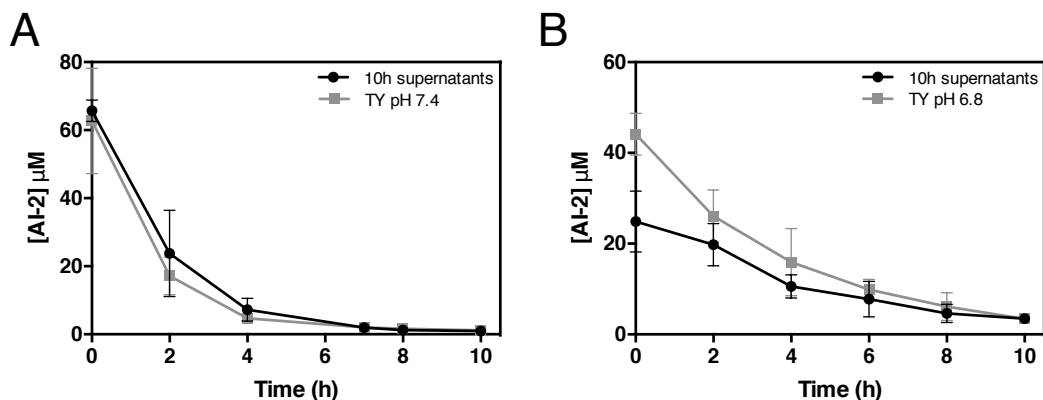


Figure S3.3 – Stability of AI-2 incubated at different pH values. (A) AI-2 concentration after incubation of 40 μM of AI-2, at time zero, in TY medium at pH 7.4 and in cell-free supernatant collected from a stationary phase culture of *C. difficile* (10h) grown in TY medium at pH 7.4. (B) AI-2 concentration after incubation of 40 μM of AI-2, at time zero, in TY-MOPS medium at pH 6.8 and in cell-free supernatant collected from a stationary phase culture of *C. difficile* (10h) grown in TY-MOPS medium at pH 6.8. Incubated mixtures were collected at the indicated time points and AI-2 concentration was assayed using a FRET assay. Error bars represent the standard deviation of three technical replicates.

3.7.2. Experimental procedures

3.7.2.1. Attempt to pull-down LuxP using streptavidin bound to agarose

235.5 μg of DB-AI-2 were incubated with 815 μg of CLPY (CFP-LuxP-YFP) in FRET buffer (25 mM of sodium phosphate buffer (pH 8.0), 35 mM NaCl, and 1 mM boric acid) to a final volume of 200 μL, overnight at 4 °C with gentle agitation. This sample was then added to 2 mL of streptavidin-agarose (Sigma-aldrich), previously equilibrated with 15 mL of FRET buffer. Three washes with the mentioned buffer were performed to remove unbound sample. Three elutions of 1 mL each were performed using FRET buffer supplemented with 2.5 mM of biotin (Fisher Scientific). To be sure that all the D-desthiobiotin was displaced from the resin, an extra elution was performed using 5 mL of FRET buffer supplemented

with 2.5 mM of biotin. 3 μ L of native loading buffer (62.5 mM Tris-HCl (pH 6.8), 25% glycerol and 1% bromophenol blue) were added to 20 μ L of each sample. From this mixture, 10 μ L were added to each of two 15% acrylamide non-denaturing gels. One gel was stained with Comassie blue (Fig. S3.2, B) and the other was used for western-blot detection using an anti-biotin antibody able to recognize D-desthiobiotin (data not shown). 5 μ L of precision plus protein all blue standards (Bio-Rad) were used as protein standards.

3.7.2.2. Attempt to pull-down LuxP using streptavidin bound to magnetic beads

117.5 μ g of DB-AI-2 or DB and 407.5 μ g of CLPY (CFP-LuxP-YFP) in FRET buffer (25 mM of sodium phosphate buffer (pH 8.0), 35 mM NaCl, and 1 mM boric acid) to a final volume of 75 μ L, were incubated overnight, at 4 °C, with gentle agitation. Next, this sample was incubated with 50 μ L of streptavidin magnetic beads (Roche) at room temperature, for 2h with agitation. Three washes with 250 μ L of FRET buffer were performed to remove unbound sample. Elution was performed by adding 75 μ L of FRET buffer supplemented with 2.5 mM of biotin to the magnetic beads. The mixture was vortexed briefly and incubated with shaking for 30 min at room temperature. Additionally, 75 μ L of FRET buffer supplemented with 2.5 mM of biotin were again added to the magnetic beads and the mixture was heated for 5 min at 99 °C to ensure complete removal of bound sample. 3 μ L of native loading buffer were added to 20 μ L of each sample. From this mixture, 10 μ L were added to a 10% acrylamide non-denaturing gel. Also, 5 μ L of the controls were added to the gel. These controls were prepared by adding 5 μ L of LuxP or 5 μ L of LuxP incubated with DB-AI-2 to 5 μ L of MiliQ water plus 3 μ L of the native loading buffer. The gel was stained with Comassie blue (Fig. S3.2, C) and 5 μ L of precision plus protein all blue standards (Bio-Rad) were used as protein standards.

3.7.2.3. Binding of commercial D-desthiobiotin to streptavidin magnetic particles

50 μ L of streptavidin magnetic particles (Roche) were incubated with either 23.5 μ g of D-desthiobiotin (Sigma-aldrich) in 50 μ L of FRET buffer or with 50 μ L of FRET buffer for 2h, at room temperature with agitation. Next, 10 μ g of anti-biotin antibody able to recognize D-desthiobiotin (PA1-30595, Invitrogen) was added to the mixture and incubated for 1h at room temperature with agitation. Three washes with 250 μ L of FRET buffer were performed to remove unbound antibody and elution was performed by adding 50 μ L of FRET buffer supplemented with 2.5 mM of biotin to the magnetic beads. The mixture was vortexed briefly and incubated with shaking for 30 min at room temperature. Additionally, 50 μ L of FRET buffer supplemented with 2.5 mM of biotin were again added to the magnetic beads and the mixture was heated for 5 min at 99 °C to ensure complete removal of bound sample. 10 μ L of each sample was incubated with 2 μ L of SDS loading buffer (240 mM Tris-HCl (pH 6.8), 40% glycerol, 8% SDS, 0.04% bromophenol blue and 5% β -mercaptoethanol) and heated at 99 °C for 5 min. 5 μ L of these mixtures were loaded into a 15% acrylamide denaturing gel. The gel was run at 180 V, for approximately 70 min, at room temperature and stained with Coomassie blue (Fig. S3.2, D). 5 μ L of precision plus protein all blue standards (Bio-Rad) were used as protein standards.

3.7.2.4. Reversion of the binding order to streptavidin

50 μ L of streptavidin magnetic particles (Roche) were incubated with either 117.5 μ g of DB-AI-2 or DB for 1h, at room temperature with agitation. Three washes with 250 μ L of FRET buffer were performed to remove unbound DB-AI-2 or DB and 407.5 μ g of CsLsrB were added to the beads and incubated for 2h, at room temperature with shaking. Three more washes using 250 μ L of FRET buffer were performed. 100 μ L of FRET buffer supplemented with 2.5 mM of biotin was

added and allowed to incubate for 30 min, at room temperature, with agitation to elute the samples. To ensure that nothing stayed bound to the beads, 100 μ L of FRET buffer supplemented with 2.5 mM of biotin were added to the beads and heated for 5 min at 99 °C. 10 μ L of each sample was incubated with 3 μ L of SDS loading buffer, heated at 99 °C for 5 min and loaded into a 15% acrylamide denaturing gel. The gel was run at 180 V, for approximately 70 min, at room temperature and stained with Comassie blue (Fig. S3.2, E). 5 μ L of SigmaMarker wide range (Sigma-aldrich) were used as protein standards.

3.7.2.5. Determination of the AI-2 stability

C. difficile culture at 10h of growth was collected and filtered using a 0.2 μ m filter (Acrodisc, Pall Life Sciences) originating cell-free supernatant that was frozen at -20 °C. The supernatant or fresh TY MOPS pH 6.8 medium were supplemented with synthetic AI-2 to a final concentration of 40 μ M and incubated for 2h, 4h, 6h, 8h or 10h at 37 °C without agitation in a transparent, flat bottom, 96 well plate (Costar 3596). After incubation, supernatants were frozen at -20 °C, overnight. AI-2 concentration was determined in triplicate using a FRET assay, similarly to what has been done previously (157,172). Briefly, CLPY was diluted to the working concentration of 12.5 μ g/mL in 25 mM of sodium phosphate buffer (pH 8.0), 35 mM NaCl and 1 mM boric acid. 280 μ L of 12.5 μ g/mL CLPY were added to 2.5 μ L of the cell-free supernatants. The resultant fluorescence was measured in a multilabel counter (1420 Victor 3, Perkin Elmer) and the FRET ratio was calculated (535/440 nm). AI-2 concentration in the cell-free supernatants collected was determined by comparing the results against a calibration curve obtained with different DPD concentrations.

Chapter



Evaluation of the biological activity of an optically pure autoinducer-2 agonist

Unpublished data and data published in:

Ascenso O. S. & **Torcato I. M.**, Miguel A. S., Marques J. C., Xavier, K. B., Ventura R. M. and Maycock C. D. (2019) Synthesis and biological activity of a potent optically pure autoinducer-2 quorum sensing agonist. *Bioorg. Chem.* **85** 75-81. DOI: 10.1016/j.bioorg.2018.12.022.

Author contributions:

Inês M. Torcato, M. Rita Ventura, João C. Marques and Karina B. Xavier conceptualized the study. M. Rita Ventura, Osvaldo S. Ascenso, A. Sofia Miguel and Christopher D. Maycock designed the synthetic route. Osvaldo S. Ascenso synthesized the analogs studied. Inês M. Torcato tested the analogs for their stability and activity.

4.1 Abstract

Quorum sensing is a cell-cell communication mechanism that regulates population-dependent bacterial behaviours, such as toxin production, biofilm formation and virulence. Autoinducer-2 (AI-2) is a quorum sensing molecule known to foster inter-species bacterial communication being produced and recognized by distantly related bacterial species. As manipulation of quorum sensing has been regarded as a possible alternative to conventional antibiotics, recognition of AI-2 by different bacteria makes this system an attractive target for manipulation. An important step to implement this strategy is the development of AI-2 analogs and the subsequent study of their activity. In this work, pure enantiomers of 4-propoxy-HPD and 4-ethoxy-HPD, known AI-2 analogs, synthesized by our collaborators in the Bioorganic Chemistry laboratory, were tested for their activity and stability. The (4*S*) enantiomer of 4-propoxy-HPD was the most active compound being approximately twice as efficient as AI-2 and ten-times more potent than the (4*R*) enantiomer. It is, also, slightly more stable than AI-2 at higher pH values. Additionally, the specificity of this analog to bacteria with LuxP receptors makes it a good candidate for clinical applications, because it is not susceptible to scavenging by LsrB-containing bacteria that degrade the natural AI-2. All in all, this work provides a study of the activity of the isomerically pure analogs of the most active AI-2 agonist described so far.

4.2. Introduction

Cell-to-cell communication between bacteria, known as quorum sensing allows bacteria to sense the critical density of their population, and consequently synchronize their behavior in order to adapt to changing environmental conditions (2). Quorum sensing is mediated by signaling molecules called autoinducers. Among autoinducers, AI-2 is recognized by many bacterial species, allowing bacteria from one species to sense and respond to the signal produced by

other species (21,50,80). AI-2 comprises a family of molecules that exist in equilibrium and which can be rapidly interconverted in aqueous solution (46,48). AI-2 is produced as the linear molecule (S)-4,5-dihydroxypentanedione ((4S)-DPD) in a reaction catalyzed by LuxS. In aqueous solution, DPD exists in a dynamic equilibrium with cyclic forms that ultimately form (2R,4S)-THMF and (2S,4S)-THMF-borate (46,51). These cyclic forms are the active AI-2 molecules recognized by the two types of AI-2 receptors known: LsrB and LuxP, respectively (46,47). AI-2 has been shown to regulate important bacterial processes, such as biofilm formation, toxin production and virulence, which are implicated in human bacterial infections (50). The fact that quorum sensing modulates organized mechanisms of offense (as expression of virulence factors) and defense (as biofilm formation) in bacteria, and the increasing resistance to conventional antibiotics, has heightened the interest in manipulating quorum sensing as an alternative antibiotic strategy (24). Moreover, as quorum sensing modulation does not suppress growth or induce killing, the selective pressure to develop resistance to this strategy is thought to be smaller (24). Thus, the discovery of new molecules for quorum sensing modulation has a huge potential for developing new strategies to manipulate bacterial behaviors and treat bacterial infections.

Several analogs of AI-2 have been described with some synthesized both in optically pure form and as racemic mixtures (157,176-181). It has been shown that small structural differences in the analogs greatly influence their efficacy as quorum sensing modulators (180). Actually, the configuration of DPD itself is very important for quorum sensing activity as its (4R) enantiomer is approximately 100 and 6 times less potent than the natural (4S)-DPD, for LuxP or LsrB responses, respectively (51,182). Accordingly, the (4S) isomers of DPD analogs with a new stereocenter at C5 were stronger agonists for LuxP receptors than their (4R) isomers. This reinforces the influence of the (4S) configuration in the quorum sensing activity of DPD analogs (157).

More recently, Tsuchikama et al reported a new family of DPD analogs, the C4-alkoxy-HPDs (181). The quorum sensing modulator activity of these analogs was tested using two reporter systems for the AI-2 receptors known: LsrB and LuxP. Interestingly, these analogs seemed to have specific agonist activity for LuxP receptors, as no activity in the LsrB reporter assay was detected (181). However, it was not possible to conclude if the analogs were able to bind LsrB because the *lsr*-dependent β -galactosidase reporter assay measures the expression of the *lsr* operon and not direct binding to LsrB. Intracellular phosphorylation of AI-2 by the kinase LsrK triggers induction of expression of the *lsr* operon through binding of AI-2-phosphate to the repressor protein (52,67). The authors observed that the C4-hydroxyl group is a critical factor for the LsrK-mediated phosphorylation, rendering the C4-hydroxy analogs unable to be phosphorylated (181). Thus, the lack of phosphorylated DPD alone can be accountable for the lack of activity for the β -galactosidase LsrB-reporter assay. From the analogs tested, 4-ethoxy-5-hydroxy-2,3-pentadione (4-ethoxy-HPD) and 4-propoxy-5-hydroxy-2,3-pentadione (4-propoxy-HPD) were the most potent quorum sensing agonists for LuxP mediated responses and, significantly, were even more potent than DPD, with racemic 4-propoxy-HPD (rac-4-propoxy-HPD) exhibiting the greatest LuxP-dependent quorum sensing activity (181).

In pharmacology, it has been established that the use of the eutomer (more active isomer) rather than the racemic mixture is advantageous as only half the concentration is needed and there is no unbound distomer (less active isomer) free to act on other targets (183,184). So, our collaborators developed an optimized synthesis process for the pure (4*S*) enantiomers of 4-propoxy-HPD and 4-ethoxy-HPD analogs and we characterized their activity. The optically pure enantiomer (4*S*)-4-propoxy-HPD was the most efficient compound having approximately double the activity of (4*S*)-DPD and being slightly more stable at higher pH values. Significantly, unlike DPD, this agonist is not processed by LsrB-harbours

bacteria, being a good clinical candidate for quorum sensing manipulation of bacteria with LuxP receptors.

4.3. Results

4.3.1. Chemical synthesis

The group of Rita Ventura had previously developed an efficient synthesis route for DPD (51). For this work, they adapted the synthesis of the most potent LuxP agonist known - 4-propoxy-HPD (181) - to obtain the pure enantiomers. Moreover, to show that this synthetic route could also be employed to other AI-2 analogs, they synthesized the pure enantiomer of a similar, although less potent, LuxP agonist: 4-ethoxy-HPD (181). From previous studies it was known that the enantiomer (4*S*) should be the most active enantiomer (51,157,182) and thus, for 4-ethoxy-DPD only this enantiomer was synthesized. The chemical structure of the (4*S*) enantiomers of 4-propoxy-HPD and 4-ethoxy-HPD are presented on Fig. 4.1.

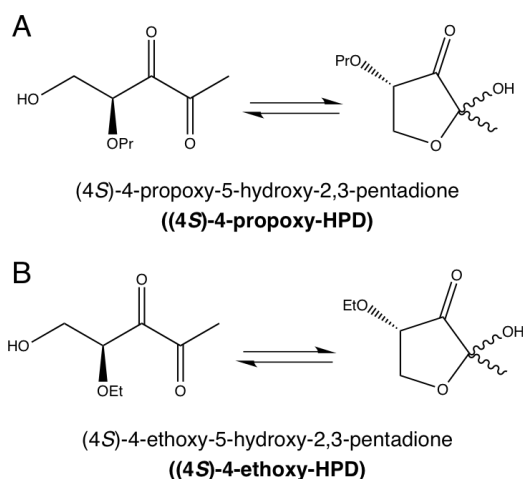


Figure 4.1 – Chemical structures of the more active enantiomers of the AI-2 analogs. Linear and cyclic structure of (A) (4*S*)-4-propoxy-5-hydroxy-2,3-pentadione abbreviated to (4*S*)-4-propoxy-HPD and (B) (4*S*)-4-ethoxy-5-hydroxy-2,3-pentadione abbreviated to (4*S*)-4-ethoxy-HPD.

4.3.2. Binding affinity of the DPD analogs to LuxP

The binding affinity of (4*S*)-4-propoxy-HPD, (4*R*)-4-propoxy-HPD, rac-4-propoxy-HPD and (4*S*)-4-ethoxy-HPD compounds was assessed and compared to that of (4*S*)-DPD. To assess the affinity of these compounds to the LuxP receptor, a LuxP-FRET assay was performed. This assay uses a LuxP receptor that has a Cyan Fluorescent Protein (CFP) and a Yellow Fluorescent Protein (YFP) fused to each of its termini. A decrease in fluorescence resonance energy transfer is detected upon ligand binding to this modified LuxP, due to the increase in distance between the two termini (171,185). The lowest the concentration able to induce this decrease, the better is the compound at binding LuxP.

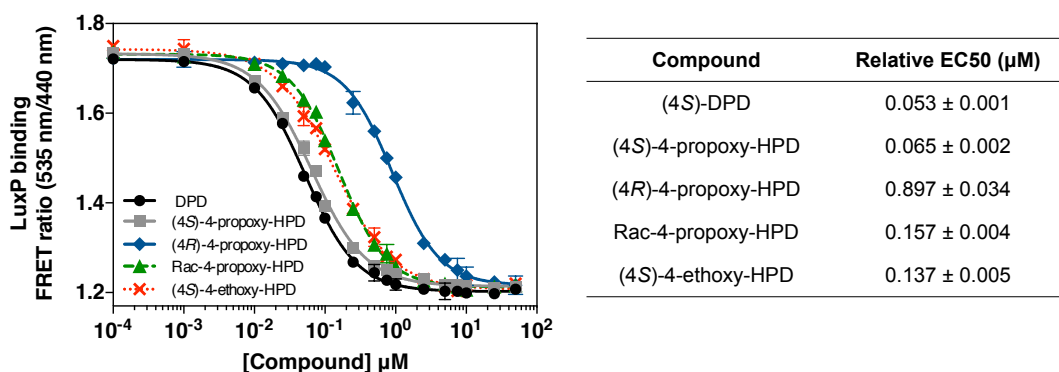


Figure 4.2 - Affinity of DPD and DPD analogs to LuxP. Values of half maximal effective concentration (EC₅₀) were obtained from the fitted curves presented. A representative of two independent experiments is shown. Additional fitting parameters, and data from the repeated experiment, are shown in Table S4.1. Error bars represent the standard deviation of three technical replicates.

Our results showed that (4*S*)-DPD and (4*S*)-4-propoxy-HPD were the compounds with the highest affinity for LuxP (Fig. 4.2). The response curves and half maximal effective concentration (EC₅₀) values for these compounds were very similar, showing that their affinity to bind LuxP was comparable. (4*R*)-4-propoxy-HPD was the compound with lowest affinity having an EC₅₀ more than ten-times higher than its (4*S*) enantiomer. This observation showed that the R-

configured propoxy group on C4 was less favorable for binding to LuxP. Rac-4-propoxy-HPD had an affinity in between the two enantiomers, as expected. (4S)-4-ethoxy-HPD also had high affinity for LuxP, with an EC₅₀ value (0.137 ± 0.005) very similar to the one obtained for rac-4-propoxy-HPD (0.157 ± 0.004) but higher than the one obtained for (4S)-DPD (0.053 ± 0.001).

4.3.3. Biological activity of the DPD analogs

To assess the activity of the studied compounds in bacterial cells, a *V. harveyi* bioassay was carried out. In *V. harveyi*, luminescence is produced in response to increasing concentrations of DPD. The MM32 *V. harveyi* strain, used in this assay, is a mutant that does not produce DPD, being, therefore, a good sensor to test the activity of DPD and its analogs. If an analog is able to induce higher levels of light production at lower concentrations than DPD, then it is more active than the natural ligand.

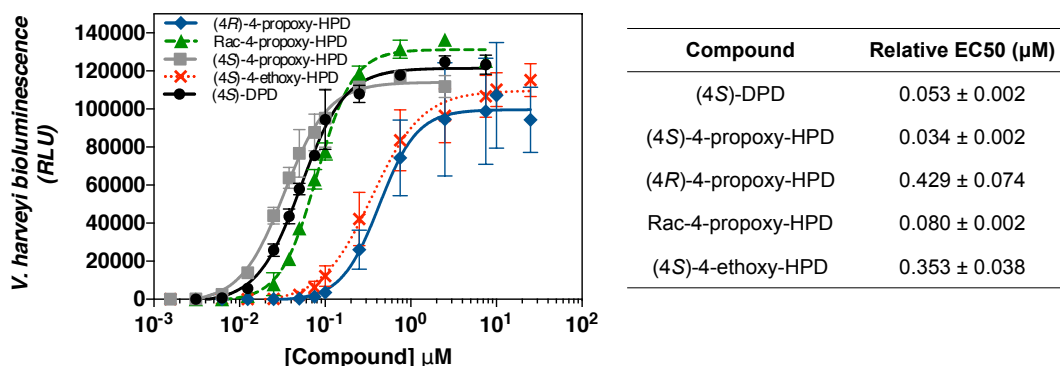


Figure 4.3 - *In vivo* activity of the compounds tested. Bioluminescence produced by *V. harveyi* MM32 was measured in the presence of different concentrations of the tested compounds. A representative of 3 independent experiments is shown. The EC₅₀ values were obtained from the fitted curves presented. Additional fitting parameters, and results from the repeats of this experiment, are shown in Table S4.2. Error bars represent the standard deviation of three technical replicates.

