UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR

TESIS DOCTORAL

New strategies for the design and development of protein antimicrobials based on phage products

Nuevas estrategias para el diseño y desarrollo de antimicrobianos proteicos basados en productos fágicos

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Roberto Vázquez Fernández

DIRECTOR

Pedro García González

Centro de Investigaciones Biológicas Margarita Salas
Consejo Superior de Investigaciones Científicas

Madrid, 2021

UNIVERSIDAD COMPLUTENSE MADRID

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña. **Roberto Vázquez Fernández**, estudiante en el Programa de Doctorado **Bioquímica, Biología Molecular y Biomedicina**, de la Facultad de **Ciencias Biológicas** de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

New strategies for the design and development of protein antimicrobials based on phage products
Nuevas estrategias para el diseño y desarrollo de antimicrobianos proteicos basados en productos fágicos

y dirigida por: **Pedro García González**

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 31 de mayo de 2021

Fdo.: 

Firmado digitalmente por VAZQUEZ FERNANDEZ ROBERTO - 71465089H
Fecha: 2021.05.31 12:33:12 +02'00'

φύσις κρύπτεσθαι φιλεί
Heraclitus

Like all great rationalists you believed in things that were twice as incredible as theology
Halldór Laxness, *Under the Glacier*

AGRADECIMIENTOS

Obviamente, debo empezar por Pedro, mi director. Le estoy eternamente agradecido porque, sin su confianza en mí, y sin su debida lucha por la posibilidad material de que este trabajo llegase a buen puerto, no estaría escribiendo estas líneas. A veces se presentan extrañas oportunidades, algunas que no son tan obvias, o que crean dudas. Pero de lo que hoy dudo sinceramente es de que hubiese podido encontrar un compañero mejor, o un director con mejores capacidades para capear mis miedos, obsesiones o mis misteriosos procedimientos. Gracias. En segundo lugar, tengo que dar las gracias a un reducido grupo de gente en el CIB que se abrieron a comprenderme, y que compartieron conmigo fatigas científicas y personales. También compartimos otras cosas: humo, cerveza, ginebra, bailes e incluso viajes. Y aprendí con ellos que no hay manera de hacer buena ciencia (o nada bueno en la vida, en realidad) si no se hace junto a un grupo de colegas alegres y comprensivos. Amigos. Fran (no puedo obviar que todo empezó contigo), y, por supuesto, mis adorados Cris, Natalia, Sergio y Vir. Gracias. Y gracias también a nuestro santo ortodoxo de referencia.

Estoy, por supuesto, en deuda con los que me acogieron en primer lugar: Miri, Susana (la única que aún sigue defendiendo conmigo el 341), Pilar, Tere. Mis primeros profesores en el laboratorio: Alba y Blas. Noemí y Manu, del 'Roca', y, después, Noelia. A pesar de lo que pudiera parecer, también tengo un agradecimiento sincero para todos los Rubenes. Sin importar si tenemos afinidad personal o no, aún mantenemos el caballeroso honor de regalar nuestra ayuda científica allá donde se pide, sin esperar nada a cambio. Enhorabuena a todos nosotros. También doy las gracias a "mis estudiantes", porque sé que no siempre soy fácil de llevar. Pero tenéis que admitir que soy un buen profesor. Fran, Víctor, Sofía. Evidentemente, no puedo olvidar a Ernesto: escucharle es casi como leer un poema de Machado, pero en prosa y sin rodeos con la (incómoda) verdad. Me lo pasé fenomenal en su fiesta de jubilación. Con Marga, que tiene la capacidad de hablar sin que se perciba el final de su discurso, lo mismo que Pedro, disfruté de un inolvidable viaje a Nueva York (vaya tres compañeros de piso), y también de su consejo científico. Gracias. También a Luis Rivas, por hacer una entrada triunfal en el último minuto con sus incisivos comentarios de "*referee* #2". Por supuesto, también he tenido el placer de beneficiarme de la experiencia y de la atención de muchos de los científicos "senior" de los Rubenes: Auxi, Eduardo, Manolo, Jose, Bea, Jesús, Beatriz. Gracias. Por motivos históricos, creo que también debo darle las gracias a Juan Nogales. Parece que fue el primero en apreciar mi trabajo, aunque en un contexto completamente diferente al de mi paso por el CIB.

Gracias al aliento de Pedro, tuve la oportunidad de pasar una temporada en el laboratorio de Yves Briers, en Gante. Una ciudad preciosa para una experiencia muy enriquecedora. Aprovecho la oportunidad para disculparme públicamente con Yves por "despreciar" las patatas fritas belgas en comparación con mi propia receta. En cualquier caso, le agradezco la oportunidad de trabajar en su laboratorio, y de aprender muchas cosas sobre ciencia, sobre mí mismo y sobre cerveza. Saludaré también a mis amigos belgas, en especial a Isabelle, Dennis, Hans, Bjorn, Lenny y Diana, mi guía española en Bélgica. Gracias. También parece el momento apropiado para dar las gracias a quienes financiaron mi viaje: EMBO y CIBERES.