Comparing the EC₅₀ values obtained for all the compounds tested, the most active compound was (4S)-4-propoxy-HPD, being approximately 2 times more

active than (4S)-DPD (Fig. 4.3). (4R)-4-propoxy-HPD was the least active compound. The racemic mixture of this analog had an intermediate efficacy with an EC₅₀ roughly two-times higher than (4S)-4-propoxy-HPD (Fig. 4.3).

4.3.4. Chemical stability

The introduction of a propyl group in the hydroxyl group of the carbon 4 of the natural DPD was proposed to increase the chemical stability of the molecule as it should lower its reactivity as the hydroxyl group is protected. This increased stability could allow the analog to withstand environments with higher pH when comparing with DPD. This property could be useful for its use as a quorum sensing modulator in niches with pH higher than 7, as the mammalian gut microbiota for example. To test this hypothesis we incubated 40 μ M (4S)-DPD and (4S)-4-propoxy-HPD in MOPS buffer at pH 7 (Fig. 4.4, A) and pH 8 (Fig. 4.4, B). It was possible to observe that at pH 8 (4S)-4-propoxy-HPD was slightly more stable than (4S)-DPD, as a higher concentration of analog than (4S)-DPD was detected at 6h and onward. At pH 7, however, the decrease in the concentration of the compounds was very similar. Additionally, we wondered if the presence of bacterial cells would have an effect in the stability of these compounds. We added (4S)-DPD and (4S)-4-propoxy-HPD to a final concentration of 40 μ M to a culture of *E. coli* KX1372, a mutant for AI-2 synthase (LuxS) and kinase (LsrK) and thus, unable to produce or internalize AI-2. The use of this strain guarantees that the compounds detected are only the ones added, without interference of naturally produced AI-2 or internalization of added AI-2. To guarantee different pH values, we grew *E. coli* in LB supplemented with 100 mM MOPS buffer at pH 7 (Fig. 4.4, C) or pH 8 (Fig. 4.4, D). Moreover, we grew *E. coli* in LB without buffer (Fig. 4.4, F) to assess the effect of bacterial metabolism in pH and therefore in the stability of the compounds. Again, we compared this result to *E. coli* grown in LB supplemented with 100 mM MOPS buffer at pH 7 (Fig. 4.4, E). Measurement of pH at latter time points of the culture in LB without buffer revealed that the pH

reaches approximately 9, whereas in the cultures grown in LB with buffer, the pH remained the same for 8h and increased approximately to 7.5 and 8.5 at 9h of growth. In both conditions, at pH higher than 7, the chemical degradation of (4S)-4-propoxy-HPD seemed slower than the one of (4S)-DPD, indicating that the analog is slightly more stable than DPD.

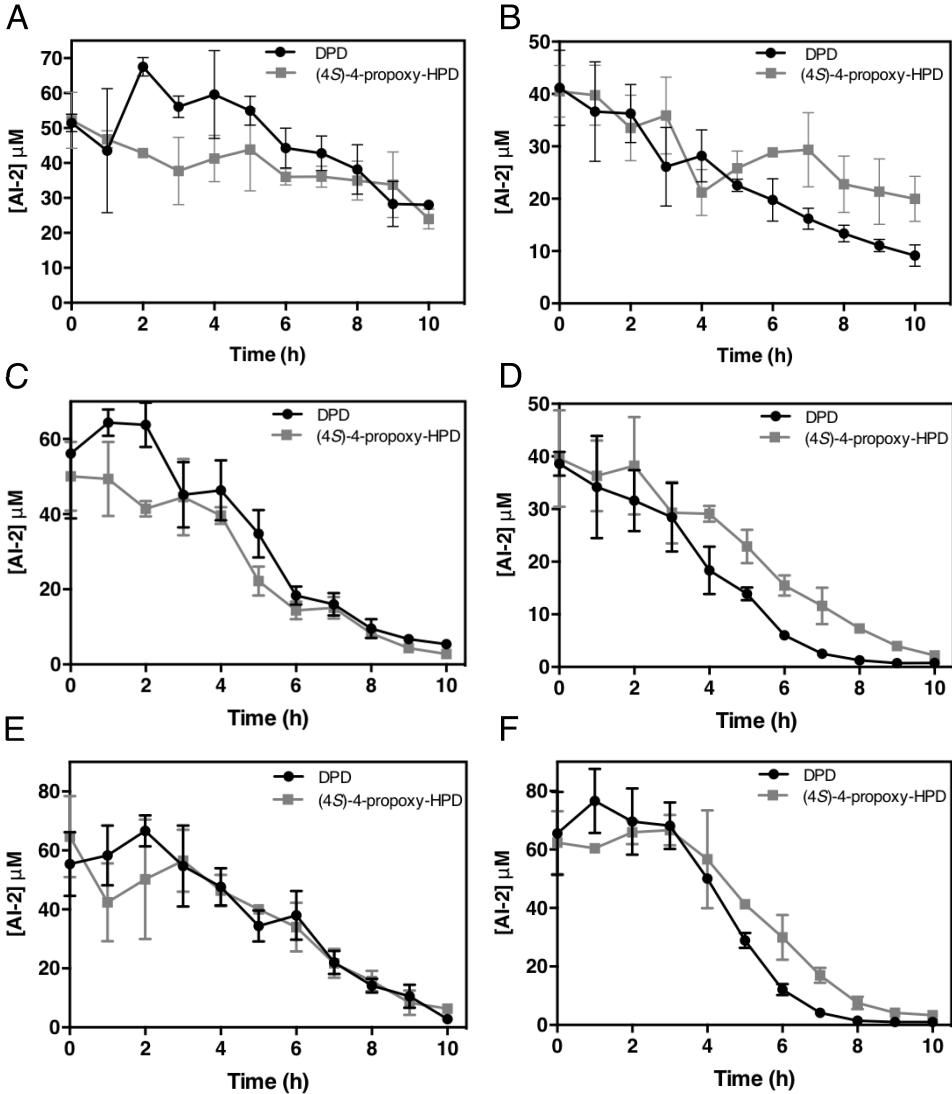


Figure 4.4 - Chemical stability of DPD and (4S)-4-propoxy-HPD *in vitro*. 40 μM of (4S)-DPD and (4S)-4-propoxy-HPD were incubated at 37 °C in 100 mM MOPS buffer pH 7 (A), pH 8 (B) or with a culture of *E. coli* *ΔluxSΔlsrK* grown in LB supplemented with 100 mM MOPS buffer pH 7 (C and E), pH 8 (D) or in unbuffered LB (F). Buffer and cell-free culture fluids were collected at various time

points to determine the concentration of the compounds. Error bars represent the standard deviation of three technical replicates.

Nonetheless, at 10h of growth in all the conditions both compounds were almost undetectable. Also, the chemical degradation of both compounds was more extensive in the presence of bacterial cells. Compound concentration reached values close to zero when incubated with *E. coli* cells but reaches a minimum of 8 μM at pH 8 and of 21 μM at pH 7 when incubated only with buffer (Fig. 4.4, A and B *versus* C-F).

4.3.5. Internalization studies

In bacteria that possess LsrB receptors, DPD increases in the extracellular medium until the concentration of AI-2 required for detection is reached. At this point, DPD is internalized and processed intracellularly (52,71). To trap DPD inside the cell, a kinase - LsrK - phosphorylates DPD. Tsuchikama et al showed that C4-hydroxy analogs cannot be phosphorylated by LsrK, contrarily to DPD (181). This led us to hypothesize that (4S)-4-propoxy-HPD would not be degraded by LsrB-containing bacteria, as it cannot be trapped inside the cell by phosphorylation and should remain in the extracellular medium. To investigate this possibility, we determined the extracellular levels of (4S)-4-propoxy-HPD and (4S)-DPD through time after adding 40 μM of these two compounds to cultures of *Escherichia coli*, a bacterium able to degrade AI-2 via a LsrB-dependent mechanism. To measure the quantities of the compounds added without the interference of DPD produced by the bacteria we used an *E. coli* strain unable to produce DPD (mutant for DPD synthase, LuxS).

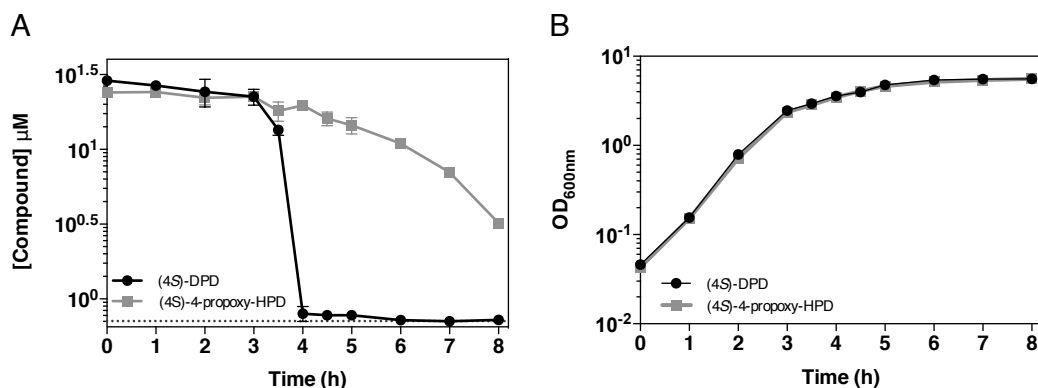


Figure 4.5 - Extracellular concentration of (4S)-DPD and (4S)-4-propoxy-HPD in *E. coli* cultures. Cell-free culture fluids and bacterial suspension were collected at various time points to determine the extracellular concentration of the compounds (A) or the bacterial growth (B). A representative of three independent experiments is shown. Error bars represent the standard deviation of three technical replicates.

Indeed, as previously observed, (4S)-DPD is readily depleted from the extracellular medium as *E. coli* reaches stationary phase (17,52). Contrarily, (4S)-4-propoxy-HPD remains in the extracellular medium (Fig. 4.6). These results show that, unlike DPP, the levels of (4S)-4-propoxy-HPD are not significantly altered by bacteria with LsrB receptors.

4.4. Discussion

The study of the activity of the pure isomers and of the racemic 4-propoxy-HPD supports the previous observations that the stereochemical configuration of DPD analogs is very important for quorum sensing activation. The LuxP binding assay showed that the R-configured propoxy group on C4 was less favorable for binding to LuxP. Surprisingly, (4S)-4-ethoxy-HPD had an EC₅₀ value (0.137 ± 0.005) higher than the one obtained for (4S)-DPD (0.053 ± 0.001) and very similar to the one obtained for rac-4-propoxy-HPD (0.157 ± 0.004). This was not expected since in the work of Tsuchikama et al (181) the biological activity of the racemic mixture of this C4-ethoxy analog was slightly higher than that of (4S)-DPD and lower than rac-4-propoxy-HPD. However, as no affinity measurement with the

purified receptor was performed in that study, no direct comparisons can be made.

The EC₅₀ values obtained in the LuxP-FRET and *V. harveyi* assays were different (compare values in Fig. 4.2 and Fig. 4.3). Nevertheless, the eudismic ratio, that reflects the difference in activity between the two enantiomers, is very similar in the two assays (13.8 ± 0.1 for LuxP-FRET and 12.6 ± 1.4 for *V. harveyi*) showing that in both assays the eutomer is approximately 10-fold more active than the distomer. (4*S*)-DPD and (4*S*)-4-propoxy-HPD had a similar affinity towards the LuxP receptor (Fig. 4.2), but the latter seems to have a higher efficacy than (4*S*)-DPD *in vivo* (Fig. 4.3). Contrarily, (4*S*)-4-ethoxy-HPD that had similar affinity with rac-4-propoxy-HPD, was approximately four times less efficient *in vivo*. The difference in the EC₅₀ of these analogs in the two assays are likely to be related to the different properties being assessed in these assays. In LuxP-FRET, only the affinity of the compounds to LuxP is determined, whereas in the *V. harveyi* bioassay we measure not only binding to LuxP, but also the *in vivo* effect of activating the quorum sensing network response to these compounds. Specifically, the EC₅₀ values obtained with the *V. harveyi* assay could be related to different conformations that LuxP might assume upon binding to these different compounds. These potentially different conformations of LuxP could influence the affinity of this receptor to LuxQ, the membrane sensor histidine/kinase, leading to differences in the activation of the quorum sensing response and ultimately differences in the induction of light production. Additionally, possible differences in the stability of these analogs could also affect their *in vivo* efficacy. Incubation of (4*S*)-DPD and (4*S*)-4-propoxy-HPD in the presence of bacteria cells in unbuffered medium revealed that degradation of the analog was slightly slower than the one observed for (4*S*)-DPD (Fig. 4.4, F).

The *V. harveyi* EC₅₀ values for rac-4-propoxy-HPD, (4*S*)-4-propoxy-HPD and (4*S*)-4-ethoxy-HPD are in the expected order of magnitude given the EC₅₀ values obtained for the racemic forms of C4-propoxy-HPD and C4-ethoxy-HPD obtained

by Tsuchikama et al (181). However, the EC₅₀ value obtained for (4S)-DPD in the present study is 20 times lower. Thus, although the analogs give somewhat similar values, our results suggest that these analogs are not so efficient when compared to (4S)-DPD. The differences observed could be due to the use of different quantification methods for the studied compounds. In the present study, NMR was used to try to overcome the errors in mass prediction or incomplete deprotection, as the final compounds are too unstable to be dried and are, therefore, obtained and used in solution. Moreover, there are significant differences among the EC₅₀ reported for DPD in different studies. Tsuchikama et al determined an EC₅₀ of 1.07 μ M for (4S)-DPD but EC₅₀s of 0.87 μ M and 0.25 μ M have also been reported (176,181,186). Furthermore, in our previous work we obtained an EC₅₀ of 0.076 μ M for (4S)-DPD that is in the same order of magnitude as the ones obtained in the present work (157). These differences are also likely to be potentiated by the intrinsic biological variability of the *V. harveyi* response. Hence, we propose that, to compare the efficacy of different analogs, they should be tested on the same day with the same bacterial culture to determine relative biological responses between them. We have measured the activity of all the compounds studied here using the same culture of *V. harveyi* and repeated this assay on three different days. In all three experiments the relation between the EC₅₀ obtained for the different compounds was the same (see results on Table S4.2 and Fig. S4.1). Importantly, our results support the major conclusion that the optically pure (4S)-4-propoxy-HPD is the most potent agonist surpassing the efficacy of DPD.

The observation that, unlike DPP, the levels of (4S)-4-propoxy-HPD are not significantly altered by bacteria with LsrB receptors could be an advantage for clinical applications, as it means that bacteria with LsrB receptors are not able to degrade this analog. Thus, lower concentrations of analog, in comparison with DPD, should be needed to induce a stable and durable quorum sensing activation. Moreover, this analog is slightly more stable than (4S)-DPD at higher pH values.

Incubation of (4S)-DPD and (4S)-4-propoxy-HPD at pH 8 (or even close to 9 in unbuffered LB medium) revealed a trend that suggests that at pH values higher than 7, this analog degrades slower than DPD (Fig. 4.4, B, D and F). This might prove to be of importance to manipulate quorum sensing in pathogens such as *Vibrio cholerae*, for example. *V. cholerae* is the causative agent of cholera in humans and it is known that at high cell densities and hence, at high autoinducers concentrations, virulence and biofilm formation are repressed (3,99,187). So, addition of synthetic autoinducers can decrease virulence in *V. cholerae*. Actually, it has been shown that addition of synthetic CAI-1 (cholera autoinducer-1) decreases the expression of toxic co-regulated pilus, a canonical virulence factor (188). As in *V. cholerae* multiple quorum sensing pathways work in parallel to regulate these behaviors, we propose that the addition of (4S)-4-propoxy-HPD together with CAI-1 would further decrease the expression of virulence factors. Also, the addition of (4S)-4-propoxy-HPD would be more advantageous than the addition of AI-2 because (4S)-4-propoxy-HPD is not labile to degradation by intestinal LsrB-containing bacteria like *E. coli*, and because of its higher efficacy inducing LuxP-mediated quorum sensing responses

4.5. Experimental procedures

4.5.1. Chemical synthesis

(4S)-DPD and the analogs 4-propoxy-HPD and (4S)-4-ethoxy-HPD were synthesized as described in (51,172).

4.5.2. LuxP-FRET assay

In vitro response of LuxP-FRET protein was measured as previously described (51,157), using a method optimized for 96 well plate reading using a multilabel counter (1420 Victor 3, Perkin Elmer). Serial dilutions of test

compounds were performed in MiliQ water and added to 12.5 $\mu\text{g ml}^{-1}$ of CFP-LuxP-YFP chimeric protein in 25mM of sodium phosphate buffer (pH 8.0), 35mM NaCl, and 1 mM boric acid. Samples (2.5 μl) were added to 280 μl of reaction volume and FRET ratio was calculated (535/440 nm). Relative EC50 values were determined by fitting a four parameter logistic model in a custom equation created in GraphPad Prism version 6 (GraphPad Software, La Jolla, California,USA) to the dose-response curves obtained. Fitting parameters for all the curves obtained using LuxP-FRET assays are presented on Table S4.1.

4.5.3. Bioluminescence assay in *Vibrio harveyi*

V. harveyi *in vivo* response was measured using MM32 reporter strain grown in AB (autoinducer bioassay medium; detailed composition in supplementary information) as previously reported (51,157). Serial dilutions of the tested compounds were performed in MiliQ water. Light emission was measured in a GloMax Explorer microplate luminometer (Promega, USA) after 5h of incubation at 30°C. Relative EC50 values were determined by fitting a four parameter logistic model in a custom equation created in GraphPad Prism version 6 (GraphPad Software, La Jolla, California,USA) to the dose-response curves obtained. The concentrations of each compound used were adjusted to have at least two concentrations before and after the linear slope (189). To test the same concentration range for all the compounds, so that we could compare their EC50 values, all the compounds shared at least 9 out of the 12 concentrations. Fitting parameters for all the curves obtained using this bioassay are presented on Table S4.2.

4.5.4. *In vitro* chemical stability testing

For the tests in buffer, 40 μM of DPD or (4S)-4-propoxy-HPD were added to 5 mL of 100 mM MOPS pH 7.0 and pH 8.0 in 50 mL Erlenmeyers and incubated at 37 °C, with shaking at 240 rpm. 80 μL aliquots were collected at the indicated time

points and stored at -20 °C, overnight. For the tests in the presence of cells, *E. coli* KX1372 ($\Delta luxS \Delta lsrK$)(52) was revived in 5 mL of the medium to be used in the experience (LB or LB supplemented with 100 mM MOPS buffer, pH 7.0 or 8.0) and incubated overnight, at 37 °C, with 240 rpm of shaking. Overnight cultures were diluted in 12.5 mL of fresh medium until an OD_{600nm} of approximately 0.05 and (4S)-DPD or (4S)-4-propoxy-HPD were added to a final concentration of 40 μ M. The bacterial cultures were grown at 37 °C, 240 rpm. At the specified time points, 100 μ L of bacterial suspensions were collected for growth assessment (at OD_{600nm}) and for compound detection. pH measurement was also carried out using pH strips (Macherey-Nagel GmbH & Co. KG). For compound detection, bacterial suspensions were filtered through multiscreen filter plates (Millipore) and stored at -20°C, overnight. A LuxP-FRET assay was employed to detect DPD and (4S)-4-propoxy-HPD following the protocol described above. To determine the concentration of compound in the cell-free supernatants, results were compared against a calibration curve obtained with different DPD concentrations.

4.5.5. Internalization studies

E. coli KX1290 ($\Delta luxS$) (52) was grown overnight in LB supplemented with 100 mM MOPS buffer, pH 7.0 at 37 °C, 240 rpm. The remaining procedure is identical to what was described for the chemical stability testing. Overnight cultures were diluted in fresh medium until an OD_{600nm} of approximately 0.05 and 40 μ M of DPD or (4S)-4-propoxy-HPD were added. The cultures were then grown at 37°C, 240 rpm. At the specified time points, bacterial suspensions were collected for growth assessment (at OD_{600nm}) and for compound detection. For compound detection, bacterial suspensions were filtered through multiscreen filter plates (Millipore) and stored at -20°C, overnight. A LuxP-FRET assay was employed to detect DPD and (4S)-4-propoxy-HPD following the protocol described above. To determine the concentration of compound in the cell-free supernatants, results were compared against a calibration curve obtained with different DPD

concentrations.

4.6. Acknowledgments

We thank the Fundação para a Ciência e Tecnologia (FCT) and MostMicro Research Unit (financially supported by LISBOA-01-0145-FEDER-007660 funded by FEDER funds through COMPETE2020 (POCI) and by national funds through FCT, and FCT PTDC/BIA-MIC/4188/2014). The NMR data was acquired at CERMAX, ITQB-NOVA, Oeiras, Portugal with equipment funded by FCT, project AAC 01/SAICT/2016. Inês M. Torcato was funded by the PhD fellowship PD/BD/105736/2014 within the scope of the PhD program Molecular Biosciences PD/00133/2012.

4.7. Supplementary material

4.7.1. Autoinducer bioassay (AB) medium composition

The medium was prepared as previously (12). AB medium consists of 0.3 M NaCl, 0.05 M MgSO₄, and 0.2% (w/v) vitamin-free casamino acids adjusted to pH 7.5 with KOH. After sterilization, the medium was allowed to cool and 10 ml of sterile 1 M potassium phosphate (pH 7.0), 10 ml of 0.1 M L-arginine and 20 ml of glycerol 50% (v/v) per liter were added.

4.7.2. Figures

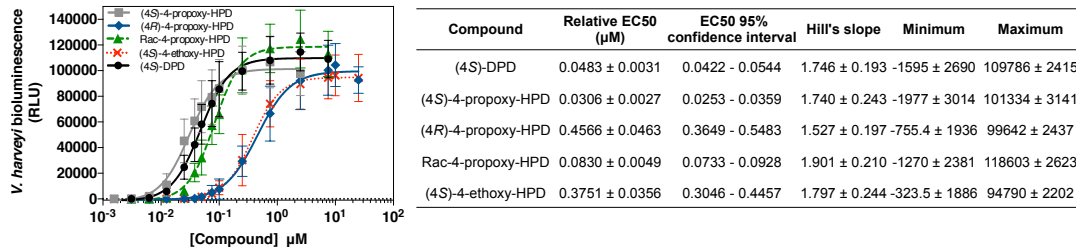


Figure S4.1 - Determination of the EC50 values using the data from the three independent experiments. By joining the data from the three biological replicates (each with technical triplicates) we observe that the EC50 value of DPD is 1.6 times superior to the one of (4S)-4-propoxy-HPD, confirming what we observe in each of the experiments. This reflects the robustness of our results.

4.7.3. Tables

Table S4.1. Fitting parameters for the determination of the relative EC50 of the studied compounds using data from the LuxP-FRET assay.

Compound	Relative EC50 (μM)	EC50 95% confidence interval	Hill's slope	Minimum	Maximum
(4S)-DPD	0.052 ± 0.002	0.048 - 0.056	1.241 ± 0.061	1.211 ± 0.004	1.736 ± 0.006
	0.053 ± 0.001	0.050 - 0.055	1.191 ± 0.033	1.202 ± 0.002	1.720 ± 0.003
(4S)-4-propoxy-HPD	0.071 ± 0.006	0.060 - 0.082	1.417 ± 0.160	1.226 ± 0.008	1.717 ± 0.011
	0.065 ± 0.002	0.061 - 0.069	1.125 ± 0.037	1.214 ± 0.003	1.732 ± 0.004
(4R)-4-propoxy-HPD	0.857 ± 0.053	0.751 - 0.963	1.067 ± 0.066	1.229 ± 0.009	1.769 ± 0.005
	0.897 ± 0.034	0.829 - 0.965	1.293 ± 0.062	1.216 ± 0.006	1.720 ± 0.003
Rac-4-propoxy-HPD	0.154 ± 0.004	0.146 - 0.161	1.228 ± 0.031	1.220 ± 0.003	1.760 ± 0.003
	0.157 ± 0.004	0.149 - 0.165	1.277 ± 0.032	1.212 ± 0.003	1.732 ± 0.003
(4S)-4-ethoxy-HPD	0.147 ± 0.004	0.140 - 0.154	1.149 ± 0.026	1.212 ± 0.002	1.744 ± 0.003
	0.137 ± 0.005	0.127 - 0.146	1.037 ± 0.032	1.207 ± 0.004	1.743 ± 0.004

Table S4.2. Fitting parameters for the determination of the relative EC50 of the studied compounds using data from the *Vibrio harveyi* MM32 assay.

Compound	Relative EC50 (μM)	EC50 95% confidence interval	Hill's slope	Minimum	Maximum
(4S)-DPD	0.038 ± 0.001	0.036 - 0.041	2.052 ± 0.125	-835.1 ± 1550	117120 ± 1319
	0.053 ± 0.002	0.048 - 0.057	1.692 ± 0.119	-1737 ± 1862	121426 ± 1715
	0.062 ± 0.004	0.053 - 0.070	1.597 ± 0.177	-1727 ± 2179	91483 ± 2113
(4S)-4-propoxy-HPD	0.024 ± 0.002	0.021 - 0.028	2.072 ± 0.227	-1824 ± 2540	112679 ± 2421
	0.034 ± 0.002	0.030 - 0.037	1.687 ± 0.148	-1762 ± 2105	114026 ± 2276
	0.039 ± 0.002	0.035 - 0.043	1.556 ± 0.127	-1895 ± 1340	77879 ± 1539
(4R)-4-propoxy-HPD	0.444 ± 0.056	0.330 - 0.559	2.230 ± 0.463	-205.7 ± 2756	104781 ± 3478
	0.429 ± 0.074	0.278 - 0.580	1.956 ± 0.516	-689.7 ± 3565	99665 ± 4388
	0.545 ± 0.096	0.349 - 0.741	2.152 ± 0.695	-252.3 ± 2601	70788 ± 3406
Rac-4-propoxy-HPD	0.076 ± 0.004	0.068 - 0.084	1.815 ± 0.171	-2028 ± 2365	131047 ± 2508
	0.080 ± 0.002	0.076 - 0.084	2.055 ± 0.107	-850.4 ± 1202	131249 ± 1312
	0.099 ± 0.005	0.090 - 0.109	1.871 ± 0.148	-990.4 ± 1299	93572 ± 1535
(4S)-4-ethoxy-HPD	0.296 ± 0.039	0.217 - 0.374	1.955 ± 0.418	-566.8 ± 2925	95788 ± 3184
	0.353 ± 0.038	0.275 - 0.430	1.474 ± 0.192	-1704 ± 2473	109617 ± 2763
	0.528 ± 0.046	0.434 - 0.622	3.022 ± 0.578	762.6 ± 1424	78542 ± 1865

Chapter



**Detection and quantification of the
quorum sensing signal
autoinducer-2 in mice caecum**

Author contributions:

Inês M. Torcato, Vanessa Miranda, M. Rita Ventura and Karina Xavier conceptualized the study. Osvaldo Ascenso suggested the ethyl acetate extraction of derivatized AI-2 and synthesized the tag used to derivatize AI-2. A. Rita Oliveira and Jessica A. Thompson colonized and collected the caecum of mice used to prepare samples. M. Rosário Bronze, Noélia Duarte, Andreia Bento-Silva and Joana Q. Carrola established the liquid chromatography - tandem mass spectrometry (LC-MS/MS) method and Joana Q. Carrola ran the last batch of samples. Vanessa Miranda helped with the preparation of the samples. Inês M. Torcato optimized the protocol for sample preparation and prepared the samples. Inês M. Torcato and Vanessa Miranda ran the samples in the LC-MS equipment.

5.1. Abstract

Autoinducer-2 (AI-2) is a quorum sensing autoinducer signal involved in interspecies communication that comprises a group of molecules derived from the linear (S)-4,5-dihydroxypentane-2,3-dione (DPD). These isomers exist in equilibrium in aqueous solution, being rapidly interconvertible. Recently, it was shown that AI-2 manipulation in the mouse intestines influences the species composition of the microbiota in the mouse gut. Increasing levels of AI-2 in the mouse using an *E. coli* engineered bacterium led to changes in the ratio between the two major bacterial phyla in the gut, promoting the expansion of Firmicutes over Bacteroidetes after antibiotic-induced dysbiosis. Thus, as manipulation of the AI-2 signal causes changes in the microbiota composition, the ability to quantify AI-2 in this complex niche would allow a more comprehensive understanding of the effects of different AI-2 concentrations in the mammalian gut microbiota. However, the existence of multiple AI-2 isomers, their reactivity, instability and the high energy needed to ionize these molecules make their quantification challenging. Here, a method to purify AI-2 from mouse caecum was optimized and a previously developed HPLC-MS/MS method was adjusted to detect and quantify AI-2 in these extracts. We were able to quantify low micromolar concentrations of AI-2 in the caecum of mice with a manipulated microbiota, but further work to validate this method is still needed. Upon validation, this process will provide a direct quantification method that can be applied to complex biological samples where AI-2 mediated communication is suggested to be relevant.

5.2. Introduction

Quorum sensing is a signalling mechanism mediated by the production and detection of small chemical molecules named autoinducers (2,8,9). Bacteria use quorum sensing to sense their environment and regulate their gene expression to

control group behaviors such as luminescence, virulence and biofilm formation (3,50). Autoinducer-2 (AI-2) is an autoinducer produced and detected by both Gram-positive and -negative bacteria, which can mediate interspecies interactions (3,17,50). AI-2 was shown to have an important role in helping the mammalian gut microbiota maintaining equilibrium after dysbiosis (21). Thompson, Oliveira et al showed that upon antibiotic treatment, the equilibrium between the two major phyla of the gut, Bacteroidetes and Firmicutes, was affected. Namely, in response to streptomycin administration, members of the microbiota belonging to the Firmicutes were greatly decreased and Bacteroidetes expanded. However, the presence of an *Escherichia coli* engineered to accumulate AI-2, favored the colonization of the Firmicutes, counteracting the effect of the antibiotic on the ratio of Bacteroidetes and Firmicutes (21). Given that AI-2 seems to have a relevant role in the composition of the mammalian gut microbiota, a method to quantify AI-2 in this niche would allow an in-depth analysis of the influence of different concentrations of AI-2 in species composition of the gut and the study of the potential physiological implications of different concentrations of this signal molecule. The chemical instability of AI-2 and the fact that it actually comprises a family of interconverting molecules has hindered the development of methods to detect and quantify AI-2 in biological samples (48,186). AI-2 is produced as a linear molecule - (S)-4,5-dihydroxypentane-2,3-dione, DPD - by bacteria and can acquire different cyclic forms (46,182). In aqueous solution, these molecules exist in a chemical equilibrium being readily interconverted (46,182). So, there are multiple forms of AI-2 to detect. Moreover, AI-2 is stable at pH values equal or lower than 7 and is sensible to temperature, which means that some extraction procedures could lead to the degradation of the molecule (182).

Two assays are commonly used to detect and quantify AI-2. The most used method is the *Vibrio harveyi* luminescence bioassay, which is based on bacterial reporter strains and that measures the induction of light in transposon mutants of *V. harveyi* in response to AI-2 (14,185). This method is useful for fast detection of

AI-2 activity in biological samples but it is not very useful for quantification purposes. *V. harveyi* culture variability affects its reproducibility and other molecules present in the samples can inhibit the growth of *V. harveyi* preventing reliable quantification of AI-2 (190). Moreover, assay conditions like pH, media preparation and growth conditions can influence the measurements (190). The other assay available to detect AI-2 is an *in vitro* assay based on LuxP, the AI-2 receptor protein from *V. harveyi* and employs Förster resonance energy transfer (FRET) (171). For this assay, LuxP is bound to two fluorescent proteins, one in each domain. Binding of AI-2 triggers a conformational change in the receptor that alters the distance between the two fused fluorescent proteins, causing an alteration in FRET (171). This assay is more useful than the *V. harveyi* bioassay for quantification purposes, allowing quantification of AI-2 in the presence of an AI-2 calibration curve. However, the sensitivity of this assay only allows detection of AI-2 in concentrations from approximately 1 μM to 60 μM (185). Moreover, it also depends on the integrity of the protein. So, if the samples have components that cause changes in protein folding or that are able to degrade or denature the protein, the assay is rendered useless. Also, expression and purification of the labeled protein is time consuming (191). Additionally, both assays report only the fraction of DPD that has converted to the cyclic-borated form of AI-2 that binds the LuxP receptor (191,192). To overcome these disadvantages, three methods for quantifying AI-2, employing mass spectrometry, were developed (191-193). The use of mass spectrometry coupled with a separation technique allows the direct quantification of AI-2 in complex biological samples. The first method uses an *o*-diaminobenzene as a DPD-derivatizing agent to overcome the high energy needed to ionize AI-2, followed by separation by gas chromatography and detection by mass spectrometry (191). The other two methods also rely on DPD derivatization and use either *o*-diaminobenzene derived reagents (192) or the commercially available 4,5-dimethylbenzene-1,2-diamine (193) to derivatize DPD. In these two methods, to detect the derivatized DPD, a high performance liquid

chromatography-tandem mass spectrometry (HPLC-MS/MS) protocol was performed. These methods were used to detect and quantify AI-2 production in extracellular medium of AI-2 producing bacterial cultures. Importantly, only Campagna et al tested the developed protocol in a complex biological sample, quantifying AI-2 in human saliva (192). As this was the first report of successful quantification of AI-2 in a biological sample of mammal origin we used the method developed by Campagna et al as basis to detect and quantify AI-2 in mice caecum. To this purpose, we optimized the preparation of extracts from mice caecum samples to quantify AI-2 in the intestines using ultrahigh pressure liquid chromatography (UHPLC)-MS/MS. We were able to quantify AI-2 in the mammalian gut. Further samples have to be tested to determine if the variability observed is due to biological differences or inherent to the method. Nevertheless, upon optimization, this method should allow to precisely determine the effect of AI-2 concentration on species composition and physiological responses to potential changes of the gut microbiota.

5.3. Results and discussion

5.3.1. Development of a protocol for AI-2 purification from caecum content

Mice caecum is a complex sample with many components besides small molecules such as AI-2. So, an extensive purification procedure was developed to facilitate detection of this quorum sensing signal. To minimize the number of animals sacrificed, the tests for the extraction procedure were performed in caecum of specific pathogen free (SPF) mice utilized for other experiments. SPF mice contain a complex microbiota whose members produce various signals and metabolites. Additionally, it also contains signaling molecules and metabolites produced by the host and dietary compounds. A scheme of the protocols tested is shown in Fig. 5.1, A.

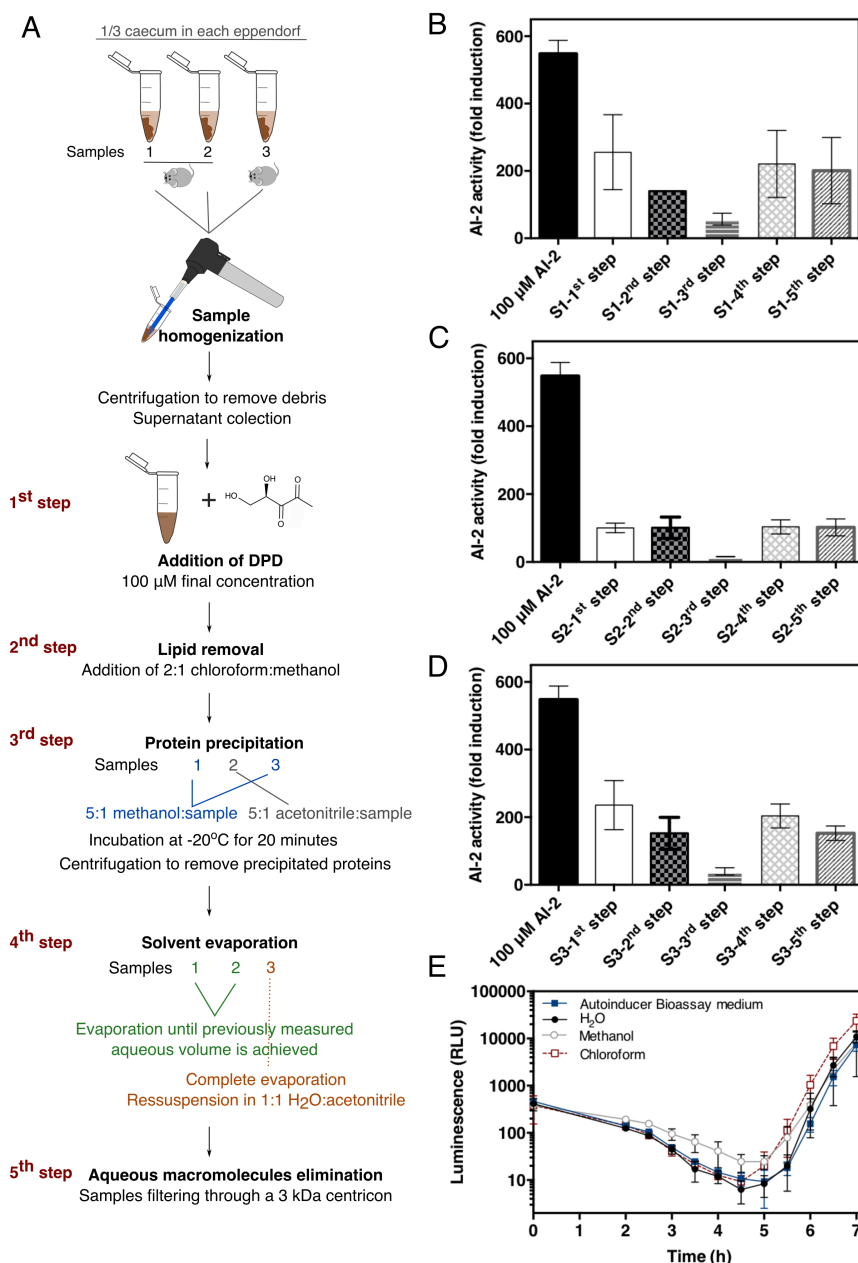


Figure 5.1 - Detection of AI-2 in the different steps of purification. (A) Scheme of the protocol followed to eliminate high molecular weight components from caecum content. (B) Sample 1 where methanol was used to precipitate proteins and the sample was not completely dried. (C) Sample 2 that was obtained from the same caecum as mouse 1 and for which acetonitrile was used to precipitate proteins. (D) Sample 3, obtained from the caecum of a different mouse than the one used to prepare samples 1 and 2; this sample was completely dried and resuspended in a 1:1 mixture of H₂O:Acetonitrile. Fold induction was calculated by dividing the light produced by the sample for the light produced by the appropriate control (methanol for step 3 and water for every other step). The

fold induction presented was obtained at 4h30 of incubation with *V. harveyi*, the lowest point of cognate AI-2 production. (E) Test of solvents used in the purification process to assess if they would cause *V. harveyi* growth inhibition.

We started by diluting approximately one third of the caecum of SPF mice, for each sample, in MiliQ water using a pellet pestle. Samples 1 and 2 were obtained from the caecum of the same mouse and sample 3 from the caecum of another mouse. The resuspended caeca were then subjected to a centrifugation to remove the debris and upon collection of the supernatants their precise volume was measured. For these tests, the necessary volume from a stock solution of chemically synthesized AI-2 was added to obtain a final concentration of 100 μ M (Fig. 5.1, B-D; 1st step tested). The addition of synthetic AI-2 allowed us to be sure that we would have detectable levels of AI-2 in these samples and to assess if any of the steps of the protocol would cause the degradation of the molecule. To remove lipids, the same volume of a 2:1 mixture of chloroform:methanol was added to the samples (Fig. 5.1, B-D; 2nd step tested) (194,195). To precipitate proteins we tested both methanol and acetonitrile. Thus, to samples 1 and 3 (that were obtained from different mice) 5 times the volume of the sample of ice-cold methanol was added and the mixtures were incubated at -20 °C for 20 min (196-198). Likewise, 5 volumes of acetonitrile were added to sample 2 and the mixture was also incubated at -20 °C for 20 min (196,198). After a centrifugation to remove the precipitated proteins, the supernatants (Fig. 5.1, B-D; 3rd step tested) were added to a SpeedVac equipment to evaporate methanol or acetonitrile. A program without heating was selected to prevent degradation of AI-2. Whereas samples 1 and 2 were subjected to evaporation until they reached approximately the same volume as they had before methanol or acetonitrile addition, sample 3 was completely evaporated to test if a higher degree of evaporation would degrade AI-2. Sample 3 was then resuspended in a 1:1 mixture of H₂O:Acetonitrile (Fig. 5.1, B-D; 4th step tested). Following this step, all the samples were submitted to a filtration through a centricon with a 3 kDa membrane (Fig. 5.1, B-D; 5th step tested) to remove aqueous macromolecules. To assess the presence of AI-2 in all of the

steps of the isolation process, AI-2 activity was measured using the *V. harveyi* BB170 bioassay that produces light in response to AI-2 (Fig. 5.1, B-E). This reporter strain produces AI-2 and, thus, fold induction of light produced by the added samples is analyzed at the point of lower light production induced by *V. harveyi* AI-2 to ensure the lowest interference of the cognate AI-2. An advantage of using this strain is that it allows to assess if samples components are inhibiting *V. harveyi* and, therefore, interfering with the assay. So, besides testing the samples we also tested if the solvents utilized in the purification process would cause inhibition of the *V. harveyi* luminescence (Fig. 5.1, E). The lowest point of light induction for this culture was at 4h30 of incubation when considering water, the solvent where the samples were dissolved, as our control (Fig. 5.1, E). By looking at the differences in luminescence induced by the various samples, it is possible to observe that AI-2 was detectable in every step of the purification, indicating that the process did not degrade AI-2 (Fig. 5.1, B-D). Sample 2 induced lower luminescence values than the other samples (compare Fig. 5.1 C with Fig. 5.1 B and D). This difference might be due to technical variance or to a non-homogeneous separation of the caecum components in the portion of caecum collected. Nevertheless, luminescence differences between the different steps tested was very similar to the one observed for the remaining samples making the conclusions taken from these tests common for every sample. In step 3, there was a lower fold induction that can be easily justified by the fact that the samples were diluted due to the addition of 5 volumes of methanol or acetonitrile. Although no clear difference was observed between the samples 1 and 3, we decided to proceed with protein purification using methanol instead of acetonitrile because during the procedure, it seemed to be more efficient in causing visible protein precipitation. We performed further tests and decided not to completely dry the samples to avoid AI-2 degradation.

The luminescence induced by the control, 100 μ M of AI-2 in water, was higher than the one obtained for the different purification steps including the first step where AI-2 was added to the soluble portion of the caecum. This observation

indicates that something in the caecum content inhibits the ability of *V. harveyi* to produce luminescence. Likewise, although methanol and chloroform did not kill *V. harveyi*, the addition of methanol changed the quorum sensing curve of this bacterium. This reinforces the utility of this method to identify the presence of AI-2 but the difficulty in using it for precise AI-2 quantification.

5.3.2. Optimization of the removal of high molecular weight components from caecum extracts

Although lipids and proteins were removed with the previous methodology, filtration through a 3 kDa membrane was still very lengthy and caused filter clogging, reflecting the presence of other interfering high molecular weight components. In an attempt to minimize filter clogging and the time necessary to filtrate samples, an extra filtration step through a 10 kDa membrane was added before the 3 kDa filtration. We prepared a third of a SPF caecum with addition of 50 μ M AI-2 (sample 4) and a third of a caecum of a different mouse without addition of AI-2 (sample 5; schematic representation on Fig. 5.1, A). By using these samples we were able to assess the effect of adding an extra filtration step as well as assess the levels of AI-2 in non-manipulated SPF caecum. Again, the *V. harveyi* bioassay was employed to test if we could still detect AI-2 after both filtration procedures (Fig. 5.2, B). It was possible to observe that, in the SPF caecum supplemented with AI-2 in the beginning of the extraction, AI-2 was detected in the three final steps (Fig. 5.2, B). Actually, there seemed to be a slight increase in AI-2 activity with the increase in sample processing. This might be due to interference of other molecules in the aqueous environment, which as mentioned above might either inhibit *V. harveyi*'s response or interact with AI-2. Thus, removal of these molecules through filtration seemed to enhance the detection of AI-2 via the *Vibrio* bioassay.