Me temo que esto se alarga de más, pero todavía quedan cuentas por saldar. Por ejemplo, con nuestros colaboradores del ICTP, que hicieron posible un final feliz para la aventura de la encapsulación, contra todo pronóstico. Curra, Julio, Javier, gracias. Mark van Raaij y Mateo Seoane fueron indispensables para explorar el fascinante mundo de las estructuras de proteínas. Nuestro núcleo CIBERES en el Centro Nacional de Microbiología también debe ser reconocido. Especialmente Jose, que en el medio de una pandemia mundial tuvo tiempo para retomar los ratones para nosotros. Me encantó su labrador cuando nos encontramos fortuitamente en Rueda. Guillermo, del Servicio de Bioinformática del CIB fue fundamental para desatar todo el potencial de nuestro "pequeño" proyecto informático, que tardó en conceptualizarse y ejecutarse prácticamente cinco años. También debo agradecer a Carmen Doñoro su trabajo con los cultivos celulares y su buena disposición para ayudarnos. Gracias también a Sebastián Albertí y su laboratorio por atreverse a juguetear con nuestros compuestos. Y, por ir acabando, también tengo un recuerdo especial de las reuniones presenciales de la red FAGOMA, y de las complutenses Cristina Casals, Olga Cañadas y Belén Patiño. También tengo que nombrar a los donadores de cepas: Carmen Ardanuy, Sara Martí, Rosa del Campo, Juan Miguel Rodríguez.

Por supuesto que tengo palabras para mis amigos eternos, los que caminan conmigo el sendero que empieza en Santa María del Páramo (Reino de León). A Alex, Leti (y a Fox), Marta, Eva, Isa, Franco, Alba, Alicia, Sandra (e incluso puede que nombre a Victoria). Porque sin ninguno de vosotros, estoy convencido de que no habría sido capaz de hacer nada de esto. Mis más sinceras gracias. Mi vida en Madrid no habría sido igual sin aquellos que gustamos de llamar "los consortes". Así que también saludo a Gonzalo, Héctor, Julián, y a Víctor.

Es una obviedad que tengo muchas cosas que agradecer a mis padres y demás parientes. Siempre termino por mitificar dos elementos fundacionales en mi vida: los cuentos que mi padre me contaba antes de dormir y la desvergonzada fuerza de mi madre al enfrentarse a todos los golpes y desafíos de la vida. Mi forma de vestir también es culpa

de mi madre, y pedimos sinceras disculpas por ello. En cualquier caso, gracias a ambos por no dejar que me arrastrara la corriente. Este es también el momento apropiado para dar las gracias a mi tía María Elena, y para recordar a mis abuelos. Ellos tampoco se conformaron. Amadeo, Carmen, Ángel, Nieves. Sé que estaríais orgullosos de nosotros.
Quizás también ¿agradezca? Al SARS-CoV-2 por darme tiempo para reflexionar pausadamente, y para elaborar la parte obviamente teletrabajada de esta tesis. También para reconocer mis propios límites. Aunque habría preferido, sinceramente, que nunca hubiera existido. Mi tío y mi tía no lo merecían, y siempre los llevo en la memoria. Hasta siempre, Poldy y Miguel, y gracias por todo. Gracias por abrirle las puertas de Madrid a un jovencísimo e ingenuo provinciano que solo empezaba a chapotear en el barro de la gran ciudad.
Ya casi termino, lo prometo, pero siento que debo hacer un pequeño ejercicio de *name-dropping* para reconocer las lecturas que me han dado forma a mí mismo en estos últimos años: Heráclito, Freud, Marx, Wittgenstein, Foucault, Marzoa, Žižek y, por supuesto, a mis profesores de Filosofía de la UNED. Gracias por ayudarme a descubrir que hay muchísimo más que ciencia pura y dura en la vida académica.
Finalmente, te doy las gracias a ti, por llegar al final de esta incontinencia verbosa. Porque, al contrario de lo que pudiera parecer, no soy tímido ni callado. Sencillamente me tomo mi tiempo para decidir si lo que voy a decir es o no bienvenido, o si me importa que lo sea. Y como esta es mi tesis, *I cry if I want to*.