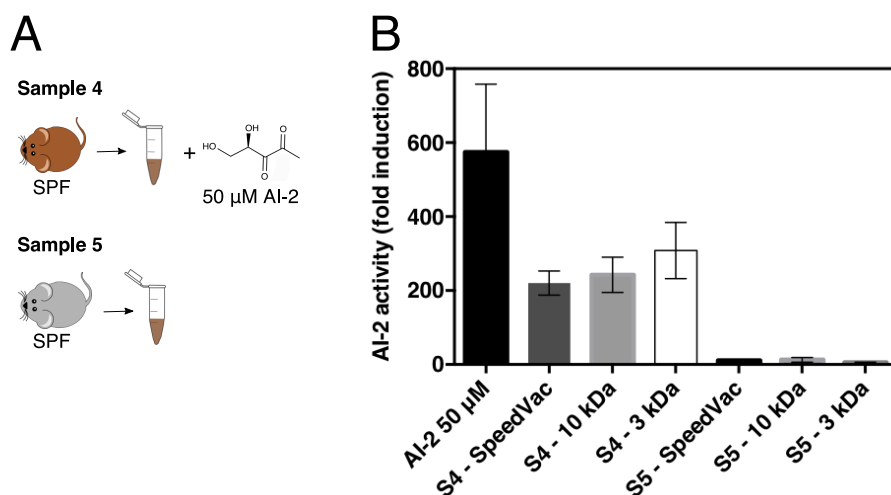


Figure 5.2 – Levels of AI-2 in specific pathogen free (SPF) caecum with or without the addition of AI-2. (A) Schematic representation of the samples prepared. (B) Assessment of AI-2 activity in SPF caecum with (S4 – sample 4) or without (S5 - Sample 5) supplementation of 50 μ M of AI-2 at the beginning of the extraction process. During the preparation of these samples, an extra filtration step through a 10 kDa filter was added before the filtration through a 3 kDa filter. Fold induction was calculated by dividing the light produced by the sample for the light produced by the appropriate control. The fold induction presented was calculated with luminescence obtained at 4h30 of incubation with *V. harveyi*, the lowest point of cognate AI-2 production. SpeedVac: samples after evaporation of methanol in a SpeedVac; 10 kDa: samples after filtration through a 10 kDa filter; 3 kDa: samples after filtration through a 3 kDa filter.

Moreover, the addition of the filtration through a 10 kDa filter reduced the time needed for the filtration through the 3 kDa filter and the amount of sample lost due to filter clogging. Therefore, the extra filtration step through a 10 kDa filter was added to the protocol. In SPF caecum without any AI-2 addition, no AI-2 was detected. A 1/20 dilution were performed to the samples to prevent saturation due to the addition of AI-2. To be able to compare the results directly, the same dilution was also performed to the samples without addition of AI-2. This dilution could be responsible for the inability to detect AI-2. However, a 1/2 dilution was also tested and we observed no signal in the SPF samples where AI-2 was not added (data not shown). This might mean that the concentration of AI-2 is low, being below the limit of detection or that AI-2 is absent in the caecum of these animals. The inability to detect AI-2 might be an indicative of the complex

microbiota that inhabits the gut of SPF mice, which might be able to produce and consume AI-2, avoiding the accumulation of this molecule.

At this point, we wondered if we could detect AI-2 in mice with a defined microbiota. Thus, we prepared samples from germ free (GF) mice inoculated with an *E. coli* strain that had been engineered to accumulate AI-2 in the extracellular medium (Fig. 5.3).

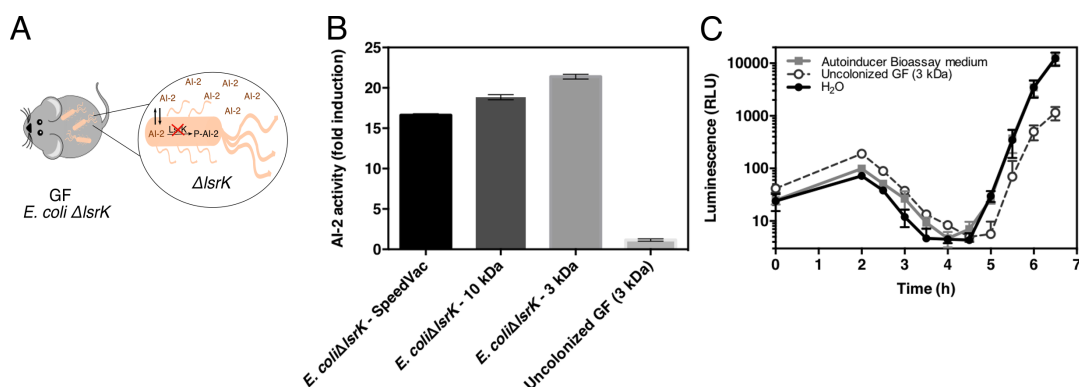


Figure 5.3 - Levels of AI-2 in monocolonized germ-free (GF) mice. (A) Schematic representation of the molecular mechanism that drives AI-2 accumulation in the extracellular medium upon colonization of GF mice with *E. coli* Δ *lsrK*. P-AI-2: phosphorylated AI-2. (B) Assessment of AI-2 activity in caecum of GF mice uncolonized or monocolonized with *E. coli* Δ *lsrK*. Fold induction was calculated by dividing the light produced by the sample for the light produced by the appropriate control. The fold induction presented was obtained at 4h30 of incubation with *V. harveyi*, the lowest point of cognate AI-2 production. SpeedVac: samples after evaporation of methanol in a SpeedVac; 10 kDa: samples after filtration through a 10 kDa filter; 3 kDa: samples after filtration through a 3 kDa. (C) Quorum sensing curves of *V. harveyi* BB170 obtained in the presence of the controls.

The AI-2 producing strain (*E. coli* Δ *lsrK*) has a mutation in LsrK, the kinase responsible for phosphorylating intracellular AI-2 and, consequently, this mutant is unable to trap AI-2 inside the cell and will lead to the extracellular accumulation of AI-2 (52,62). In GF mice monocolonized with this *E. coli* we could detect AI-2, indicating that in a caecum where the microbiota accumulates AI-2, we could detect AI-2 even after the extensive purification procedures (Fig. 5.3, B). However, the levels of AI-2 induction were much lower than the ones obtained with SPF caecum supplemented with 50 μ M of AI-2 (Fig. 5.2, B), indicating that the concentration of AI-2 might be significantly lower in this condition. Finally,

similarly to what was observed for caecum of SPF mice, an increase in AI-2 activity was also observed with the increase of sample processing (Fig. 5.3, B). This reinforces the suggestion that compounds from caecum extracts inhibit detection of AI-2 through the *Vibrio* bioassay and that the protocol used is capable of removing some, although not all, inhibitory compounds.

All in all, we can conclude that the purification steps employed did not lead to the degradation or loss of AI-2. Still, the loss of small amounts of AI-2 in these steps can not be excluded.

5.3.3. AI-2 derivatization and ethyl acetate extraction for mass spectrometry analysis

After having established a protocol to remove lipids, proteins and molecules with high molecular weight from the caecum extracts, we needed to derivatize AI-2 to detect it by mass spectrometry. Smaller molecules like AI-2 linear precursor, DPD, usually have higher ionization energies. This higher energy requirement makes their ionization a difficult process and therefore renders their analysis by mass spectrometry inapplicable (192). To overcome this hindrance, DPD was derivatized with a tag (dimethyl 2,2'-((4,5-diamino-1,2-phenylene)bis(oxy))diacetatdimethyl) originating a quinoxaline derivate (dimethyl 2,2'-((2-(1,2-dihydroxyethyl)-3-methylquinoxaline-6,7-diyl)bis(oxy))(R)-diacetate; DPD-Q; Fig. 5.4), as described by Campagna et al (192).

The formation of DPD-Q makes detection of AI-2 by mass spectrometry possible and enhances the stability of this molecule (191,192,199). Moreover, DPD-Q is more apolar than DPD allowing a subsequent extraction with ethyl acetate to further isolate AI-2 for quantification. Thus, after reaction with the tag, an extraction with ethyl acetate was performed and the organic phase that contains DPD-Q was evaporated and resuspended in acetonitrile. This procedure was only possible due to the higher stability of DPD-Q when compared with the one of DPD alone. For the initial samples, we used a large excess of tag to derivatize AI-2. As a standard, we prepared 2 mM of DPD in water and derivatized it with 2.8 mM of tag. To test the ability to detect DPD in GF caecum extracts, 2 mM or 50 μ M of DPD was added to these extracts before the ethyl acetate extraction and derivatized with 2.8 mM or 280 μ M of tag, respectively. As negative controls, we prepared GF caecum without addition of DPD or tag and GF caecum derivatized with 2.8 mM or 280 μ M of tag. Next, all these samples were subjected to HPLC-MS/MS. In the samples of DPD added to GF caecum extracts, a peak with a retention time of approximately 5.3 min (Fig. 5.5, B) similar to the one observed for the DPD derivatized in water (Fig. 5.5, A) was observed. However, derivatized GF caecum without added DPD, also had peaks with a retention time very similar to the one obtained for DPD-Q (compare Fig. 5.5, C with A and B). As these GF mice were not colonized with bacteria, samples from these mice should not have AI-2, as no microbes able to produce DPD were present. Thus, we reason that a

contaminant with a similar chemical structure to DPD exists in the caecum of GF mice. From both DPD-Q in water (Fig. 5.5, A) and in the biological matrix (GF extract; Fig. 5.5, B), we expected a retention time for DPD-Q between 5.2 – 5.4 min. However, in GF extract where only tag was added, we could detect two peaks in that area with retention times of 4 and 5.29 (Fig. 5.5, C). Therefore, the presence of these peaks can interfere with our measurements/quantifications. Still, analysis of these samples indicated that concentration of the contaminant was low and we reasoned that if we would avoid using tag in excess, the contribution of the contaminant could be negligible. Thus, we adjusted the concentration of tag used. From then on, 3 molar equivalents of tag were added for controls with known DPD concentration. For samples with unknown DPD concentrations, a concentration of 5 μ M was assumed and three additions of 5 μ M of tag were performed. After this concentration adjustment, we observed that the mass chromatogram obtained for GF caecum derivatized with 15 μ M tag (Fig. 5.5, D) was identical to the one obtained for GF caecum without addition of tag (Fig. 5.5, E), showing that at this lower concentration of tag, the compounds with peaks with a similar retention time to DPD-Q were no longer detectable.

Importantly, the observation that the peaks obtained when the contaminant reacts with the tag have a similar retention time to the one obtained for DPD-Q, suggests that this contaminant molecule is likely to have a structure similar to DPD. Recently, Ismail et al reported an AI-2 mimic produced by epithelial cells of mice in response to interaction with bacteria (117). We considered the hypothesis that this host-derived AI-2 mimic was our contaminant. However, the detection of the contaminant in GF caecum where no bacteria is present, might indicate that either this mimic does not need bacterial induction for its production, or, that another molecule with a similar chemical structure to DPD is present in the caecum of GF mice. Upon discovery of the chemical structure of the AI-2 mimic, a mass profile of the derivatized mimic can be determined and compared with the

Chapter 5

one obtained for the contaminant and, consequently, determine if the contaminant and the mimic are the same molecule.

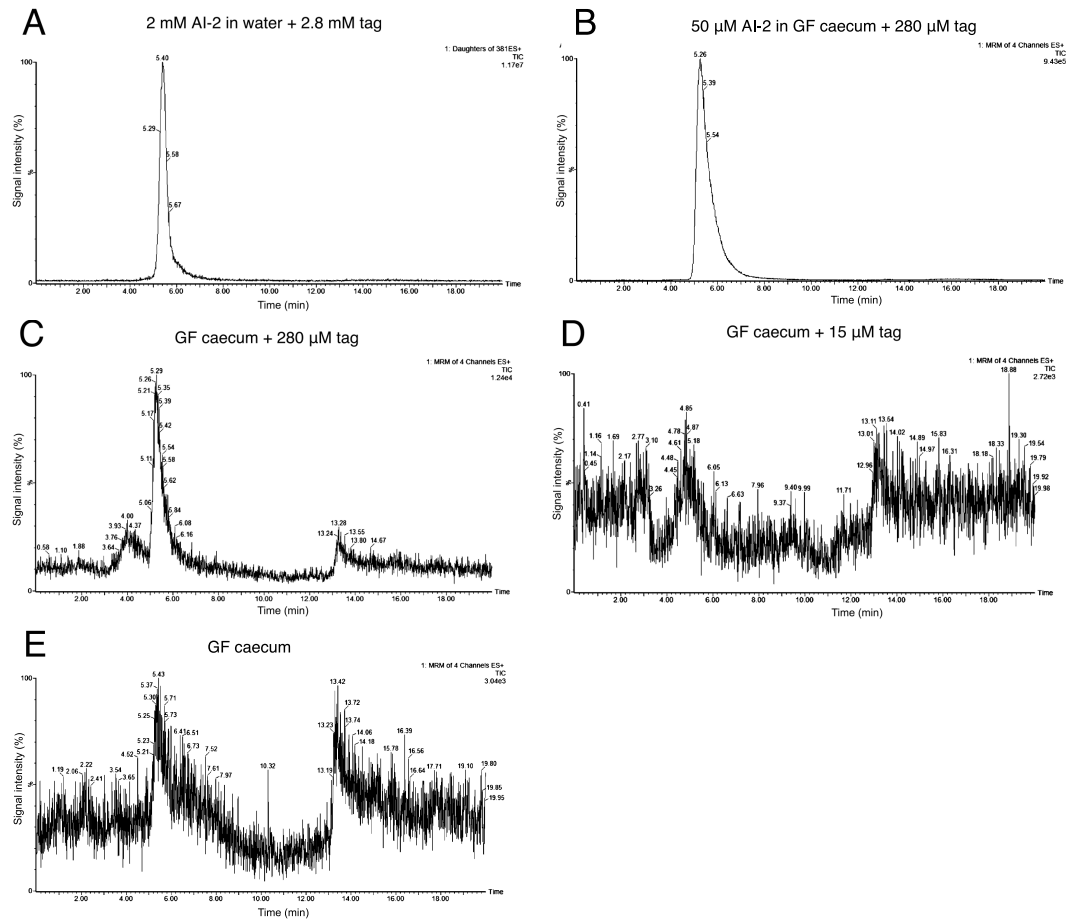


Figure 5.5 – Mass chromatograms of AI-2 derivatized with large excess or adjusted concentrations of tag. (A) Standard sample prepared with 2 mM of DPD in water derivatized with 2.8 mM of tag. (B) Positive control sample with 50 μ M of DPD added to the extract prepared from a germ free (GF) mouse before derivatization with 280 μ M of tag. (C) Control sample with addition of 280 μ M of tag to the extract prepared from the caecum of a GF mouse. (D) Control sample with addition of an adjusted concentration of tag; addition of 15 μ M of tag to the extract prepared from the caecum of a GF mouse. (E) Extract from GF caecum without addition of DPD or tag. An ethyl acetate extraction was performed before analyzing the samples by HPLC-MS/MS.

5.3.4. Choosing the collision energy

To detect DPD-Q in the caecum extracts prepared, a triple quadrupole mass analyzer that allows the use of multiple reaction monitoring (MRM) was employed(200). In this case, the first quadrupole (Q1) will select the molecular cation of interest (m/z 381 for DPD-Q) that will collide with the neutral collision gas (argon) in Q2 and undergo further fragmentation. The fragment ions will be accelerated and enter Q3 where they will be monitored. By employing MRM, pre-determined molecular and fragment ions are chosen, increasing the specificity and sensitivity of detection, indispensable characteristics for compound quantification (200). From a previous work, it was known that the fragment ion m/z 202 was the one with higher specificity for DPD-Q (192). However, since it suffers a higher number of fragmentations, quantification could be less sensitive. Thus, we decided to monitor multiple transitions (381>363, 381>261, 381>231 and 381>202) at the same time. Transition 381>202 would be used for identification purposes and one of the others would be used for quantification. To optimize detection, the more efficient collision energy needed to be selected. To achieve this goal, we analyzed the chromatograms of the 4 chosen transitions in three samples with different concentrations of DPD in the biological matrix, using two different collision energies (Table 5.1).

It was possible to observe that by using a collision energy of 21 eV the signal-to-noise ratio (S/N) was lower for the transition 381>202 than when using a collision energy of 28 eV. Moreover, when using a collision energy of 21 eV, we were unable to identify the DPD-Q peak for the sample with lower concentration of AI-2 (GF + 0.5 μ M; Table 5.1) using the transition 381>202. In contrast, when using a collision energy of 28 eV, we were able to identify this peak in all the samples. As detection of DPD-Q in the 381>202 transition is important for specific identification of the compound we decided to use a collision energy of 28 eV. Choosing this collision energy did not compromise the quantification of the

sample as transitions 381>231 and 381>261 had a S/N ratio higher than 10 for all the three samples, allowing its quantification (Table 5.1).

Table 5.1 – Signal-to-noise (S/N) ratios obtained for the four transitions analyzed when employing 21 or 28 eV as collision energy. Caecum of GF mice supplemented with 50 μ M of AI-2 before tag derivatization was used. Derivatization was performed with approximately 150 μ M of tag. Dilutions of this sample in acetonitrile were performed to obtain the samples with 5 and 0.5 μ M of AI-2. ^a The noise was clipped during chromatogram acquisition thus, it would not be correct to calculate a S/N. ^b The noise was too intense to correctly identify the peak.

Collision energy	Sample	Transition	Retention Time	S/N
21 eV	GF + 50 μ M	381>202	4,94	> 10 ^a
		381>231	4,84	87,39
		381>261	4,88	43,9
		381>363	4,82	159,32
	GF + 5 μ M	381>202	4,94	2,36
		381>231	4,94	24,02
		381>261	4,98	18,97
		381>363	4,92	45,07
	GF + 0.5 μ M	381>202	- ^b	- ^b
		381>231	5,2	8,5
		381>261	5,33	6,24
		381>363	5,26	10,01
	GF + 50 μ M	381>202	5,19	63,42
		381>231	5,17	136,46
		381>261	5,18	118,74
		381>363	5,16	94,99
28 eV	GF + 5 μ M	381>202	4,96	13,45
		381>231	5,02	74,19
		381>261	4,94	37,52
		381>363	4,95	14,53
	GF + 0.5 μ M	381>202	4,9	2,56
		381>231	4,91	13,29
		381>261	5,01	10,51
		381>363	- ^b	- ^b

5.3.5. Specificity of AI-2 detection in the biological matrix

As mentioned in section 5.3.1, we prepared the samples by using approximately a third of mice caecum for each sample. To ensure that we were measuring a homogeneously distributed AI-2 concentration in the mice caecum, we started to prepare samples from whole caeca.

For the following experiment, we prepared three caeca of GF mice colonized with *E. coli* Δ *lsrK* and one of a SPF mouse. As negative controls, we prepared one caecum of a GF mouse colonized with *E. coli* Δ *luxS* Δ *lsrK*, which is unable to produce or accumulate AI-2 as it has the AI-2 synthase and the AI-2 kinase mutated, and a caecum of an uncolonized GF mouse. To quantify the amount of AI-2 in the samples, we needed a sample that would take into account the concentration of AI-2 that is lost during the preparation of the caecum extracts. Thus, we prepared caecum from an uncolonized GF mouse and added 50 μ M of AI-2 at the beginning of the purification process (sample GF + 50 μ M AI-2). This sample was then subjected to the exact same treatment. For all the samples prepared, we maintained the same proportion of caecum matter/water and, after filtration through a 3 kDa filter, we made sure that the suspension obtained was homogeneous. Next, we used a 1 mL aliquot of each sample for AI-2 derivatization and ethyl acetate extraction. We took two aliquots of the extract prepared from the caecum of a GF mouse and, to one of them, added tag in the same quantity as for the biological samples with unknown AI-2 concentration, 15 μ M. Before analysis by HPLC-MS/MS, dilutions in acetonitrile of the sample GF + 50 μ M AI-2 were performed to obtain a calibration curve in the biological matrix that could be used to calculate the AI-2 concentration in the samples of unknown concentrations. However, after analyzing the samples through HPLC-MS/MS, we verified that, although we could detect AI-2 in the samples used to construct the calibration curve, we could not detect AI-2 in samples derived from the caecum of any of the GF mice colonized with an *E. coli* Δ *lsrK* (data not shown).

Hypothesizing that AI-2 would be below detection in these caeca, we tried to concentrate the samples using two different approaches (Fig. 5.6). In the first approach (process A, Fig. 5.6), the remaining volume of the samples that had been processed only until derivatization, were derivatized with tag and subjected to the ethyl extraction and subsequent evaporation. At this point, we tried to resuspend the whole samples in 200 μL but the formation of a precipitate was observed. We increased the volume to 500 μL and collected the samples. Analysis of the samples through HPLC-MS/MS confirmed that AI-2 had indeed precipitated, as we could not detect AI-2 in any sample (data not shown).

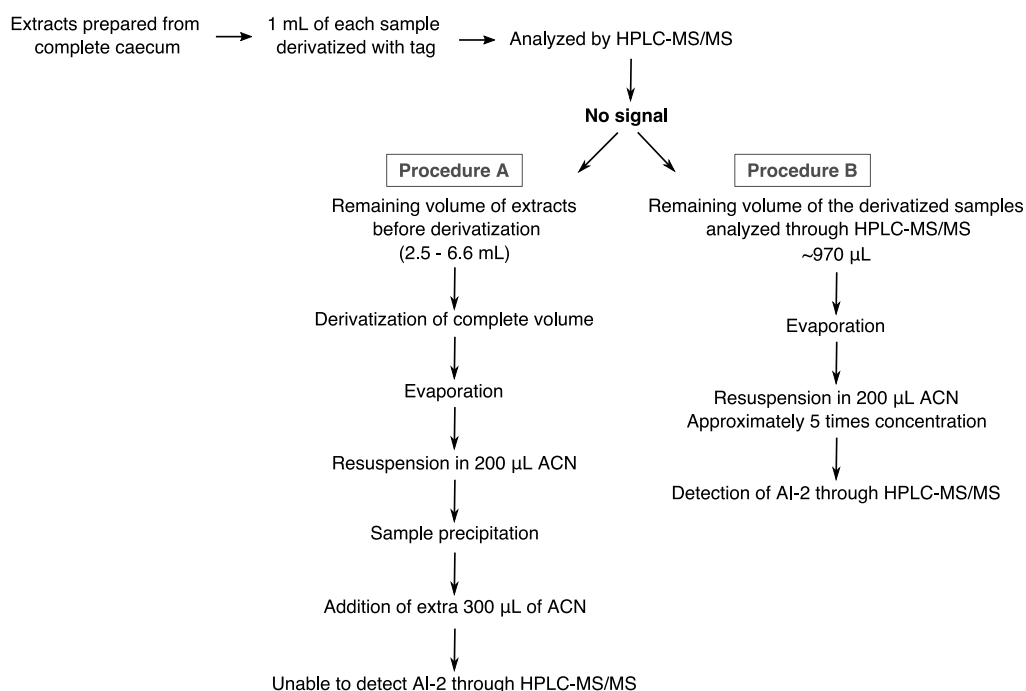


Figure 5.6 – Fluxogram of the two different sample concentration processes employed.

In the second approach (process B, Fig. 5.6), we recovered the remaining 970 μL (approximately) of the samples tested and evaporated the acetonitrile in a rotary evaporator. We then resuspended them in 200 μL of acetonitrile, performing a concentration of approximately 5 times. By analyzing these samples through

HPLC-MS/MS, we could detect AI-2 in one of the caecum of GF mice monocolonized with *E. coli* Δ *lsrK* (mouse 1, Table 5.2) with a S/N ratio equal or higher than 10, condition necessary for compound quantification. AI-2 in the caecum of GF mice 2 and 3 monocolonized with *E. coli* Δ *lsrK* was detected, but could not be correctly quantified, as the S/N ratio obtained was below 10 (Table 5.2). Moreover, using the samples from the calibration curve, we observed that we could quantify AI-2 until a minimum concentration of 2.5 μ M. The sample with 1 μ M of DPD was not included in the calibration curve as the S/N ratio obtained was below 10 (Table 5.2, Fig. 5.7).

Table 5.2 - Areas and signal-to-noise (S/N) ratios obtained for the peaks of the mass chromatograms of the tested samples and of the samples used to construct the calibration curve. The values are presented as average of three technical replicates and associated standard deviation. ^a The S/N is below 10 and a correct quantification is not possible. ^b The values used are out of the calibration curve. The transition used for quantification was 381>231.

Sample	Area (a.u.)	S/N	[DPD] (μ M)	Sample [DPD] (μ M)
GF + 50 μ M DPD	13374.2 $\pm$ 2910.8	158.3 $\pm$ 50.2	-	-
GF + 20 μ M DPD	4051.6 $\pm$ 55.8	43.7 $\pm$ 7.2	-	-
GF + 10 μ M DPD	2203.3 $\pm$ 221.3	44.3 $\pm$ 9.2	-	-
GF + 8.5 μ M DPD	2034.9 $\pm$ 135.8	36.2 $\pm$ 4.7	-	-
GF + 7 μ M DPD	1548.9 $\pm$ 108.3	30.2 $\pm$ 2.5	-	-
GF + 5 μ M DPD	1348.2 $\pm$ 59.5	23.1 $\pm$ 3.3	-	-
GF + 2.5 μ M DPD	652.6 $\pm$ 46.7	15.2 $\pm$ 1.1	-	-
GF + 1 μ M DPD	220.9 $\pm$ 16.7	8.8 $\pm$ 0.7 ^a	-	-
GF mouse 1 + Δ <i>lsrK</i>	622.51 $\pm$ 184.88	12.7 $\pm$ 3.3	3.6	0.7
GF mouse 2 + Δ <i>lsrK</i>	431.93 $\pm$ 188.9 ^b	7.5 $\pm$ 0.8 ^a	~2.8	~0.6
GF mouse 3 + Δ <i>lsrK</i>	472.03 $\pm$ 27.56 ^b	7.1 $\pm$ 0.6 ^a	~3.0	~0.6

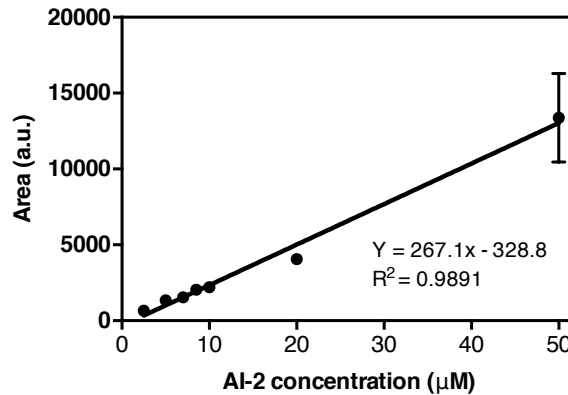


Figure 5.7 – Calibration curve obtained from the data on table 5.2.

Nevertheless, it is important to keep in mind that we assume 2.5 μM considering that no AI-2 was lost during the purification and extraction processes. As this is highly improbable, our lower limit of detection should actually be smaller. For the sample GF mouse 1 + ΔlsrK we quantified 3.6 μM of AI-2. Taking into account the volume used to prepare this sample, it would mean that our sample had approximately 0.7 μM . This quantification showed us that the level of AI-2 in these samples and, consequently, in the caecum, is very close to our limit of detection. It also could provide a justification to why we could not quantify AI-2 in the other two samples of GF + ΔlsrK as a slightly lower AI-2 concentration, inherent to biological variability, will be out of the calibration curve and, therefore, below our detection limit. Since the attempt to concentrate the extracts more than 5 times resulted in precipitation and loss of AI-2, a method/equipment with higher sensitivity was needed.

These alterations to the protocol were also performed for the sample prepared from a SPF mouse, but AI-2 was never detected. This observation is in line with what was observed in the *V. harveyi* bioassay (Fig. 5.2, B) where AI-2 activity was not detected in SPF extracts. This inability to detect AI-2 might be the consequence of a complex microbiota where AI-2 is constantly being produced

and internalized, not allowing its accumulation to levels able to be detected and quantified, as mentioned before.

5.3.6. Improvement of the liquid chromatography and mass spectrometry methods

Given our need for higher sensitivity, our collaborator Joana from the Laboratory of Structural Analysis of the Faculty of Pharmacy from the Universidade de Lisboa, tested a different equipment. The new equipment used was also a triple quadrupole but it is associated with an ultra-pressure liquid chromatography (UPLC) instead of a HPLC. Moreover, we had been using an electrospray ionization source (ESI) and she tested an atmospheric-pressure chemical ionization source (APCI) as well. However, DPD-Q ionization was better using ESI so there was no change in the ionization conditions. Also, our aminopropyl NH₂ column was giving problems due to extensive use with dirtier samples so a C18 and a Hilic column were tested. The separation of DPD-Q was better with a C18 column. A collision energy of 25 eV was used for all transitions but for the 381>363 transition, for which a collision energy of 20 eV was used. The collision energies follow the same rationale as described before (Table 5.1), as higher S/N ratio was obtained for the 381>363 transition with a lower collision energy.

With these improvements, Joana re-ran the 10 samples of the calibration curve with concentrations of AI-2 that range from 20 to 0.25 μ M. The sample with lowest concentration, 0.25 μ M, was still possible to quantify (Table 5.3 and Fig. 5.8). Again, it was possible to quantify AI-2 in GF mouse 1 + *E. coli* Δ lsrK. The value obtained was lower (2.2 μ M instead of 3.6 μ M that taking into account the concentration step, results in a caecum concentration of 0.4 μ M instead of 0.7 μ M) but it was in the same order of magnitude. The GF + *E. coli* Δ lsrK samples from mice 2 and 3, had a significantly lower area than before. However, as the S/N was better than before, it is probably more accurate. The area values are below the ones

used to construct the calibration curve and thus could not be correctly quantified. This observation suggests that there might be high biological variance in the levels of AI-2 in the caecum of different mice, although small experimental differences cannot be excluded. Further tests with other caeca from GF+ *E. coli* Δ lsrK mice will shed a light on biological variance and experimental reproducibility.

Table 5.3 - Areas and signal-to-noise (S/N) ratios obtained for the peaks of the mass chromatograms of the tested samples and of the samples used to construct the calibration curve. The values are presented as average of three technical replicates and associated standard deviation. ^a The values used are out of the calibration curve. The transition used for quantification was 381>363.

Sample	Area (a.u.)	S/N	[DPD] (μ M)	Sample [DPD] (μ M)
GF + 20 μ M DPD	4995.2 $\pm$ 104.7	1163.0 $\pm$ 191.0	-	-
GF + 10 μ M DPD	2485.3 $\pm$ 23.4	517.3 $\pm$ 39.1	-	-
GF + 8.5 μ M DPD	1973.7 $\pm$ 22.8	424.0 $\pm$ 50.0	-	-
GF + 7 μ M DPD	1764.9 $\pm$ 88.6	297.4 $\pm$ 80.8	-	-
GF + 5 μ M DPD	1294.1 $\pm$ 59.5	232.9 $\pm$ 14.2	-	-
GF + 2.5 μ M DPD	592.4 $\pm$ 21.4	128.0 $\pm$ 41.6	-	-
GF + 1 μ M DPD	201.6 $\pm$ 12.3	42.8 $\pm$ 6.8	-	-
GF + 0.5 μ M DPD	120.0 $\pm$ 18.0	24.7 $\pm$ 5.2	-	-
GF + 0.25 μ M DPD	127.7 $\pm$ 3.3	16.5 $\pm$ 2.7	-	-
GF mouse 1 + Δ lsrK	537.9 $\pm$ 62.7	91.7 $\pm$ 28.5	2.2	0.4
GF mouse 2 + Δ lsrK	108.5 $\pm$ 8.7 ^a	10.6 $\pm$ 3.4	~0.5	~0.1
GF mouse 3 + Δ lsrK	71.7 $\pm$ 13.8 ^a	11.1 $\pm$ 1.6	~0.3	~0.1

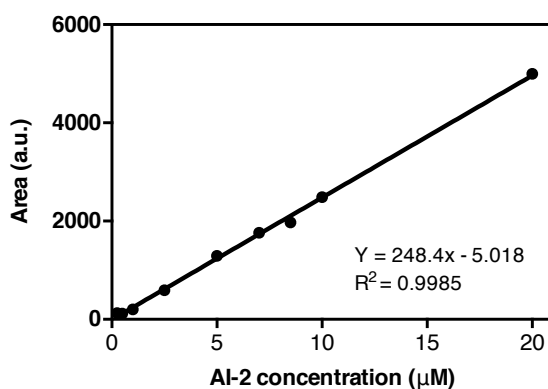


Figure 5.8 - Calibration curve obtained from the data on table 5.3.

5.4. Conclusion

In this work, we optimized the preparation of mice caecum samples for analysis through UHPLC-MS/MS and were able to quantify AI-2 in this biological matrix up to a concentration of 0.25 μ M. It was also possible to detect and quantify AI-2 in one of our samples from mice monocolonized with an AI-2-producing *E. coli*. To the best of our knowledge, this is the first description of the detection of AI-2 from mammalian caecum using a mass spectrometry approach. Although procedures for extraction of metabolites from caecal and faecal matter exist (201,202), the unique characteristics of AI-2 limit the conditions that can be used. Moreover, mass spectrometry-based quantification of AI-2 has only been applied to less complex biological samples such as cell-free bacterial supernatants and human saliva. Further samples have to be tested to determine if the lack of reproducibility in detecting of AI-2 from mice caecum is due to biological differences or inherent to the method. Still, upon optimization, this method will allow the study of the impact of different concentrations of AI-2 on species composition and physiological responses of the gut microbiota

5.5. Experimental procedures

5.5.1. Optimized preparation of caecum samples for UHPLC-MS/MS analysis

To prepare the samples for mass spectrometry-based analysis the caecum content to be treated was weighted and the amount of water needed to add (5 mL of water for 1 g of caecum) was calculated (195). Samples were kept on ice through the whole procedure. A portion of the calculated MilliQ H₂O volume was added and the caecum was suspended in solution using a pellet pestle (Sigma-aldrich). Once homogeneity was achieved, the remaining volume of MilliQ H₂O was added and the sample was vortexed. To remove debris the sample was then centrifuged

at 20 000 g for 15 min at 4°C. The supernatant was transferred to 15 mL falcon tubes and the precise volume was registered. For the initial tests and for the sample used to construct the calibration curve, the appropriate volume of DPD, for the desired concentration, was added in this step. To remove lipids, the same volume registered before of a 2:1 mixture of chloroform:methanol was added (194,195). The samples were vortexed for 1 min and immediately centrifuged for 5 min at 4000 rpm, 4 °C. The upper phase (the aqueous phase where AI-2 is) was collected to a new tube and the process was repeated two more times. To precipitate proteins, 5 volumes of sample (taking into account the sample volume precisely measured before) of ice-cold methanol were added to the samples and the mixture was vortexed for 30 seconds and incubated at -20 °C for 20 min (196-198). To remove the precipitated proteins, the samples were centrifuged at 4000 rpm for 30 min at 4 °C (196-198). The resulting supernatant was collected and divided into 2 mL eppendorfs and placed in a SpeedVac concentrator (Thermo Scientific Savant SPD131DDA) device to evaporate methanol. The equipment used had a coupled refrigerator system that prevented sample heating (and DPD degradation). When the samples reached approximately the same volume registered before, the concentration process was terminated. The samples were then sequentially passed through 10 kDa and 3 kDa membranes by centrifugation at 4000 rpm, at 4 °C in centricon tubes (Millipore), according to manufacturer recommendation. The samples were then stored at -80 °C until further processing. To derivatize DPD, 1 mL of each sample was taken and the initial pH was measured and adjusted to pH values below 3 with HCl. The same mass (in mol) of tag was added and the mixture was incubated at room temperature for 30 min. When the DPD concentration was unknown, a concentration of 5 µM was assumed. This addition/incubation was performed for two more times. Next, the pH of the samples was adjusted to values between 8 and 9 with NaOH and an ethyl acetate extraction was performed. The solvent was evaporated in a rotary evaporator and the resulting film was resuspended in 200 µL of HPLC-grade

acetonitrile. Until analysis through mass spectrometry the samples were stored at -80 °C. To prepare the samples used to construct the calibration curve, dilutions of the sample with 50 µM of DPD added to GF caecum in the beginning of the extraction process, were performed using HPLC-grade acetonitrile.

5.5.2. Dimethyl 2,2'-((4,5-diamino-1,2-phenylene)bis(oxy))diacetate synthesis

Dimethyl 2,2'-((4,5-diamino-1,2-phenylene)bis(oxy))diacetate, the tag used to derivatize AI-2 was synthesized as described in (192).

5.5.3. *Vibrio harveyi* BB170 bioluminescence assay

V. harveyi response was measured using BB170 reporter strain grown in autoinducer bioassay medium, as previously reported (51,157). For SPF derived samples and respective controls, a 20 fold dilution was performed to prevent saturation due to AI-2 addition. Light production was measured over time in a microplate luminometer (1420 Victor 3, Perkin Elmer). AI-2 activity in the samples was calculated from luminescence values at the time point where the light produced by *V. harveyi* in the presence of the correct negative control was the lowest. AI-2 activity was reported as the induction of light production compared with the background light obtained with the appropriate control. In Fig. 5.1 the control was water for every sample except the ones from 3rd step of extraction. For these samples, methanol was used as control. In Figs. 5.2 and 5.3 the control was always water to be able to compare the differences between the different samples when normalized by the same control. Nevertheless, the control samples without AI-2 are also represented in the figure. Standard deviation was calculated from three technical replicates. Propagation of uncertainty was employed to calculate the standard deviation after normalization.

5.5.4. HPLC-MS/MS method used during sample optimization

The liquid chromatography procedure was performed in an Alliance 2695 HPLC (Waters) with a Phenomenex Luna NH₂ reversed phase column (5 μ m pore size, 100 Å particle size, 250 x 2 mm). The mobile phases were HPLC-grade water (solvent A) and acetonitrile (solvent B) used to construct a 20 min gradient elution profile (t= 0, 15% solvent A, 85% solvent B; t=2 min, 15% solvent A, 85% solvent B; t=4 min, 95% solvent A, 5% solvent B; t=15, 95% solvent A, 5% solvent B; t=16, 15% solvent A, 85% solvent B; t=20, 15% solvent A, 85% solvent B). The flow rate was constant at 150 μ L/min, the column was at 35 °C and the injection volume was 10 μ L. This separation module is connected to a Micromass Quatromicro API Tandem Quadrupole Mass Spectrometer (Waters) equipped with an ESI source (Waters). The instrument was operated in the positive electrospray mode, ESI capillary voltage was optimized to 2.5 kV, cone voltage was set to 25 V and source and desolvation temperature were set to 120 °C and 350 °C, respectively. Nitrogen gas was generated from a nitrogen source (Gasman G4800) and used to supply the cone with a flow of 50 L/Hr and as desolvation gas with a flow of 750 L/Hr. For collision induced dissociation, argon was employed and the collision energy was optimized to 28 eV. For data acquisition and processing, MassLynx version 4.1 software was used.

5.5.5. Final UHPLC-MS/MS method used to quantify AI-2

The liquid chromatography procedure was performed in an ACQUITY UPLC (Waters) with a LiChrospher 100 RP-18 column (5 μ m pore size, 250 x 4 mm). The mobile phases were HPLC-grade formic acid 0.1% (solvent A) and acetonitrile (solvent B) used to construct a 20 min gradient elution profile (t= 0, 95% solvent A, 5% solvent B; t=5 min, 50% solvent A, 50% solvent B; t=5.1 min, 5% solvent A, 95% solvent B; t=10, 5% solvent A, 95% solvent B; t=10.1, 95% solvent A, 5% solvent B; t=20, 95% solvent A, 5% solvent B). The flow rate was constant at 300 μ L/min, the column was at 25 °C and the injection volume was 10 μ L. This

separation module is connected to an ACQUITY TQ Detector Tandem Quadrupole Mass Spectrometer (Waters) equipped with a ZSpray ESI source. The instrument was operated in the positive electrospray mode, ESI capillary voltage was optimized to 3 kV, cone voltage was set to 25 V and source and desolvation temperature were set to 120 °C and 350 °C, respectively. Nitrogen gas was generated from a nitrogen source (Gasman G4800) and used to supply the cone with a flow of 50 L/Hr and as desolvation gas with a flow of 750 L/Hr. For collision induced dissociation, argon was employed and the collision energy was optimized to 20 eV for transition 381>202 and to 25 eV for the remaining transitions. For data acquisition and processing, MassLynx version 4.1 software was used.

5.6. Acknowledgments

The work performed in the laboratory of Karina B. Xavier was supported by Portuguese National funds from Fundação para a Ciência e Tecnologia (FCT) PTDC/BIA-MIC/4188/14. Karina B. Xavier and Inês M. Torcato acknowledge FCT for individual grants IF/00831/2015 and PD/BD/105736/2014, being the latter within the scope of the PhD program Molecular Biosciences PD/00133/2012. Vanessa Miranda acknowledges FCT for individual PhD grant SFRH/BD/102159/2014.

Chapter



General discussion

The chemistry of autoinducer-2 (AI-2) is very complex and interesting. Synthesis of AI-2 as a linear precursor by all bacteria that express the AI-2 synthase, LuxS, allows it to be used in interspecies communication. By being highly reactive, this linear form will then acquire different forms that seem to be influenced by the chemical environment in which bacteria live. Accordingly, proteins with different binding sites bind the different AI-2 isomers. LuxP, which binds the borated cyclic form of AI-2 (S-THMF-borate) was, so far, only ever found in the *Vibrio* genus (47,50). These bacteria are predominantly aquatic where the concentration of boron is relatively high (47,50). In sea water the concentration of boron is around 400 μM (47), whereas humans typically possess 60 μM of boron in plasma, which is easily and almost completely absorbed and rapidly eliminated in urine being the half-life of boron estimated to be of approximately 20 h in the human organism (203). Thus, it seems reasonable that enteric bacteria like *Salmonella enterica* serovar Typhimurium and *Escherichia coli* bind the non-borated form of AI-2 (R-THMF). Accordingly, it was in these enteric bacteria, which belong to the Proteobacteria phylum, that LsrB receptors, which bind R-THMF, were first identified.

From previous work, it was already known that LsrB receptors were more widespread than initially thought, as events of lateral gene transfer were suggested to be responsible for the transition of LsrB receptors from Proteobacteria to Firmicutes (17). Confirmation that functional LsrB receptors existed in the *Bacillus* genus established the presence of this type of AI-2 receptors in Gram-positive bacteria, belonging to the Firmicutes phylum (17). The identification of functional AI-2 receptors in Gram-positive bacteria was important for the consolidation of AI-2 as an interspecies signal that fosters communication across bacteria from different phyla. In the present work, we show that the binding site of LsrB receptors is more plastic than what was previously thought, as naturally occurring mutations in the amino acid residues that bind AI-2 were identified in a functional LsrB-like receptor. Although this receptor has two

alterations (P225N and A227S) in the AI-2 binding amino acids, it still binds R-THMF. The discovery of this noncanonical receptor expanded the criteria that can be used to identify more LsrB receptors. Through bioinformatic analysis we identified 95 putative, noncanonical LsrB receptors, expanding significantly the number of bacteria with the potential to sense AI-2. These hits were identified mainly in the Firmicutes but also in the Spirochaetes and Actinobacteria phyla. Moreover, we identified noncanonical LsrB receptors with only one of the mutations (A277S) that are also proposed to be functional LsrB receptors as in our site directed mutagenesis study, this mutant was able to bind AI-2. Nevertheless, *in vitro* AI-2 binding assays with some of these putative LsrB receptors are needed to confirm our predictions. Tests with the proteins found in *Treponema primitia* and *Treponema azotonutricium* would be especially interesting as these putative LsrB receptors are next to a sensor histidine kinase, similar to what happens with LuxP. Upon *in vitro* confirmation that these proteins bind AI-2, the determination of the concentration of extracellular AI-2 throughout time could also be informative. Since LuxS homologues were identified for these strains, determining if AI-2 would accumulate or be internalized in these organisms could give an indication if these putative LsrBs are interacting with another ABC transporter or with the nearby histidine kinase. Moreover, methods to investigate protein-protein interactions such as co-immunoprecipitation or even isothermal titration calorimetry (ITC) could be employed to determine if this putative receptor is interacting with the neighboring sensor histidine kinase.