TABLE OF CONTENTS

ABBREVIATIONS

Besides the abbreviations and units accepted by the International System of Units (SI, https://physics.nist.gov/cuu/Units/units.html) and the International Union of Pure and Applied Chemistry (IUPAC, https://iupac.org/nomenclature-and-terminology), the following abbreviations have been used throughout this thesis:

aa	Amino acid/s
AMP	Antimicrobial peptide
AMR	Antimicrobial resistance
AP	Ampicillin
AUC	Area under the curve
AZI	Azithromycin
BHI	Brain heart infusion
C+Y	C medium pH 8.0 with 0.08% yeast extract
CAP	Community-acquired pneumonia
CBD	Choline binding domain
CBP	Choline binding protein
CBR	Choline binding repeat
CHL	Chloramphenicol
CID	Collision induced dissociation
CLSI	Clinical and Laboratory Standards Institute
CLSM	Confocal laser scanning microscopy
CV	Crystal violet
CWBD	Cell wall binding domain
DEAE	Diethylaminoethanol/diethylaminoethyl
DLS	Dynamic light scattering
EAD	Enzymatically active domain
ECDC	European Centre for Disease Prevention and Control
ERY	Erythromycin
FIC	Fractional inhibitory concentration
FICI	Fractional inhibitory concentration index
FPR	False positive rate
G−	Gram-negative
G+	Gram-positive
GAM	Generalized additive model
GAS	Group A streptococci, *Streptococcus pyogenes*
GBS	Group B streptococci, *Streptococcus agalactiae*
GEN	Gentamycin
GEWL	Goose egg white lysozyme
HAP	Hospital-acquired pneumonia
HCAP	Healthcare-associated pneumonia
HEWL	Hen egg white lysozyme
HM	Hydrophobic moment

HMM	Hidden Markov Model
ICU	Intensive care unit
IPTG	Isopropyl-β-D-thiogalactopyranoside
KAN	Kanamycin
*k*NN	*k*-Nearest Neighbours
LB	Lysogeny broth
LPS	Lipopolysaccharide
LRTIs	Lower respiratory tract infections
LVX	Levofloxacin
MDR	Multi-drug resistant
MIC	Minimum inhibitory concentration
MRSA	Methicillin-resistant *Staphylococcus aureus*
MSA	Multiple sequence alignment
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
MurNAc	N-acetylmuramic acid
NAG	N-acetylglucosamine
NAM-amidase	N-acetylmuramoyl-L-alanine amidase
NaPiB	Sodium phosphate buffer
NCBI	National Center for Biotechnology Information
NCPR	Net charge per residue
NHS	N-hydroxysuccinimidyl
NP	Nanoparticle
NPN	N-phenyl-1-naphthylamine
OD	Optical density
OM	Outer membrane
PBS	Phosphate buffer saline
PDI	Polydispersity index
PDR	Pandrug resistant
PF	Pfam
PI	Propidium iodide
RBB	Remazol brilliant blue
ROC	Receiver operating characteristic
RP-HPLC	Reverse-phase high-performance liquid chromatography
SAR	Signal-arrest-release
SDS-PAGE	Sodium dodecylsulphate-polyacrylamide gel electrophoresis
SMG	*Streptococcus* milleri group
SSN	Sequence similarity network
TFE	2,2,2-trifluoroethanol
TPP	Tripolyphosphate
TPR	True positive rate
TSB	Trypticasein soy broth
VAP	Ventilator-associated pneumonia
WHO	World Health Organization
XDR	Extensively drug resistant

SUMMARY

Nowadays, antibiotic resistance is one of the most urgent global health issues to be tackled. Among the infective bacteria, Gram-negative ones are the main causative agents of nosocomial infections, causing an important death burden, partly due to antimicrobial resistance. Currently, phage therapy research is experiencing a renaissance as a viable alternative or complement to common antimicrobial chemotherapy. Not only complete virion particles but also phage-derived products have a series of advantages over antibiotics. In particular, phage (endo)lysins are the phage-encoded proteins responsible for bacterial host lysis and subsequent death due to their cell wall lytic activity. Lysins are being repurposed to be exogenously applied, in a purified form, against bacteria, functioning as lytic antimicrobials, and thus are also called 'enzybiotics'. Among the many advantages of enzybiotics, we may cite their lower chance to provoke resistance in the target bacteria and their versatility to be screened and engineered using bioinformatic and biotechnological tools. The current literature already proves that enzybiotics are efficient in killing Gram-positive bacteria. However, they have been deemed less effective against Gram-negative ones due to the presence of the outer membrane, which acts as a permeability barrier that prevents the access of the exogenously added enzyme to its target (the bacterial peptidoglycan). However, not only some recent engineering efforts have rendered lysins active against Gram-negative 'from without', but also there are reports of some lysins from phages that infect Gram-negatives that are intrinsically active against such bacteria. Such an intrinsic activity has been related to antimicrobial peptide-like regions located within the lysins. However, the evolutionary importance of those regions or their actual spread among the Gram-negative phage lysins is still unknown.