Further, as DNA sequencing techniques are constantly evolving and the number of bacterial genomes sequenced and deposited is increasing, a novel search might identify more noncanonical LsrB receptors. Similarly, by employing the previously established criteria, canonical LsrB receptors might also be identified in different bacterial species. To expand this search, sequential searches using other LsrB proteins can also be performed. For example, if the ability to bind AI-2 was confirmed for the LsrB of a bacterial species belonging to the

Actinobacteria phylum like *Orsonella urininfantis*, a new search using this protein as query could result in the identification of other LsrB-like receptors in this phylum.

Experimentally, a comprehensive site-directed mutagenesis study to identify all the mutations that would still allow binding of AI-2 could be performed. Then, when looking for other putative LsrB receptors all the functional alterations would be known. However, this would be expensive and a vast amount of work. It would be more effective to see which substitutions are recurrent and perform those substitutions. Moreover, it is important to keep in mind that the amino acid residues that surround the binding site are extremely important. Even if an amino acid can bind AI-2, if its side chain cannot be accommodated by the surrounding amino acids, an extensive change in the overall structure of the protein can occur and prevent proper binding. Following this line of thought, *in silico* comparison of the identity of the amino acid residues surrounding the binding sites of LsrB receptors that possess different amino acids involved in AI-2 binding could shed some light on the reason why some amino acid switches are allowed. During our study, we also identified homologs to the noncanonical LsrB, with a novel substitution in the amino acids that bind AI-2. Sequence alignments indicate that these homologs have a cysteine instead of a proline (canonical LsrB) or asparagine (noncanonical LsrB) on top of the A227S substitution that we had identified as AI-2 binding residues in noncanonical receptors. These homologs were identified in the bacterial strains *Dorea* sp. 27, *Clostridium* sp. Marseille P2538 and *Clostridium hylemonae*, which belong to the Firmicutes phylum. All of them shared 56% of amino acid sequence identity with the LsrB of *C. saccharobutylicum* (CsLsrB) higher than the approximately 40% that CsLsrB shared with the canonical LsrBs. Further tests with these proteins would be interesting to assess if this would be a viable amino acid substitution.

Another approach to test the functionality of putative LsrB receptors was developed in this work. An isothermal titration calorimetry (ITC) assay was optimized to measure the dissociation constant of AI-2 binding to LsrB receptors. In these assays, we observed that, in the absence of dialysis, the N-value obtained in these assays was between 0.2 and 0.4. This value is usually interpreted as a measure of the stoichiometry of the reaction. From the crystal structure of CsLsrB, we knew that there was only one binding site able to bind only one molecule of AI-2. Thus, we expected this value to be close to 1. However, the N-value also is influenced by the concentrations of the protein and ligand, as evidenced in the following equation:

$$N - value = number\ of\ binding\ sites \times \frac{AF_{cell}}{AF_{syringe}} \quad (Eq. 1)$$

where AF_{cell} represents the concentration of active protein in the cell and $AF_{syringe}$ represents the concentration of active ligand, AI-2 in our case, in the syringe (204). Thus, if the estimated concentrations of LsrB or AI-2 were not accurate, the N-value would not be exactly 1. Usually, for ITC assays, data obtained with a N-value with 20% of variance is considered to be valid. As our N-value had 60 to 80% of variance and the values were below 1, either we had a fraction of the protein unable to bind AI-2 or we had much more AI-2 than we thought. As we did not see protein precipitation and AI-2 had been recently synthesized in the absence of boron (and thus, shifting the equilibrium to the non-borated cyclic forms) and quantified by NMR, we hypothesized that something could be blocking the binding sites. In fact, a run with the LsrB from *Bacillus anthracis* that had been submitted to a size exclusion chromatography gave an N-value of 0.87, which is included in the 20% of variance usually accepted, supporting our hypothesis. Accordingly, by submitting CsLsrB to dialysis, values of N equal or above 0.8 were obtained. As AI-2 is a small 5-carbon molecule with a structure similar to pentose sugars and the AI-2 receptors are substrate-binding

proteins with a fold similar to sugar binding proteins, we hypothesize that it was a pentose sugar or another similar intracellular metabolite that was bound to CsLsrB. Nevertheless, the affinity of this substrate could not be as high as the one observed for AI-2. AI-2 stays bound through extensive purification and crystallization processes and the unknown ligand was dislocated by the presence of AI-2. Thus, since CsLsrB and possibly other LsrB receptors have the ability to bind other compounds, it is likely that the receptors for these other metabolites are also able to bind AI-2. However, the affinity of these proteins for AI-2 should be lower than to its cognate ligand. Therefore, the developed ITC assay has potential to enable the distinction between proteins that evolved to be AI-2 receptors and other proteins that have higher affinity for other ligands but are able to bind AI-2 with lower affinity, as different dissociation constants should be obtained. From our results for LsrB receptors and from Zhu et al (144) for LuxP, we know that AI-2 receptors have dissociation constants in the submicromolar range. This suggests that every protein that binds AI-2 with higher dissociation constants is likely not an AI-2 cognate receptor. This assay could be applied to other ligand binding proteins, such as AibA and AibB, which were shown to bind AI-2 and interact with a chemoreceptor in *Helicobacter pylori* (89), similarly to what has been observed for the LsrB receptor of *E. coli*. However, these proteins share no significant similarity with the *E. coli* LsrB, which might suggest that these proteins could be a novel AI-2 detection system. Application of the ITC assay to study the binding affinity of AI-2 in these proteins can confirm their primary function and help to precisely distinguish between AI-2 receptors and proteins able to bind AI-2. Furthermore, the confirmation that other proteins can bind AI-2 with lower affinity than the cognate receptors could explain why bacterial mutants on LsrB receptors are still able to internalize AI-2, albeit at a slower rate.

Sequence comparison coupled with binding confirmation through biochemical assays, such as the ones employed in this work, allow the discovery

of receptors related to the classes of receptors known. However, some bacteria where no known receptor has been identified have phenotypes regulated by AI-2 (50,57). Discovery of a novel class of receptors would increase the knowledge on AI-2 mediated quorum sensing and thus, on the approaches to modulate it. Nevertheless, the discovery of new receptors presents some challenges. The fold of AI-2 receptors is not unique, being shared by sugar binding proteins (46,47). Moreover, the amino acid sequences of the two classes of receptors known share low similarity, a very distinct binding site and they bind different AI-2 isomers (46,47). Thus, to identify new AI-2 receptors, we developed a chemical probe with the potential to pull-down AI-2 receptors. The chemistry that makes AI-2 an interesting autoinducer, also makes its study a difficult task. Besides being highly reactive, AI-2 is unstable at high concentrations and high pH values, limiting the experimental conditions that can be employed (182). Moreover, the hypothesis that other forms of AI-2 could be the active ligands of unknown receptors has to be considered. Thus, when designing a probe with the goal of identifying novel AI-2 receptors we had to take into account that the coupled DPD would need some flexibility. To ensure that DPD would be able to cyclize, D-desthiobiotin was bound to the carbon 1 of this molecule, as DPD analogs with substituents on this carbon were able to bind both LsrB and LuxP receptors (180). Furthermore, D-desthiobiotin is a small molecule and should be less prone to cause structural hindrances than other molecules such as rhodamine previously tested by others (161). Indeed, we observed that this probe was able to bind LuxP as well as canonical and noncanonical LsrB receptors. This is the first report of a functional AI-2 probe able to bind both types of receptors. Garner et al had previously developed an AI-2 probe based on the structure of rhodamine (161). This probe was shown to bind LsrB receptors from *Salmonella* serovar Typhimurium and from *Bacillus cereus* by imaging (161). Upon binding, this rhodamine-based probe behaved as an antagonist for LsrB receptors (161). The large size of rhodamine might explain the antagonist role of the probe, as complete binding of the receptor

to DPD would not be possible. Nonetheless, coupling of DPD to rhodamine has the advantage of probably circumvent the need of a transporter for AI-2, due to the ability of rhodamine to enter the cells through passive diffusion (205). However, in contrast to our probe, it has the drawback of not being able to bind LuxP receptors. That might be due to a structural constrain that would prevent the formation of S-THMF-borate, as this borated AI-2 adduct is larger than R-THMF due to the incorporation of boron. Although this probe could be useful for the identification of bacteria with LsrB receptors by microscopy, DPD-bound to desthiobiotin (DB-AI-2) can be employed to discover new AI-2 receptors, as DPD can cyclize into the different active forms that bind the receptors. DB-AI-2 also possess the potential to pull-down new receptors from cell extracts through immunoprecipitation. As a proof of concept, pull-down by immunoprecipitation was performed with partially purified LuxP receptor. Preliminary tests with cell extracts from *Clostridioides difficile* were also performed. We confirmed that *C. difficile* produces AI-2 and it does not seem to behave as LsrB-harboring bacteria, as no AI-2 internalization was observed, in the growth conditions employed. However, this bacterium has phenotypes regulated by AI-2 (57,158,159) and therefore it is likely to have an AI-2 receptor. Accordingly, in an experiment performed by Diogo Martins in Adriano O. Henriques' laboratory, we observed that supplementing growing cultures of *C. difficile* with AI-2 increased the production of toxin A, one of the main virulence factors of this bacterium (unpublished data). This observation is in agreement with another study where expression of toxins A and B was proposed to be regulated by AI-2 in *C. difficile* CCUG19126 (57). Moreover, biofilm formation has also been suggested to be regulated by AI-2 in *C. difficile* (158,159). Thus, upon identification of the new receptor, receptor mutants will be constructed to confirm that toxin A production and biofilm formation are regulated by AI-2. The use of receptor mutants prevents the alteration of metabolic functions, which might be affected when using a *luxS* mutant. Also, RNAseq can be employed to compare the transcriptome between

the mutant and the wild type. This would allow the identification of genes regulated by AI-2 in this organism and, consequently, the regulated traits. Furthermore, a characterization of the new receptor similar to the one performed for CsLsrB, will allow the identification of this type of receptor in other bacterial strains. Alternatively, DB-AI-2 could be used to identify bacteria with new AI-2 receptors from microbial complex niches. By optimizing the length of the linker we might be able to employ streptavidin-bound fluorophores (173), bound to D-desthiobiotin-AI-2 to label cells with receptor proteins. This optimization would not require the development of a new synthesis process as the binding of AI-2 to TentaGel resins already employed a bigger linker and should, therefore, be relatively straightforward.

The already mentioned chemical characteristics of AI-2 also make its isolation from complex biological samples and subsequent quantification, a challenge. Steps involving heating or increasing the pH of the samples has to be avoided (51,182). The conversion of DPD to a quinoxaline derivative, DPD-Q, however, stabilizes the molecule allowing evaporation of the solvent without loss of AI-2. Nevertheless, it also introduced a new challenge. The caecum of GF mice seemed to contain a contaminant that reacts with the tag and that might be similar to AI-2 as it has a peak near the peak observed for DPD-Q. This observation made the optimization of the concentration of tag used a priority, as the derivatization of DPD is a fundamental step to be able to detect AI-2 through mass spectrometry. We observed that we could not derivatize DPD in the beginning of the experiment (to increase the stability of the molecule) as the other large present in the caecum seemed to interfere with this reaction possibly because they bind DPD or react with the tag, preventing the specific formation of DPD-Q. Actually, we tested derivatizing DPD after performing only the filtration step through a 10 kDa membrane. Thus, we eliminated the filtration through a 3 kDa filter, the last step in the purification process before derivatization. This alteration rendered us

unable to identify AI-2. Nevertheless, filtration through a 3 kDa membrane of these samples, followed by a new derivatization allowed detection of AI-2. This observation suggests that the presence of high molecular weight compounds interferes with the ability to derivatize AI-2. Therefore, the filtration revealed to be essential for the derivatization of AI-2 present in caecum samples. Additionally, these filtration steps also seem to be removing contaminants that inhibit the ability to detect AI-2 through the *Vibrio* bioassay, as a slight increase in AI-2 detection seemed to occur as samples were submitted to more filtration processes.

By optimizing the purification of AI-2 from caecum extracts and by adapting the mass spectrometry method described by Campagna et al, AI-2 could be detected in the caecum of mice monocolonized with an *E. coli* strain able to accumulate AI-2. However, we could not detect AI-2 in the caecum of SPF mice either due to its absence or due to its presence at a concentration below the limit of detection. On a positive note, we were able to obtain a calibration curve of DPD-Q in GF caecum extracts that will allow quantification of this compound in biological samples. DPD-Q was quantified in samples with concentration ranging from 0.25 μ M to 50 μ M without losing linearity. In the future, lower concentrations of DPD-Q will also be tested to define the limit of detection. The optimization and validation of this quantification method will allow the determination of AI-2 concentration in bacterial communities of the gut with different compositions. Thus, it will be interesting to test if there are conditions where AI-2 concentration increases due to microbiota composition changes. For example, if the ratio between Firmicutes and Bacteroidetes was altered and more Firmicutes were present, the AI-2 concentration could be higher than in a community with a ratio of approximately 1. This expectation is based on the observation that around 80% of the Firmicutes are potential AI-2 producers whereas only approximately 20% of the Bacteroidetes have homologs for the AI-2 synthase (21). Nevertheless, as we showed that Firmicutes members are able to internalize AI-2, we cannot exclude the possibility of being unable to detect AI-2 due to fluctuating levels of AI-2

caused by the production and internalization of AI-2 by different bacterial species, similarly to what we hypothesized happens in the caecum of SPF mice. Moreover, mice subjected to different diets might have different concentrations of AI-2. The Lsr system is regulated by catabolite repression, being repressed in the presence of glucose. The repression of the AI-2 transporter system leads to the accumulation of AI-2 in the extracellular medium. Thus, by adding high levels of glucose either to mice food or to the drinking water, an increase in extracellular AI-2 concentration might be observed.

Additionally, modulatory effects of AI-2 analogs in the microbiota could also be studied. One of the analogs that could be tested is (4S)-4-propoxy-HPD, which is two times more active than AI-2 and is not degraded by bacteria with LsrB receptors. As it is specific for LuxP receptors, it would be interesting to investigate if *Vibrio cholerae*, a well-known human pathogen, would produce less virulence factors in the presence of the analog in comparison to the same concentration of AI-2. The mass spectrometry-based method to quantify AI-2 could be adjusted to quantify the (4S)-4-propoxy-HPD as the modification on the carbon 4 should not hinder derivatization. Thus, upon determination of the molecular and fragmentation ions, the concentration of (4S)-4-propoxy-HPD in the intestines could be compared to the one of AI-2.

All together, the work developed in this doctoral thesis increases the knowledge about LsrB receptors and provides the basis for the discovery of novel AI-2 receptors. The discovery of AI-2 receptors is the first step to understand the molecular mechanisms that underlie AI-2 mediated communication. Thus, identification of AI-2 receptors in bacteria belonging to complex, biological niches will help to understand how interspecies communication shapes these environments. Moreover, it will provide the basis for the development of new antibacterial strategies that alter this communication process. As interfering with quorum sensing does not induce cell death, these strategies are less likely to

promote selection for resistance, one of the biggest threats to global health today. Additionally, we studied the biological activity of a potential modulator and started the development of a methodology that will allow the study of the influence of AI-2 in one of these complex niches: the mammalian gut microbiota.

References

1. Nealson, K. H., Platt, T., and Hastings, J. W. (1970) Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol.* **104**, 313-322
2. Waters, C. M., and Bassler, B. L. (2005) Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell. Dev. Bi.* **21**, 319-346
3. Papenfort, K., and Bassler, B. L. (2016) Quorum sensing signal-response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* **14**, 576-588
4. Bassler, B. L., and Losick, R. (2006) Bacterially speaking. *Cell* **125**, 237-246
5. Tomasz, A., and Hotchkiss, R. D. (1964) Regulation of the Transformability of *Pneumococcal* Cultures by Macromolecular Cell Products. *Proc Natl Acad Sci U S A* **51**, 480-487
6. Engebrecht, J., and Silverman, M. (1984) Identification of genes and gene products necessary for bacterial bioluminescence. *Proc Natl Acad Sci U S A* **81**, 4154-4158
7. Eberhard, A., Burlingame, A. L., Eberhard, C., Kenyon, G. L., Nealson, K. H., and Oppenheimer, N. J. (1981) Structural identification of autoinducer of *Photobacterium fischeri* luciferase. *Biochemistry* **20**, 2444-2449
8. Whiteley, M., Diggle, S. P., and Greenberg, E. P. (2017) Progress in and promise of bacterial quorum sensing research. *Nature* **551**, 313-320
9. Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* **176**, 269-275
10. Havarstein, L. S., Coomaraswamy, G., and Morrison, D. A. (1995) An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A* **92**, 11140-11144
11. Ji, G., Beavis, R. C., and Novick, R. P. (1995) Cell density control of *staphylococcal* virulence mediated by an octapeptide pheromone. *Proc Natl Acad Sci U S A* **92**, 12055-12059
12. Bassler, B. L., Wright, M., and Silverman, M. R. (1994) Multiple signalling systems controlling expression of luminescence in *Vibrio harveyi*: sequence and function of genes encoding a second sensory pathway. *Mol. Microbiol.* **13**, 273-286
13. Greenberg, E. P. H., J. W.; Ulitzur, S. (1979) Induction of Luciferase Synthesis in *Beneckeia harveyi* by Other Marine Bacteria. *Arch Microbiol* **120**, 87-91

14. Bassler, B. L., Greenberg, E. P., and Stevens, A. M. (1997) Cross-species induction of luminescence in the quorum-sensing bacterium *Vibrio harveyi*. *J. Bacteriol.* **179**, 4043-4045
15. Bassler, B. L. (1999) How bacteria talk to each other: regulation of gene expression by quorum sensing. *Curr. Opin. Microbiol.* **2**, 582-587
16. Federle, M. J. (2009) Autoinducer-2-based chemical communication in bacteria: complexities of interspecies signaling. *Contributions to microbiology* **16**, 18-32
17. Pereira, C. S., de Regt, A. K., Brito, P. H., Miller, S. T., and Xavier, K. B. (2009) Identification of functional LsrB-like autoinducer-2 receptors. *J. Bacteriol.* **191**, 6975-6987
18. Kaur, A., Capalash, N., and Sharma, P. (2019) Communication mechanisms in extremophiles: Exploring their existence and industrial applications. *Microbiol. Res.* **221**, 15-27
19. Mukherjee, S., and Bassler, B. L. (2019) Bacterial quorum sensing in complex and dynamically changing environments. *Nat. Rev. Microbiol.* **17**, 371-382
20. Nadell, C. D., Ricaurte, D., Yan, J., Drescher, K., and Bassler, B. L. (2017) Flow environment and matrix structure interact to determine spatial competition in *Pseudomonas aeruginosa* biofilms. *eLife* **6**, e21855
21. Thompson, J. A., Oliveira, R. A., Djukovic, A., Ubeda, C., and Xavier, K. B. (2015) Manipulation of the quorum sensing signal AI-2 affects the antibiotic-treated gut microbiota. *Cell Rep* **10**, 1861-1871
22. Nealson, K. H., and Hastings, J. W. (1979) Bacterial bioluminescence: its control and ecological significance. *Microbiol. Rev* **43**, 496-518
23. Lee, K. H., and Ruby, E. G. (1994) Competition between *Vibrio fischeri* strains during initiation and maintenance of a light organ symbiosis. *J. Bacteriol.* **176**, 1985-1991
24. LaSarre, B., and Federle, M. J. (2013) Exploiting quorum sensing to confuse bacterial pathogens. *Microbiol. Mol. Bio. Rev.* **77**, 73-111
25. Parsek, M. R., Val, D. L., Hanzelka, B. L., Cronan, J. E., Jr., and Greenberg, E. P. (1999) Acyl homoserine-lactone quorum-sensing signal generation. *Proc Natl Acad Sci U S A* **96**, 4360-4365
26. Chan, Y. Y., Bian, H. S., Tan, T. M., Mattmann, M. E., Geske, G. D., Igarashi, J., Hatano, T., Suga, H., Blackwell, H. E., and Chua, K. L. (2007) Control of quorum sensing by a *Burkholderia pseudomallei* multidrug efflux pump. *J. Bacteriol* **189**, 4320-4324