In this thesis, a full design and development path towards the obtention of phage lysin-based antimicrobials especially directed against Gram-negative pathogens was investigated. Firstly, a comprehensive database of lysin sequences was compiled and analysed to detect structural features related to the architecture of the corresponding bacterial hosts. This analysis supported a widespread appearance of antimicrobial peptide-like subdomains within a relevant subpopulation of lysins from Gram-negative bacteria infecting phages. Making use of this database, a bioinformatic screening method was set up to predict antimicrobial peptide-like regions within the lysin sequences to select a set of enzybiotic candidates with the potential to interact (and disrupt) the Gram-negative outer membrane. In this way, an enzybiotic candidate, named Pae87, was selected for researching its intrinsic activity against Gram-negative pathogens, particularly *Pseudomonas aeruginosa*, and to serve as a scaffold for the developing of efficient antimicrobial

molecules. A three-dimensional model of the protein with a bound peptidoglycan fragment was obtained in the course of this research. In this way, a putative substrate-binding subdomain was identified within the very catalytic domain of the protein. In addition, the ability of Pae87 to bind and disrupt the Gram-negative outer membrane was proven, and it was related to a specific C-terminal region termed peptide P87. This peptide had an antimicrobial activity of its own, comparable to that of the full protein. Moreover, peptide P87 was enhanced by rational point mutation and thus the derived peptide P88 was obtained. Peptide P88 had a greater bactericidal efficiency against a similar range of Gram-negative bacteria but no dramatic increase in its cytotoxic effect against eukaryotic cells was found. In order to improve the therapeutic potential of P88, its synergistic activity with a set of antibiotics was tested. The peptide displayed a potent *in vitro* synergy when used in combination with various antibiotics with intracellular targets (namely macrolides, chloramphenicol, and tetracycline). Using this strategy, the effective concentration of peptide (and antibiotics) was 4-fold reduced. Additionally, Pae87 activity was enhanced by fusing it to a calcium-binding domain, giving rise to the Pae87F chimera. The calcium-binding activity was expected to improve the outer membrane-disrupting ability of the enzyme or its attachment to it by binding the cations that usually stabilize the outer membrane. In this way, the use of this kind of domain as enhancing modules for engineering enzybiotics was for the first time demonstrated.

Finally, a strategy for the *in vivo* enzybiotics behaviour enhancement was developed. Since lysins *in vivo* half-life is known to be rather short (30-60 min), one of the currently explored approaches for enzybiotics administration is encapsulation and controlled release. Based on the *Streptococcus pneumoniae* surface structure, which stands out for the presence of choline residues that serve as anchorage for physiologically relevant surface proteins, a chitosan-based polymer was designed. Such a polymer was grafted with diethylaminoethanol (DEAE) moieties which can act as a structural and functional analogue of choline. Therefore, the nanoparticles prepared with chitosan-DEAE were able to specifically bind choline-binding proteins. Due to this ability, the treatment of pneumococcal cultures with chitosan-DEAE nanoparticles provoked the chaining of the bacterial cells because of the competitive inhibition of housekeeping choline-binding proteins responsible for daughter cell separation. The chitosan-DEAE nanoparticles were thus also able to bind choline-binding enzybiotics, such as the antipneumococcal Cpl-711, and release it in a controlled manner, although with some cytotoxic effect against eukaryotic cells.

Altogether, the results presented in this thesis investigate state-of-the-art tools and methods for the development of alternative, phage-derived antimicrobials to tackle some of the most relevant public health issues that we have to face in the years to come.

RESUMEN

Hoy en día, la resistencia a antibióticos es uno de los problemas de salud mundial más urgentes que deben abordarse. Entre las bacterias infecciosas, las Gram-negativas son los principales agentes causantes de infecciones nosocomiales, provocando una importante mortalidad, en parte debido a la resistencia a antimicrobianos. Actualmente, la investigación en terapia fágica está experimentando un renacimiento, entendida como alternativa viable o complemento de la quimioterapia antimicrobiana común. No solo las partículas fágicas completas sino también los productos derivados de fagos tienen una serie de ventajas sobre los antibióticos. En particular, las (endo)lisinas fágicas son las proteínas codificadas por los fagos que son responsables de la lisis y muerte bacteriana debido a su actividad enzimática que hidroliza la pared celular. Las lisinas se están re-utilizando para aplicarlas exógenamente en forma purificada contra bacterias, funcionando como antimicrobianos líticos y, por ello, también se denominan "enzibióticos". Entre las muchas ventajas de los enzibióticos, podemos citar su menor probabilidad de provocar resistencias en las bacterias diana y su versatilidad para ser detectadas y diseñadas utilizando herramientas bioinformáticas y biotecnológicas. La literatura actual prueba que los enzibióticos son altamente eficaces contra bacterias Gram-positivas. Sin embargo, se han considerado menos efectivos contra las Gram-negativas, debido a la presencia de la membrana externa que actúa como una barrera de permeabilidad que impide el acceso de la enzima exógenamente añadida a su diana (el peptidoglicano bacteriano). Sin embargo, no solo existen recientes estrategias de ingeniería de proteínas que han producido lisinas activas contra Gram-negativas "desde fuera", sino que, además, hay ejemplos en la literatura de algunas lisinas de fagos que infectan a Gram-negativas que son intrínsecamente activas contra dichas bacterias. Tal actividad intrínseca se ha relacionado con regiones similares a péptidos antimicrobianos presentes en las propias lisinas. Sin embargo, aún se desconoce la importancia evolutiva de esas regiones o su propagación real entre las lisinas de fagos que infectan a bacterias Gram-negativas.

En esta tesis se ha llevado a cabo una estrategia completa de diseño y desarrollo de antimicrobianos basados en lisinas fágicas, especialmente dirigidos contra patógenos Gram-negativos. En primer lugar, se compiló y analizó una base de datos exhaustiva de secuencias de lisinas para detectar características estructurales relacionadas con la arquitectura de los hospedadores bacterianos correspondientes. En este análisis se observó la presencia generalizada de subdominios similares a péptidos antimicrobianos dentro de una subpoblación relevante de lisinas de fagos de Gram-negativas. Haciendo uso de esta base de datos, se estableció un método de cribado bioinformático para

predecir regiones similares a péptidos antimicrobianos dentro de las secuencias de lisinas con el objetivo de obtener un conjunto de candidatos a enzibióticos con el potencial de interactuar (y desestructurar) la membrana externa Gram-negativa. De esta manera, un enzibiótico candidato, denominado Pae87, fue seleccionado para investigar su actividad intrínseca contra patógenos Gram-negativos, particularmente *Pseudomonas aeruginosa*, y para servir como base para el desarrollo de moléculas antimicrobianas eficientes. En el curso de la investigación, se obtuvo un modelo tridimensional de esta proteína conteniendo un fragmento de peptidoglicano unido. De esta manera, se identificó un posible subdominio de unión a sustrato dentro del propio dominio catalítico de la proteína. Además, se demostró la capacidad de Pae87 para unirse y permeabilizar la membrana externa Gram-negativa, y esta actividad se relacionó con una región C-terminal específica denominada péptido P87. Este péptido mostró actividad antimicrobiana propia, comparable a la de la proteína completa. Además, el péptido P87 se potenció mediante mutaciones puntuales racionales y como resultado se obtuvo el péptido derivado P88. El péptido P88 mostró una mayor eficacia bactericida contra un rango similar de bacterias Gram-negativas, pero sin un aumento importante en su efecto citotóxico contra células eucariotas. Para mejorar el potencial terapéutico de P88, se probó su actividad sinérgica con un conjunto de antibióticos. El péptido mostró una potente sinergia *in vitro* cuando se usó en combinación con varios antibióticos con dianas intracelulares (a saber, macrólidos, cloranfenicol y tetraciclina). Usando esta estrategia, la concentración efectiva de péptidos (y antibióticos) se redujo 4 veces. La actividad de Pae87 también se mejoró, en este caso, fusionándola con un dominio de unión a calcio, dando lugar a la quimera Pae87F, esperando que la actividad de unión al calcio mejorara la capacidad de alteración de la membrana externa de la enzima o bien su unión a ella mediante la unión de los cationes que normalmente estabilizan la membrana externa. De esta forma, se ha demostrado por primera vez el posible uso de este tipo de dominios como módulos para mejorar enzibióticos mediante ingeniería de proteínas.

Finalmente, se desarrolló una estrategia para mejorar el comportamiento *in vivo* de enzibióticos. Dado que se sabe que la vida media *in vivo* de las lisinas es bastante corta (30-60 min), uno de los enfoques actualmente explorados para la administración de enzibióticos es su encapsulación y liberación controlada. Basándonos en la estructura superficial de *Streptococcus pneumoniae*, que destaca por la presencia de residuos de colina que sirven de anclaje para proteínas superficiales fisiológicamente relevantes, se diseñó un polímero a base de quitosano. Dicho polímero se derivatizó con dietilaminoetanol (DEAE) que pueden actuar como un análogo estructural y funcional de la colina. Por ello, las nanopartículas preparadas con quitosano-DEAE resultaron ser capaces de unir específicamente proteínas de unión a colina. Debido a esta capacidad, el

tratamiento de cultivos neumocócicos con nanopartículas de quitosano-DEAE provocó el encadenamiento de las células bacterianas debido a la inhibición competitiva de las proteínas de unión a colina responsables de la separación de las células hijas. Las nanopartículas de quitosano-DEAE también fueron capaces de unir enzibióticos de unión a colina, como el antineumocócico Cpl-711, y liberarlo de forma controlada, aunque con cierto efecto citotóxico en las células eucariotas.