References

27. Stevens, A. M., Dolan, K. M., and Greenberg, E. P. (1994) Synergistic binding of the *Vibrio fischeri* LuxR transcriptional activator domain and RNA polymerase to the lux promoter region. *Proc Natl Acad Sci U S A* **91**, 12619-12623
28. Monnet, V., Juillard, V., and Gardan, R. (2016) Peptide conversations in Gram-positive bacteria. *Crit. Rev. Microbiol.* **42**, 339-351
29. Zhou, L., Zhang, L. H., Camara, M., and He, Y. W. (2017) The DSF Family of Quorum Sensing Signals: Diversity, Biosynthesis, and Turnover. *Trends in microbiology* **25**, 293-303
30. Deng, Y., Wu, J., Yin, W., Li, P., Zhou, J., Chen, S., He, F., Cai, J., and Zhang, L. H. (2016) Diffusible signal factor family signals provide a fitness advantage to *Xanthomonas campestris* pv. *campestris* in interspecies competition. *Environ. Microbiol.* **18**, 1534-1545
31. Dean, S. N., Chung, M. C., and van Hoek, M. L. (2015) *Burkholderia* Diffusible Signal Factor Signals to *Francisella novicida* To Disperse Biofilm and Increase Siderophore Production. *Appl. Environ. Microbiol.* **81**, 7057-7066
32. Vilchez, R., Lemme, A., Ballhausen, B., Thiel, V., Schulz, S., Jansen, R., Sztajer, H., and Wagner-Dobler, I. (2010) *Streptococcus mutans* inhibits *Candida albicans* hyphal formation by the fatty acid signaling molecule trans-2-decenoic acid (SDSF). *ChemBioChem* **11**, 1552-1562
33. Boon, C., Deng, Y., Wang, L. H., He, Y., Xu, J. L., Fan, Y., Pan, S. Q., and Zhang, L. H. (2008) A novel DSF-like signal from *Burkholderia cenocepacia* interferes with *Candida albicans* morphological transition. *The ISME journal* **2**, 27-36
34. He, Y. W., Xu, M., Lin, K., Ng, Y. J., Wen, C. M., Wang, L. H., Liu, Z. D., Zhang, H. B., Dong, Y. H., Dow, J. M., and Zhang, L. H. (2006) Genome scale analysis of diffusible signal factor regulon in *Xanthomonas campestris* pv. *campestris*: identification of novel cell-cell communication-dependent genes and functions. *Mol. Microbiol.* **59**, 610-622
35. Slater, H., Alvarez-Morales, A., Barber, C. E., Daniels, M. J., and Dow, J. M. (2000) A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. *Mol. Microbiol.* **38**, 986-1003
36. Tao, F., He, Y. W., Wu, D. H., Swarup, S., and Zhang, L. H. (2010) The cyclic nucleotide monophosphate domain of *Xanthomonas campestris* global regulator Clp defines a new class of cyclic di-GMP effectors. *J. Bacteriol.* **192**, 1020-1029
37. Zhou, L., Yu, Y., Chen, X., Diab, A. A., Ruan, L., He, J., Wang, H., and He, Y. W. (2015) The Multiple DSF-family QS Signals are Synthesized from Carbohydrate and Branched-chain Amino Acids via the FAS Elongation Cycle. *Sci. Rep.* **5**, 13294

38. Deng, Y., Lim, A., Wang, J., Zhou, T., Chen, S., Lee, J., Dong, Y. H., and Zhang, L. H. (2013) Cis-2-dodecenoic acid quorum sensing system modulates N-acyl homoserine lactone production through RpfR and cyclic di-GMP turnover in *Burkholderia cenocepacia*. *BMC microbiology* **13**, 148
39. Suppiger, A., Eshwar, A. K., Stephan, R., Kaever, V., Eberl, L., and Lehner, A. (2016) The DSF type quorum sensing signalling system RpfF/R regulates diverse phenotypes in the opportunistic pathogen *Cronobacter*. *Sci. Rep.* **6**, 18753
40. Deng, Y., Schmid, N., Wang, C., Wang, J., Pessi, G., Wu, D., Lee, J., Aguilar, C., Ahrens, C. H., Chang, C., Song, H., Eberl, L., and Zhang, L. H. (2012) Cis-2-dodecenoic acid receptor RpfR links quorum-sensing signal perception with regulation of virulence through cyclic dimeric guanosine monophosphate turnover. *Proc Natl Acad Sci U S A* **109**, 15479-15484
41. Amari, D. T., Marques, C. N., and Davies, D. G. (2013) The putative enoyl-coenzyme A hydratase Dspl is required for production of the *Pseudomonas aeruginosa* biofilm dispersion autoinducer cis-2-decenoic acid. *J. Bacteriol.* **195**, 4600-4610
42. Rahmani-Badi, A., Sepehr, S., Fallahi, H., and Heidari-Keshel, S. (2015) Dissection of the cis-2-decenoic acid signaling network in *Pseudomonas aeruginosa* using microarray technique. *Front. Microbiol.* **6**, 383
43. Bi, H., Yu, Y., Dong, H., Wang, H., and Cronan, J. E. (2014) *Xanthomonas campestris* RpfB is a fatty Acyl-CoA ligase required to counteract the thioesterase activity of the RpfF diffusible signal factor (DSF) synthase. *Mol. Microbiol.* **93**, 262-275
44. Espinosa-Urgel, M. (2016) Learning when (and how) to shut up: intercellular signal turnover in *Xanthomonas*. *Environ. Microbiol.* **18**, 314-315
45. Bassler, B. L., Wright, M., Showalter, R. E., and Silverman, M. R. (1993) Intercellular Signaling in *Vibrio-Harveyi* - Sequence and Function of Genes Regulating Expression of Luminescence. *Mol. Microbiol.* **9**, 773-786
46. Miller, S. T., Xavier, K. B., Campagna, S. R., Taga, M. E., Semmelhack, M. F., Bassler, B. L., and Hughson, F. M. (2004) *Salmonella typhimurium* recognizes a chemically distinct form of the bacterial quorum-sensing signal AI-2. *Mol. Cell.* **15**, 677-687
47. Chen, X., Schauder, S., Potier, N., Van Dorsselaer, A., Pelczer, I., Bassler, B. L., and Hughson, F. M. (2002) Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* **415**, 545-549
48. Schauder, S., Shokat, K., Surette, M. G., and Bassler, B. L. (2001) The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Mol. Microbiol.* **41**, 463-476

References

49. Xavier, K. B., and Bassler, B. L. (2003) LuxS quorum sensing: more than just a numbers game. *Curr. Opin Microbiol.* **6**, 191-197
50. Pereira, C. S., Thompson, J. A., and Xavier, K. B. (2013) AI-2-mediated signalling in bacteria. *FEMS microbiology reviews* **37**, 156-181
51. Ascenso, O. S., Marques, J. C., Santos, A. R., Xavier, K. B., Ventura, M. R., and Maycock, C. D. (2011) An efficient synthesis of the precursor of AI-2, the signalling molecule for inter-species quorum sensing. *Bioorg. Med. Chem.* **19**, 1236-1241
52. Xavier, K. B., and Bassler, B. L. (2005) Regulation of uptake and processing of the quorum-sensing autoinducer AI-2 in *Escherichia coli*. *J. Bacteriol.* **187**, 238-248
53. Kamaraju, K., Smith, J., Wang, J., Roy, V., Sintim, H. O., Bentley, W. E., and Sukharev, S. (2011) Effects on membrane lateral pressure suggest permeation mechanisms for bacterial quorum signaling molecules. *Biochemistry* **50**, 6983-6993
54. Herzberg, M., Kaye, I. K., Peti, W., and Wood, T. K. (2006) YdgG (TqsA) controls biofilm formation in *Escherichia coli* K-12 through autoinducer 2 transport. *J. Bacteriol.* **188**, 587-598
55. Rettner, R. E., and Saier, M. H., Jr. (2010) The autoinducer-2 exporter superfamily. *J. Mol. Microbiol. Biotechnol.* **18**, 195-205
56. De Araujo, C., Balestrino, D., Roth, L., Charbonnel, N., and Forestier, C. (2010) Quorum sensing affects biofilm formation through lipopolysaccharide synthesis in *Klebsiella pneumoniae*. *Res. Microbiol.* **161**, 595-603
57. Lee, A. S., and Song, K. P. (2005) LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. *Biochem. Biophys. Res. Commun.* **335**, 659-666
58. Rossmann, F. S., Racek, T., Wobser, D., Puchalka, J., Rabener, E. M., Reiger, M., Hendrickx, A. P., Diederich, A. K., Jung, K., Klein, C., and Huebner, J. (2015) Phage-mediated dispersal of biofilm and distribution of bacterial virulence genes is induced by quorum sensing. *PLoS Pathog.* **11**, e1004653
59. Neiditch, M. B., Federle, M. J., Miller, S. T., Bassler, B. L., and Hughson, F. M. (2005) Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. *Mol. Cell.* **18**, 507-518
60. Neiditch, M. B., Federle, M. J., Pompeani, A. J., Kelly, R. C., Swem, D. L., Jeffrey, P. D., Bassler, B. L., and Hughson, F. M. (2006) Ligand-induced asymmetry in histidine sensor kinase complex regulates quorum sensing. *Cell* **126**, 1095-1108
61. Surette, M. G., Miller, M. B., and Bassler, B. L. (1999) Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proc Natl Acad Sci U S A* **96**, 1639-1644

62. Taga, M. E., Miller, S. T., and Bassler, B. L. (2003) Lsr-mediated transport and processing of AI-2 in *Salmonella typhimurium*. *Mol. Microbiol.* **50**, 1411-1427
63. Taga, M. E., Semmelhack, J. L., and Bassler, B. L. (2001) The LuxS-dependent autoinducer AI-2 controls the expression of an ABC transporter that functions in AI-2 uptake in *Salmonella typhimurium*. *Mol. Microbiol.* **42**, 777-793
64. Kavanaugh, J. S., Gakhar, L., and Horswill, A. R. (2011) The structure of LsrB from *Yersinia pestis* complexed with autoinducer-2. *Acta Crystallogr F* **67**, 1501-1505
65. Pereira, C. S., McAuley, J. R., Taga, M. E., Xavier, K. B., and Miller, S. T. (2008) *Sinorhizobium meliloti*, a bacterium lacking the autoinducer-2 (AI-2) synthase, responds to AI-2 supplied by other bacteria. *Mol. Microbiol.* **70**, 1223-1235
66. Wang, L., Li, J., March, J. C., Valdes, J. J., and Bentley, W. E. (2005) luxS-dependent gene regulation in *Escherichia coli* K-12 revealed by genomic expression profiling. *J. Bacteriol.* **187**, 8350-8360
67. Xavier, K. B., Miller, S. T., Lu, W. Y., Kim, J. H., Rabinowitz, J., Pelczar, I., Semmelhack, M. F., and Bassler, B. L. (2007) Phosphorylation and processing of the quorum-sensing molecule autoinducer-2 in enteric bacteria. *Acs Chem. Biol.* **2**, 128-136
68. Wu, M., Tao, Y., Liu, X., and Zang, J. (2013) Structural basis for phosphorylated autoinducer-2 modulation of the oligomerization state of the global transcription regulator LsrR from *Escherichia coli*. *J. Biol. Chem.* **288**, 15878-15887
69. Xue, T., Zhao, L., Sun, H., Zhou, X., and Sun, B. (2009) LsrR-binding site recognition and regulatory characteristics in *Escherichia coli* AI-2 quorum sensing. *Cell Res.* **19**, 1258-1268
70. Marques, J. C., Lamosa, P., Russell, C., Ventura, R., Maycock, C., Semmelhack, M. F., Miller, S. T., and Xavier, K. B. (2011) Processing the interspecies quorum-sensing signal autoinducer-2 (AI-2): characterization of phospho-(S)-4,5-dihydroxy-2,3-pentanedione isomerization by LsrG protein. *J. Biol. Chem.* **286**, 18331-18343
71. Marques, J. C., Oh, I. K., Ly, D. C., Lamosa, P., Ventura, M. R., Miller, S. T., and Xavier, K. B. (2014) LsrF, a coenzyme A-dependent thiolase, catalyzes the terminal step in processing the quorum sensing signal autoinducer-2. *P Natl Acad Sci USA* **111**, 14235-14240
72. Ha, J. H., Hauk, P., Cho, K., Eo, Y., Ma, X., Stephens, K., Cha, S., Jeong, M., Suh, J. Y., Sintim, H. O., Bentley, W. E., and Ryu, K. S. (2018) Evidence of link between quorum sensing and sugar metabolism in *Escherichia coli* revealed via cocrystal structures of LsrK and HPr. *Sci. Adv.* **4**, eaar7063

References

73. Pereira, C. S., Santos, A. J., Bejerano-Sagie, M., Correia, P. B., Marques, J. C., and Xavier, K. B. (2012) Phosphoenolpyruvate phosphotransferase system regulates detection and processing of the quorum sensing signal autoinducer-2. *Mol. Microbiol.* **84**, 93-104
74. Federle, M. J., and Bassler, B. L. (2003) Interspecies communication in bacteria. *J. Clin. Invest.* **112**, 1291-1299
75. Jani, S., Seely, A. L., Peabody, G. L., Jayaraman, A., and Manson, M. D. (2017) Chemotaxis to self-generated AI-2 promotes biofilm formation in *Escherichia coli*. *Microbiol-Sgm* **163**, 1778-1790
76. Hegde, M., Englert, D. L., Schrock, S., Cohn, W. B., Vogt, C., Wood, T. K., Manson, M. D., and Jayaraman, A. (2011) Chemotaxis to the quorum-sensing signal AI-2 requires the Tsr chemoreceptor and the periplasmic LsrB AI-2-binding protein. *J. Bacteriol.* **193**, 768-773
77. Laganenka, L., Colin, R., and Sourjik, V. (2016) Chemotaxis towards autoinducer 2 mediates autoaggregation in *Escherichia coli*. *Nat. Commun.* **7**, 12984
78. Li, J., Attila, C., Wang, L., Wood, T. K., Valdes, J. J., and Bentley, W. E. (2007) Quorum sensing in *Escherichia coli* is signaled by AI-2/LsrR: Effects on small RNA and Biofilm architecture. *J. Bacteriol.* **189**, 6011-6020
79. Wood, T. K., Gonzalez Barrios, A. F., Herzberg, M., and Lee, J. (2006) Motility influences biofilm architecture in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **72**, 361-367
80. Xavier, K. B., and Bassler, B. L. (2005) Interference with AI-2-mediated bacterial cell-cell communication. *Nature* **437**, 750-753
81. Krin, E., Chakroun, N., Turlin, E., Givaudan, A., Gaboriau, F., Bonne, I., Rousselle, J. C., Frangeul, L., Lacroix, C., Hullo, M. F., Marisa, L., Danchin, A., and Derzelle, S. (2006) Pleiotropic role of quorum-sensing autoinducer 2 in *Photorhabdus luminescens*. *Appl. Environ. Microbiol.* **72**, 6439-6451
82. Shao, H., James, D., Lamont, R. J., and Demuth, D. R. (2007) Differential interaction of *Aggregatibacter (Actinobacillus) actinomycetemcomitans* LsrB and RbsB proteins with autoinducer 2. *J. Bacteriol.* **189**, 5559-5565
83. Bobrov, A. G., Bearden, S. W., Fetherston, J. D., Khweek, A. A., Parrish, K. D., and Perry, R. D. (2007) Functional quorum sensing systems affect biofilm formation and protein expression in *Yersinia pestis*. *Adv. Exp. Med. Biol.* **603**, 178-191
84. Armbruster, C. E., Pang, B., Murrah, K., Juneau, R. A., Perez, A. C., Weimer, K. E., and Swords, W. E. (2011) RbsB (NTHI_0632) mediates quorum signal uptake in nontypeable *Haemophilus influenzae* strain 86-028NP. *Mol. Microbiol.* **82**, 836-850

85. Pang, B., Armbruster, C. E., Foster, G., Learman, B. S., Gandhi, U., and Swords, W. E. (2018) Autoinducer 2 (AI-2) Production by Nontypeable *Haemophilus influenzae* 86-028NP Promotes Expression of a Predicted Glycosyltransferase That Is a Determinant of Biofilm Maturation, Prevention of Dispersal, and Persistence *In Vivo*. *Infect. Immun.* **86**, e00506-18
86. Ryu, E. J., Sim, J., Sim, J., Lee, J., and Choi, B. K. (2016) D-Galactose as an autoinducer 2 inhibitor to control the biofilm formation of periodontopathogens. *J. Microbiol.* **54**, 632-637
87. Trappetti, C., McAllister, L. J., Chen, A., Wang, H., Paton, A. W., Oggioni, M. R., McDevitt, C. A., and Paton, J. C. (2017) Autoinducer 2 Signaling via the Phosphotransferase FruA Drives Galactose Utilization by *Streptococcus pneumoniae*, Resulting in Hypervirulence. *mBio* **8**, e02269-16
88. Graff, S. M., and Bentley, W. E. (2017) Mathematical model of LsrR-binding and derepression in *Escherichia coli* K12. *J. Bioinf. Comput. Biol.* **15**, 1650039
89. Anderson, J. K., Huang, J. Y., Wreden, C., Sweeney, E. G., Goers, J., Remington, S. J., and Guillemin, K. (2015) Chemorepulsion from the Quorum Signal Autoinducer-2 Promotes *Helicobacter pylori* Biofilm Dispersal. *mBio* **6**, e00379
90. Rader, B. A., Wreden, C., Hicks, K. G., Sweeney, E. G., Ottemann, K. M., and Guillemin, K. (2011) *Helicobacter pylori* perceives the quorum-sensing molecule AI-2 as a chemorepellent via the chemoreceptor TlpB. *Microbiology* **157**, 2445-2455
91. Seed, P. C., Passador, L., and Iglewski, B. H. (1995) Activation of the *Pseudomonas aeruginosa* lasI gene by LasR and the *Pseudomonas* autoinducer PAI: an autoinduction regulatory hierarchy. *J. Bacteriol.* **177**, 654-659
92. Lee, J., and Zhang, L. (2015) The hierarchy quorum sensing network in *Pseudomonas aeruginosa*. *Protein & cell* **6**, 26-41
93. Lee, J., Wu, J., Deng, Y., Wang, J., Wang, C., Wang, J., Chang, C., Dong, Y., Williams, P., and Zhang, L. H. (2013) A cell-cell communication signal integrates quorum sensing and stress response. *Nat. Chem. Biol.* **9**, 339-343
94. Hoffman, L. R., Kulasekara, H. D., Emerson, J., Houston, L. S., Burns, J. L., Ramsey, B. W., and Miller, S. I. (2009) *Pseudomonas aeruginosa* lasR mutants are associated with cystic fibrosis lung disease progression. *J. Cys. Fibros.* **8**, 66-70
95. Mukherjee, S., Moustafa, D. A., Stergioula, V., Smith, C. D., Goldberg, J. B., and Bassler, B. L. (2018) The PqsE and RhlR proteins are an autoinducer synthase-receptor pair that control virulence and biofilm development in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **115**, E9411-E9418

References

96. Freeman, J. A., and Bassler, B. L. (1999) A genetic analysis of the function of LuxO, a two-component response regulator involved in quorum sensing in *Vibrio harveyi*. *Mol. Microbiol.* **31**, 665-677
97. Lilley, B. N., and Bassler, B. L. (2000) Regulation of quorum sensing in *Vibrio harveyi* by LuxO and sigma-54. *Mol. Microbiol.* **36**, 940-954
98. Lenz, D. H., Mok, K. C., Lilley, B. N., Kulkarni, R. V., Wingreen, N. S., and Bassler, B. L. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118**, 69-82
99. Papenfort, K., Silpe, J. E., Schramma, K. R., Cong, J. P., Seyedsayamdost, M. R., and Bassler, B. L. (2017) A *Vibrio cholerae* autoinducer-receptor pair that controls biofilm formation. *Nat. Chem. Biol.* **13**, 551-557
100. Venturi, V., and Ahmer, B. M. (2015) Editorial: LuxR Solos are Becoming Major Players in Cell-Cell Communication in Bacteria. *Front. Cell. Infect. Microbiol.* **5**, 89
101. Lee, J. H., Lequette, Y., and Greenberg, E. P. (2006) Activity of purified QscR, a *Pseudomonas aeruginosa* orphan quorum-sensing transcription factor. *Mol. Microbiol.* **59**, 602-609
102. Riedel, K., Hentzer, M., Geisenberger, O., Huber, B., Steidle, A., Wu, H., Hoiby, N., Givskov, M., Molin, S., and Eberl, L. (2001) N-acylhomoserine-lactone-mediated communication between *Pseudomonas aeruginosa* and *Burkholderia cepacia* in mixed biofilms. *Microbiology* **147**, 3249-3262
103. Valente, R. S., Nadal-Jimenez, P., Carvalho, A. F. P., Vieira, F. J. D., and Xavier, K. B. (2017) Signal Integration in Quorum Sensing Enables Cross-Species Induction of Virulence in *Pectobacterium wasabiae*. *mBio* **8**, e00398-17
104. Li, H., Li, X., Song, C., Zhang, Y., Wang, Z., Liu, Z., Wei, H., and Yu, J. (2017) Autoinducer-2 Facilitates *Pseudomonas aeruginosa* PAO1 Pathogenicity *in Vitro* and *in Vivo*. *Front. Microbiol.* **8**, 1944
105. Gorelik, O., Levy, N., Shaulov, L., Yegodayev, K., Meijler, M. M., and Sal-Man, N. (2019) *Vibrio cholerae* autoinducer-1 enhances the virulence of enteropathogenic *Escherichia coli*. *Sci. Rep.* **9**, 4122
106. Sender, R., Fuchs, S., and Milo, R. (2016) Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* **14**, e1002533
107. Sommer, F., and Backhed, F. (2013) The gut microbiota--masters of host development and physiology. *Nat. Rev. Microbiol.* **11**, 227-238
108. Kendall, M. M., and Sperandio, V. (2016) What a Dinner Party! Mechanisms and Functions of Interkingdom Signaling in Host-Pathogen Associations. *mBio* **7**, e01748