En conjunto, los resultados presentados en esta tesis proporcionan herramientas y métodos de última generación para el desarrollo de antimicrobianos alternativos derivados de fagos para abordar algunos de los problemas de salud pública más relevantes que tendremos que enfrentar en los próximos años.

I. INTRODUCTION

1. BACTERIAL RESPIRATORY INFECTIONS AND ANTIMICROBIAL RESISTANCE

It is almost unnecessary these times to remark that communicable diseases are still a key player in the global health landscape. The current pandemic has taught us at least not to underestimate the threat of infectious diseases, for even though scientific actors have been long time warning about the perils of gathering epidemics, the COVID-19 has taken over and completely changed our lifestyles. Even before the pandemic, communicable diseases still represented an important percentage of the global death burden (**Figure 1.1A**). In 2016 estimates, lower respiratory tract infections (LRTIs) remained the leading infectious cause of death, with more than 2 million deaths per year (1). LRTIs are even more relevant in low-income countries, where they represented the second cause of death by burden (2). Community-acquired pneumonia (CAP) was, in fact, the single largest global bacterial infectious cause of death in children (3). *Streptococcus pneumoniae* accounted for most of CAP cases, followed by some other bacterial and viral pathogens such as respiratory syncytial virus, *Haemophilus influenzae* type b, influenza virus, *Streptococcus pyogenes* (Group A streptococci, GAS), non-typeable *H. influenzae*, *Staphylococcus aureus*, *Mycoplasma pneumoniae*, *Moraxella catarrhalis*, and *Klebsiella pneumoniae* (1, 4, 5) (**Figure 1.1B**). While COVID-19 has changed the global health game rules —at least for the time being— there are already some reports suggesting that respiratory bacterial co-infections can be a critical risk factor regarding severity and mortality in COVID-19 patients, especially in those admitted to intensive care units (ICUs) (6, 7). Nosocomial respiratory infections have always been a major concern regarding hospital-admitted patients because they present etiological and therapeutic challenges of their own (8). Indeed, hospital-acquired pneumonia (HAP) and ventilator-associated pneumonia (VAP) are the leading nosocomial infections worldwide, and, among their most frequent causative agents, we may cite many Gram-negative (G−) bacteria (*Pseudomonas aeruginosa*, *Klebsiella* sp., *Escherichia coli*, *Acinetobacter* sp. or *Enterobacter* sp.; and, with lower prevalence, *Serratia* sp. or *Stenotrophomonas maltophilia*). In the clinical setting, the most common Gram-positive (G+) causative agent is, indisputably, *S. aureus* and, in some cases, community-acquired G+ bacteria such as *S. pneumoniae* (**Figure 1.1C**) (9, 10). While HAP, VAP, and Healthcare-associated pneumonia (HCAP) represent, in percentage, less than CAP hospitalizations (**Figure 1.1D**), nosocomial pneumonia cases have a death rate that is almost four times that of CAP (11) (**Figure 1.1E**). And, getting to the root of the problem, this higher mortality is typically associated with high inherent antibiotic resistance rates especially among the G− bacteria causing HAP and VAP (10, 11).

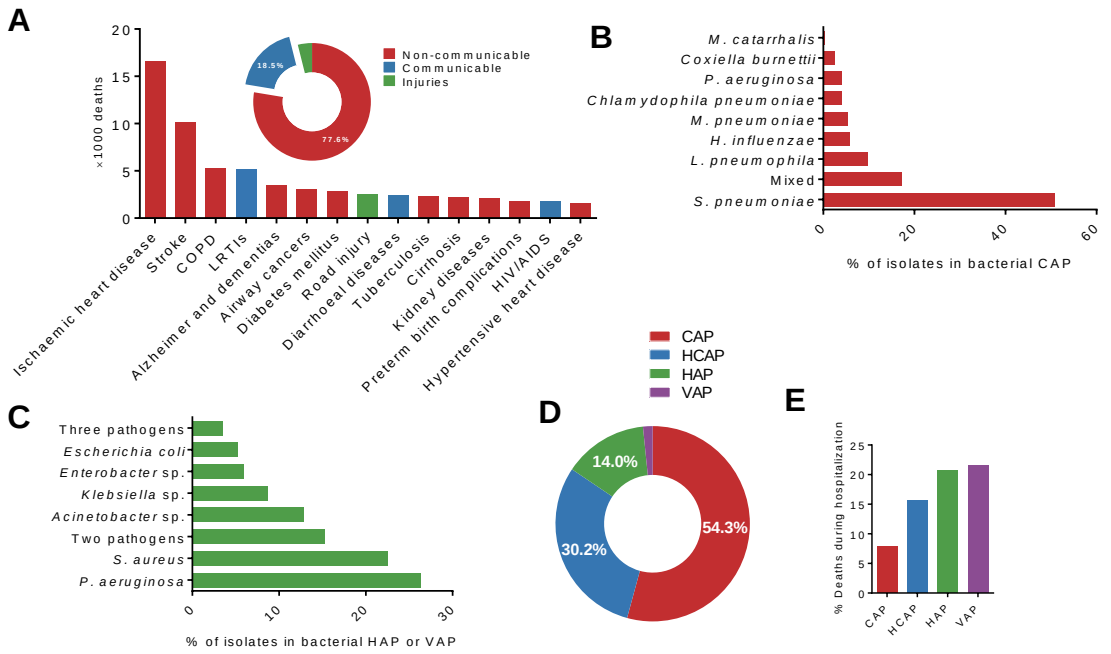


Figure 1.1. Impact and causative agents of pneumonia. A: Top 15 global death causes in 2016 (2). **B:** Causative agents of CAP (12). **C**: Causative agents of HAP or VAP (13). **D**: Percentage of hospitalizations due to CAP, HCAP, HAP, and VAP among total hospital stays due to pneumonia (11). **E**: Percentages of death during hospitalization due to CAP, HCAP, HAP, and VAP (11).

The introduction of antibiotics (1930-1940s) as antibacterial therapeutics changed forever the way we handle bacterial infections. However, since antibiotics are just a part of the complex ecological system in which bacteria are included, microorganisms have also evolved mechanisms to evade their action. This is why the discovery of antibiotic resistance (AMR) mechanisms was almost simultaneous to the beginning of their use as antimicrobials (14, 15). The extensive use of antibiotics as chemotherapeutic agents has thus provoked an additional evolutionary pressure that has historically led to the selection of resistant pathogenic strains. However, a good discovery pace of novel antibiotics allowed circumventing such problem in the clinic until rather recently. Nowadays, given the widespread use of antibiotics, AMR has become a serious health threat that has been called to be tackled by many public actors (16-20). In an already classically cited report, the Review on Antimicrobial Resistance estimated that, by 2050, the world would have as much as 10 million deaths due to AMR if the trends should continue, more than the current 8.3 million yearly deaths due to cancer (19). Although such estimations have a clear aim to stir the public opinion from a scientific point of view (and should be taken with care (21)) even such entities as the World Bank are already warning about the economic perils of the AMR threat. They estimate that the world will suffer an annual shortfall loss between $1 and $3.4 trillion by 2013 due to AMR (22).

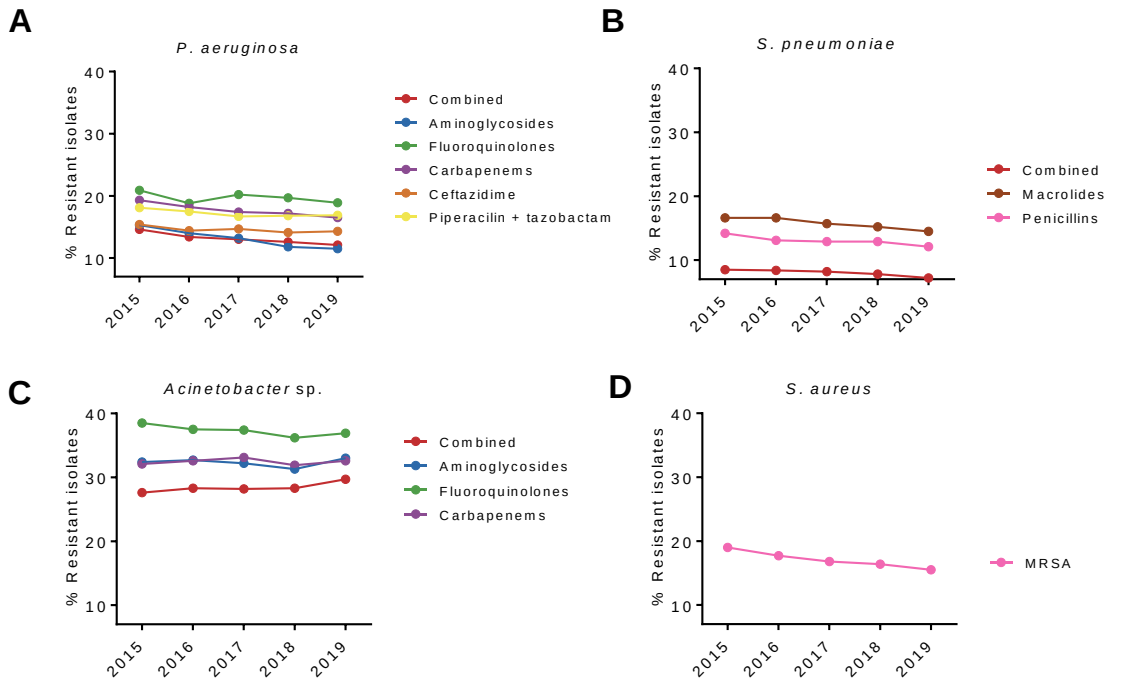


Figure 1.2. Evolution of antibiotic resistance in Europe in selected pathogens. Data adapted from European Centre for Disease Prevention and Control (ECDC) report on AMR (16).