109. Lustri, B. C., Sperandio, V., and Moreira, C. G. (2017) Bacterial Chat: Intestinal Metabolites and Signals in Host-Microbiota-Pathogen Interactions. *Infect Immun.* **85**, e00476-17
110. Sperandio, V., Torres, A. G., Jarvis, B., Nataro, J. P., and Kaper, J. B. (2003) Bacteria-host communication: the language of hormones. *Proc Natl Acad Sci U S A* **100**, 8951-8956
111. Li, Q., Ren, Y., and Fu, X. (2019) Inter-kingdom signaling between gut microbiota and their host. *Cell. Mol. Life Sci.* **76**, 2383-2389.
112. Clarke, M. B., and Sperandio, V. (2005) Transcriptional regulation of flhDC by QseBC and sigma (FliA) in enterohaemorrhagic *Escherichia coli*. *Mol. Microbiol.* **57**, 1734-1749
113. Vlisidou, I., Lyte, M., van Diemen, P. M., Hawes, P., Monaghan, P., Wallis, T. S., and Stevens, M. P. (2004) The neuroendocrine stress hormone norepinephrine augments *Escherichia coli* O157:H7-induced enteritis and adherence in a bovine ligated ileal loop model of infection. *Infect. Immun.* **72**, 5446-5451
114. Clarke, M. B., Hughes, D. T., Zhu, C., Boedeker, E. C., and Sperandio, V. (2006) The QseC sensor kinase: a bacterial adrenergic receptor. *Proc Natl Acad Sci U S A* **103**, 10420-10425
115. Hadjifrangiskou, M., Kostakioti, M., Chen, S. L., Henderson, J. P., Greene, S. E., and Hultgren, S. J. (2011) A central metabolic circuit controlled by QseC in pathogenic *Escherichia coli*. *Mol. Microbiol.* **80**, 1516-1529
116. Zaborina, O., Lepine, F., Xiao, G., Valuckaite, V., Chen, Y., Li, T., Ciancio, M., Zaborin, A., Petrof, E. O., Turner, J. R., Rahme, L. G., Chang, E., and Alverdy, J. C. (2007) Dynorphin activates quorum sensing quinolone signaling in *Pseudomonas aeruginosa*. *PLoS Pathog.* **3**, e35
117. Ismail, A. S., Valastyan, J. S., and Bassler, B. L. (2016) A Host-Produced Autoinducer-2 Mimic Activates Bacterial Quorum Sensing. *Cell Host Microbe* **19**, 470-480
118. Autieri, S. M., Lins, J. J., Leatham, M. P., Laux, D. C., Conway, T., and Cohen, P. S. (2007) L-fucose stimulates utilization of D-ribose by *Escherichia coli* MG1655 DeltafucAO and *E. coli* Nissle 1917 DeltafucAO mutants in the mouse intestine and in M9 minimal medium. *Infect. Immun.* **75**, 5465-5475
119. Smith, R. S., Harris, S. G., Phipps, R., and Iglewski, B. (2002) The *Pseudomonas aeruginosa* quorum-sensing molecule N-(3-oxododecanoyl)homoserine lactone contributes to virulence and induces inflammation in vivo. *J. Bacteriol.* **184**, 1132-1139

References

120. Zargar, A., Quan, D. N., Carter, K. K., Guo, M., Sintim, H. O., Payne, G. F., and Bentley, W. E. (2015) Bacterial secretions of nonpathogenic *Escherichia coli* elicit inflammatory pathways: a closer investigation of interkingdom signaling. *mBio* **6**, e00025
121. Skindersoe, M. E., Zeuthen, L. H., Brix, S., Fink, L. N., Lazenby, J., Whittall, C., Williams, P., Diggle, S. P., Froekiaer, H., Cooley, M., and Givskov, M. (2009) *Pseudomonas aeruginosa* quorum-sensing signal molecules interfere with dendritic cell-induced T-cell proliferation. *FEMS Immunol. Med. Microbiol.* **55**, 335-345
122. Kravchenko, V. V., Kaufmann, G. F., Mathison, J. C., Scott, D. A., Katz, A. Z., Grauer, D. C., Lehmann, M., Meijler, M. M., Janda, K. D., and Ulevitch, R. J. (2008) Modulation of gene expression via disruption of NF-kappaB signaling by a bacterial small molecule. *Science* **321**, 259-263
123. Jahoor, A., Patel, R., Bryan, A., Do, C., Krier, J., Watters, C., Wahli, W., Li, G., Williams, S. C., and Rumbaugh, K. P. (2008) Peroxisome proliferator-activated receptors mediate host cell proinflammatory responses to *Pseudomonas aeruginosa* autoinducer. *J. Bacteriol.* **190**, 4408-4415
124. Wynendaele, E., Verbeke, F., D'Hondt, M., Hendrix, A., Van De Wiele, C., Burvenich, C., Peremans, K., De Wever, O., Bracke, M., and De Spiegeleer, B. (2015) Crosstalk between the microbiome and cancer cells by quorum sensing peptides. *Peptides* **64**, 40-48
125. Yeo, S., Park, H., Ji, Y., Park, S., Yang, J., Lee, J., Mathara, J. M., Shin, H., and Holzapfel, W. (2015) Influence of gastrointestinal stress on autoinducer-2 activity of two *Lactobacillus* species. *FEMS Microbiol. Eco.* **91**, fiv065
126. Kim, M. K., Ingremeau, F., Zhao, A., Bassler, B. L., and Stone, H. A. (2016) Local and global consequences of flow on bacterial quorum sensing. *Nat. Microbiol.* **1**, 15005
127. Sandoz, K. M., Mitzimberg, S. M., and Schuster, M. (2007) Social cheating in *Pseudomonas aeruginosa* quorum sensing. *Proc Natl Acad Sci U S A* **104**, 15876-15881
128. Drescher, K., Nadell, C. D., Stone, H. A., Wingreen, N. S., and Bassler, B. L. (2014) Solutions to the public goods dilemma in bacterial biofilms. *Curr. Biol.* **24**, 50-55
129. Ozkaya, O., Balbontin, R., Gordo, I., and Xavier, K. B. (2018) Cheating on Cheaters Stabilizes Cooperation in *Pseudomonas aeruginosa*. *Curr. Biol.* **28**, 2070-2080 e2076
130. Grote, J., Krysciak, D., and Streit, W. R. (2015) Phenotypic Heterogeneity, a Phenomenon That May Explain Why Quorum Sensing Does Not Always Result in Truly Homogenous Cell Behavior. *Appl. Environ. Microbiol.* **81**, 5280-5289
131. Veening, J. W., Smits, W. K., and Kuipers, O. P. (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* **62**, 193-210

132. Jung, S. A., Chapman, C. A., and Ng, W. L. (2015) Quadruple quorum-sensing inputs control *Vibrio cholerae* virulence and maintain system robustness. *PLoS Pathog.* **11**, e1004837
133. Bansal, T., Jesudhasan, P., Pillai, S., Wood, T. K., and Jayaraman, A. (2008) Temporal regulation of enterohemorrhagic *Escherichia coli* virulence mediated by autoinducer-2. *Appl. Microbiol. Biot.* **78**, 811-819
134. Zhang, Y. J., Li, S., Gan, R. Y., Zhou, T., Xu, D. P., and Li, H. B. (2015) Impacts of gut bacteria on human health and diseases. *Int. J. Mol. Sci.* **16**, 7493-7519
135. Buffie, C. G., Bucci, V., Stein, R. R., McKenney, P. T., Ling, L. L., Gobourne, A., No, D., Liu, H., Kinnebrew, M., Viale, A., Littmann, E., van den Brink, M. R. M., Jenq, R. R., Taur, Y., Sander, C., Cross, J. R., Toussaint, N. C., Xavier, J. B., and Pamer, E. G. (2015) Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* **517**, 205-U207
136. Solbach, P., Woltemate, S., Chhatwal, P., Tacconelli, E., Buhl, M., Gerhard, M., Vehreschild, M., Jazmati, N., Rupp, J., Manns, M. P., Bachmann, O., and Suerbaum, S. (2018) BaiCD gene abundance is negatively correlated with *Clostridium difficile* infection. *J. Crohns Colitis* **12**, S558-S558
137. Abrini, J., Naveau, H., and Nyns, E. J. (1994) *Clostridium autoethanogenum*, Sp-Nov, an Anaerobic Bacterium That Produces Ethanol from Carbon-Monoxide. *Arch. Microbiol.* **161**, 345-351
138. van Sambeek, D. M., Tran, H., Fernando, S. C., Ciobanu, D. C., Miller, P. S., and Burkey, T. E. (2016) Alteration of the pig intestinal microbiome when vaccinated against or inoculated with porcine circovirus 2 using a multivariate analysis model. *J. Anim. Sci.* **94**, 387-390
139. Ames, G. F. L. (1986) Bacterial Periplasmic Transport-Systems - Structure, Mechanism, and Evolution. *Annu. Rev. Biochem.* **55**, 397-425
140. Rivers, D. M., and Oresnik, I. J. (2015) The Sugar Kinase That Is Necessary for the Catabolism of Rhamnose in *Rhizobium leguminosarum* Directly Interacts with the ABC Transporter Necessary for Rhamnose Transport. *J. Bacteriol.* **197**, 3812-3821
141. Saurin, W., Hofnung, M., and Dassa, E. (1999) Getting in or out: Early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J. Mol. Evol.* **48**, 22-41
142. Wang, B., Dukarevich, M., Sun, E. I., Yen, M. R., and Saier, M. H. (2009) Membrane Porters of ATP-Binding Cassette Transport Systems Are Polyphyletic. *J. Membrane Biol.* **231**, 1-10
143. Torres-Escobar, A., Juarez-Rodriguez, M. D., Lamont, R. J., and Demuth, D. R. (2013) Transcriptional regulation of *Aggregatibacter actinomycetemcomitans*

References

- lslACDBFG and lslRK operons and their role in biofilm formation. *J. Bacteriol.* **195**, 56-65
144. Zhu, J. G., and Pei, D. H. (2008) A LuxP-based fluorescent sensor for bacterial autoinducer II. *Acs Chem. Biol.* **3**, 110-119
145. Duncan, S. H., Aminov, R. I., Scott, K. P., Louis, P., Stanton, T. B., and Flint, H. J. (2006) Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces. *Int. J. Syst. Evol. Microbiol.* **56**, 2437-2441
146. Alou, M. T., Ndongo, S., Fregere, L., Labas, N., Andrieu, C., Richez, M., Couderc, C., Baudoin, J. P., Abrahao, J., Brah, S., Diallo, A., Sokhna, C., Cassir, N., La Scola, B., Cadoret, F., and Raoult, D. (2018) Taxonogenomic description of four new *Clostridium* species isolated from human gut: '*Clostridium amazonitimonense*', '*Clostridium merdae*', '*Clostridium massiliadielmoense*' and '*Clostridium nigeriense*'. *New Microbes New Infect.* **21**, 128-139
147. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501
148. Darwish, N. B., Kochkodan, V., and Hilal, N. (2015) Boron removal from water with fractionized Amberlite IRA743 resin. *Desalination* **370**, 1-6
149. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242
150. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68**, 352-367
151. Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., Bateman, A., and Eddy, S. R. (2015) HMMER web server: 2015 update. *Nucleic Acids Res.* **43**, W30-38
152. Katoh, K., and Standley, D. M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **30**, 772-780
153. Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164-1165

154. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313
155. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457-462
156. Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., and Ostell, J. (2016) NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614-6624
157. Rui, F., Marques, J. C., Miller, S. T., Maycock, C. D., Xavier, K. B., and Ventura, M. R. (2012) Stereochemical diversity of AI-2 analogs modulates quorum sensing in *Vibrio harveyi* and *Escherichia coli*. *Bioorg. Med. Chem.* **20**, 249-256
158. Slater, R. T., Frost, L., Jossi, S., Millard, A. D., and Unnikrishnan, M. (2019) *Clostridioides difficile* LuxS mediates inter-bacterial interactions within biofilms. *Sci. Rep.* **9**, 9903
159. Dapa, T., Leuzzi, R., Ng, Y. K., Baban, S. T., Adamo, R., Kuehne, S. A., Scarselli, M., Minton, N. P., Serruto, D., and Unnikrishnan, M. (2013) Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*. *J. Bacteriol.* **195**, 545-555
160. Torcato, I. M., Kasal, M. R., Brito, P. H., Miller, S. T., and Xavier, K. B. (2019) Identification of novel autoinducer-2 receptors in Clostridia reveals plasticity in the binding site of the LsrB receptor family. *J. Biol. Chem.* **294**, 4450-4463
161. Garner, A. L., Park, J., Zakhari, J. S., Lowery, C. A., Struss, A. K., Sawada, D., Kaufmann, G. F., and Janda, K. D. (2011) A multivalent probe for AI-2 quorum-sensing receptors. *J. Am. Chem. Soc.* **133**, 15934-15937
162. Lambert, C. R., Nijsure, D., Huynh, V., and Wylie, R. G. (2018) Hydrogels with reversible chemical environments for *in vitro* cell culture. *Biomed. Mater.* **13**, 045002
163. Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J., and Uhlen, M. (2005) The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. *Electrophoresis* **26**, 501-510
164. Green, N. M. (1975) Avidin. *Adv. Protein Chem.* **29**, 85-133
165. Zhou, Z., Xiang, Y., Tong, A., and Lu, Y. (2014) Simple and efficient method to purify DNA-protein conjugates and its sensing applications. *Anal. Chem.* **86**, 3869-3875
166. Chen, R., Folarin, N., Ho, V. H., McNally, D., Darling, D., Farzaneh, F., and Slater, N. K. (2010) Affinity recovery of lentivirus by diaminopelargonic acid mediated

References

- desthiobiotin labelling. *J. Chromatogr B Analyt. Technol. Biomed. Life Sci.* **878**, 1939-1945
167. Ngo, C., Mehta, R., Aggarwal, K., Fikes, A. G., Santos, I. C., Greer, S. M., and Que, E. L. (2019) Pull-Down of Metalloproteins in Their Native States by Using Desthiobiotin-Based Probes. *ChemBioChem* **20**, 1003-1007
168. Hirsch, J. D., Eslamizar, L., Filanoski, B. J., Malekzadeh, N., Haugland, R. P., Beechem, J. M., and Haugland, R. P. (2002) Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation. *Anal. Biochem.* **308**, 343-357
169. Quarrell, R., Claridge, T. D., Weaver, G. W., and Lowe, G. (1996) Structure and properties of TentaGel resin beads: implications for combinatorial library chemistry. *Mol. Divers.* **1**, 223-232
170. Richard Santini, M. C. G. a. M. Q. (1998) A Measure of Solvent Effects on Swelling of Resins for Solid Phase Organic Synthesis. *Tetrahedron Lett.* **39**, 8951-8954
171. Rajamani, S., Zhu, J., Pei, D., and Sayre, R. (2007) A LuxP-FRET-based reporter for the detection and quantification of AI-2 bacterial quorum-sensing signal compounds. *Biochemistry* **46**, 3990-3997
172. Ascenso, O. S., Torcato, I. M., Miguel, A. S., Marques, J. C., Xavier, K. B., Ventura, M. R., and Maycock, C. D. (2019) Synthesis and biological activity of a potent optically pure autoinducer-2 quorum sensing agonist. *Bioorg. Chem.* **85**, 75-81
173. Beeton, C., Wulff, H., Singh, S., Botsko, S., Crossley, G., Gutman, G. A., Cahalan, M. D., Pennington, M., and Chandy, K. G. (2003) A novel fluorescent toxin to detect and investigate Kv1.3 channel up-regulation in chronically activated T lymphocytes. *J. Biol. Chem.* **278**, 9928-9937
174. Carter, G. P., Purdy, D., Williams, P., and Minton, N. P. (2005) Quorum sensing in *Clostridium difficile*: analysis of a luxS-type signalling system. *J. Med. Microbiol.* **54**, 119-127
175. Miranda, V. (2019) *Synthesis of new bioactive autoinducer-2 analogues - Approaches to prodrugs, analogues and chemical probes*. Ph.D. thesis, Universidade Nova de Lisboa
176. Collins, K. C., Tsuchikama, K., Lowery, C. A., Zhu, J., and Janda, K. D. (2015) Dissecting AI-2-mediated quorum sensing through C5-analogue synthesis and biochemical analysis. *Tetrahedron* **72**, 3593-3598
177. Gamby, S., Roy, V., Guo, M., Smith, J. A., Wang, J., Stewart, J. E., Wang, X., Bentley, W. E., and Sintim, H. O. (2012) Altering the communication networks of multispecies microbial systems using a diverse toolbox of AI-2 analogues. *Acs Chem. Biol.* **7**, 1023-1030

178. Guo, M., Zheng, Y., Terrell, J. L., Ad, M., Opoku-Temeng, C., Bentley, W. E., and Sintim, H. O. (2015) Geminal dihalogen isosteric replacement in hydrated AI-2 affords potent quorum sensing modulators. *ChemComm.* **51**, 2617-2620
179. Kadirvel, M., Fanimarvasti, F., Forbes, S., McBain, A., Gardiner, J. M., Brown, G. D., and Freeman, S. (2014) Inhibition of quorum sensing and biofilm formation in *Vibrio harveyi* by 4-fluoro-DPD; a novel potent inhibitor of signalling. *ChemComm.* **50**, 5000-5002
180. Lowery, C. A., Park, J., Kaufmann, G. F., and Janda, K. D. (2008) An unexpected switch in the modulation of AI-2-based quorum sensing discovered through synthetic 4,5-dihydroxy-2,3-pentanedione analogues. *J. Am. Chem. Soc.* **130**, 9200-9201
181. Tsuchikama, K., Zhu, J., Lowery, C. A., Kaufmann, G. F., and Janda, K. D. (2012) C4-alkoxy-HPD: a potent class of synthetic modulators surpassing nature in AI-2 quorum sensing. *J. Am. Chem. Soc.* **134**, 13562-13564
182. Semmelhack, M. F., Campagna, S. R., Federle, M. J., and Bassler, B. L. (2005) An expeditious synthesis of DPD and boron binding studies. *Org. Lett.* **7**, 569-572
183. Lien, E. J. (1995) Chirality and drug targeting: pros and cons. *J. Drug Target.* **2**, 527-532
184. Waldeck, B. (1993) Biological significance of the enantiomeric purity of drugs. *Chirality* **5**, 350-355
185. Taga, M. E., and Xavier, K. B. (2011) Methods for analysis of bacterial autoinducer-2 production. *Current protocols in microbiology* **Chapter 1**, Unit1C 1
186. Meijler, M. M., Hom, L. G., Kaufmann, G. F., McKenzie, K. M., Sun, C., Moss, J. A., Matsushita, M., and Janda, K. D. (2004) Synthesis and biological validation of a ubiquitous quorum-sensing molecule. *Angewandte Chemie* **43**, 2106-2108
187. Zhu, J., Miller, M. B., Vance, R. E., Dziejman, M., Bassler, B. L., and Mekalanos, J. J. (2002) Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **99**, 3129-3134
188. Higgins, D. A., Pomianek, M. E., Kraml, C. M., Taylor, R. K., Semmelhack, M. F., and Bassler, B. L. (2007) The major *Vibrio cholerae* autoinducer and its role in virulence factor production. *Nature* **450**, 883-886
189. Sebaugh, J. L. (2011) Guidelines for accurate EC50/IC50 estimation. *Pharm. Stat.* **10**, 128-134
190. Vilchez, R., Lemme, A., Thiel, V., Schulz, S., Sztajer, H., and Wagner-Dobler, I. (2007) Analysing traces of autoinducer-2 requires standardization of the *Vibrio harveyi* bioassay. *Anal. Bioanal. Chem.* **387**, 489-496

References

191. Thiel, V., Vilchez, R., Sztajer, H., Wagner-Dobler, I., and Schulz, S. (2009) Identification, quantification, and determination of the absolute configuration of the bacterial quorum-sensing signal autoinducer-2 by gas chromatography-mass spectrometry. *ChemBioChem*. **10**, 479-485
192. Campagna, S. R., Gooding, J. R., and May, A. L. (2009) Direct quantitation of the quorum sensing signal, autoinducer-2, in clinically relevant samples by liquid chromatography-tandem mass spectrometry. *Anal. Chem*. **81**, 6374-6381
193. Xu, F., Song, X., Cai, P., Sheng, G., and Yu, H. (2017) Quantitative determination of AI-2 quorum-sensing signal of bacteria using high performance liquid chromatography-tandem mass spectrometry. *J. Environ. Sci*. **52**, 204-209
194. Bligh, E. G., and Dyer, W. J. (1959) A rapid method of total lipid extraction and purification. *Can. J. Biochem. and Physiol*. **37**, 911-917
195. Kraus, D., Yang, Q., and Kahn, B. B. (2015) Lipid Extraction from Mouse Feces. *Bio-protocol* **5**
196. Armirotti, A., Basit, A., Realini, N., Caltagirone, C., Bossu, P., Spalletta, G., and Piomelli, D. (2014) Sample preparation and orthogonal chromatography for broad polarity range plasma metabolomics: application to human subjects with neurodegenerative dementia. *Anal. Biochem*. **455**, 48-54
197. Evans, A. M., DeHaven, C. D., Barrett, T., Mitchell, M., and Milgram, E. (2009) Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem*. **81**, 6656-6667
198. Nagana Gowda, G. A., and Raftery, D. (2014) Quantitating metabolites in protein precipitated serum using NMR spectroscopy. *Anal. Chem*. **86**, 5433-5440
199. Hauck, T., Hubner, Y., Bruhlmann, F., and Schwab, W. (2003) Alternative pathway for the formation of 4,5-dihydroxy-2,3-pentanedione, the proposed precursor of 4-hydroxy-5-methyl-3(2H)-furanone as well as autoinducer-2, and its detection as natural constituent of tomato fruit. *Biochim. Biophys. Acta*. **1623**, 109-119
200. Ho, C. S., Lam, C. W., Chan, M. H., Cheung, R. C., Law, L. K., Lit, L. C., Ng, K. F., Suen, M. W., and Tai, H. L. (2003) Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin. Biochem. Rev*. **24**, 3-12
201. Jump, R. L., Polinkovsky, A., Hurless, K., Sitzlar, B., Eckart, K., Tomas, M., Deshpande, A., Nerandzic, M. M., and Donskey, C. J. (2014) Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. *PloS one* **9**, e101267

202. Theriot, C. M., Koenigsnecht, M. J., Carlson, P. E., Jr., Hatton, G. E., Nelson, A. M., Li, B., Huffnagle, G. B., J, Z. L., and Young, V. B. (2014) Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. *Nat. Commun.* **5**, 3114
203. S. Meacham, S. K., A. Wallace and F. Altun. (2010) Boron in Human Health: Evidence for Dietary Recommendations and Public Policies. *The open mineral processing journal* **3**, 36-53
204. Milev, S. (2015) Don't throw away your "bad" N-value ITC data. Malvern Panalytical Materials talks. <https://www.materials-talks.com/blog/2015/04/23/dont-throw-away-your-bad-n-value-itc-data> (accessed August, 30th 2019)
205. Vult von Steyern, F., Josefsson, J. O., and Tagerud, S. (1996) Rhodamine B, a fluorescent probe for acidic organelles in denervated skeletal muscle. *J. Histochem. Cytochem.* **44**, 267-274

ITQB-UNL | Av. da República, 2780-157 Oeiras, Portugal
Tel (+351) 214 469 100 | Fax (+351) 214 411 277

www.itqb.unl.pt