Looking at the AMR surveillance data from the ECDC, if a trend should be underlined, is that of a generally slow but steady decay (or, at least, a generalized non-increasing tendency) (**Figure 1.2**). However, it is difficult to assess the relevance of such a general trend, given the wide variety that is also observed between regions and pathogens (16). For example, while the methicillin-resistant *S. aureus* (MRSA) prevalence or that of aminoglycosides and carbapenem resistance in *P. aeruginosa* seem to decline over the 2015–2019 period (**Figure 1.2AD**), the tendency is clearly reversed for the multi-drug resistant (MDR) *Acinetobacter* sp. (**Figure 1.2C**). On the other hand, the two bacterial pathogens responsible for most of the CAPs and HAP/VAPs, *P. aeruginosa*, and *S. pneumoniae*, show somewhat stable proportions of antibiotic resistance, in the range of 10–20% (**Figure 1.2AB**). These percentages put them into the World Health Organization (WHO) priority list of antibiotic-resistant bacteria for the development of novel antimicrobials (18), together with other relevant pathogens such as *S. aureus*, *Acinetobacter baumannii* or other members of the so-called ESKAPE bugs (*Enterococcus faecium*, *S. aureus*, *K. pneumoniae*, *A. baumannii*, *P. aeruginosa, Enterobacter* sp.).

As for the specific situation in Spain, a recent study yielded that, among *P. aeruginosa* clinical isolates, 17.3% were extensively drug-resistant (XDR), 8.8% MDR and non-XDR,

and 0.1% pandrug resistant (PDR)[1] (24). Altogether, these numbers mean that Spain is at the European average percentage of resistance for *P. aeruginosa*, as confirmed by the ECDC latest report (16). For comparison, the lowest carbapenem-resistant *P. aeruginosa* percentage recorded in Europe is below 1% (Iceland) and the highest one is above 50% (Romania) (**Figure 1.3A**).

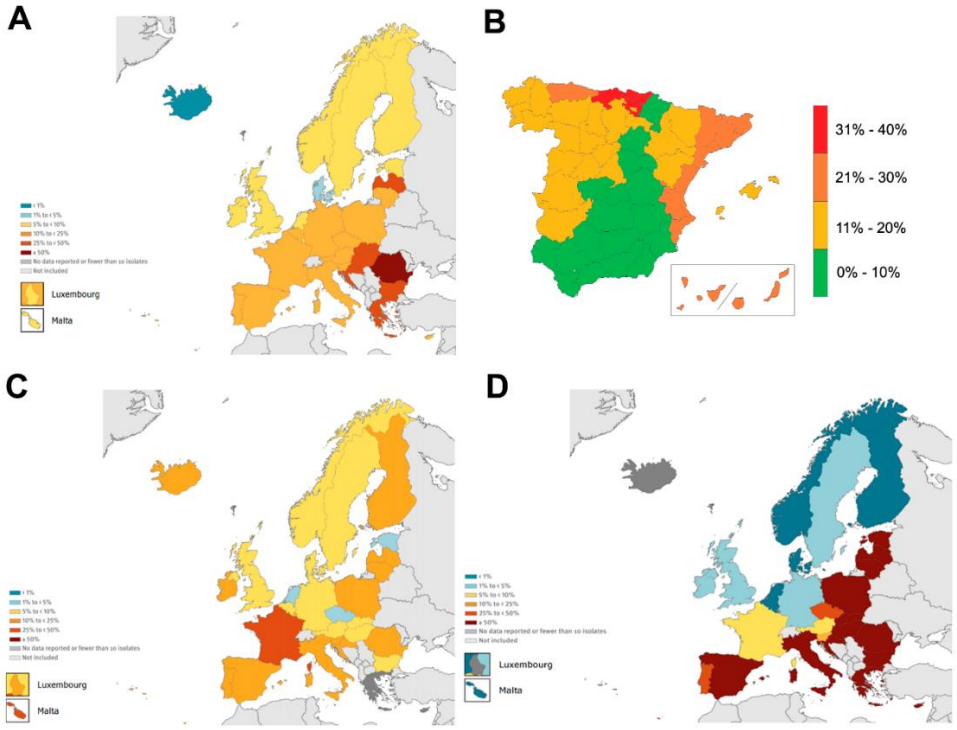


Figure 1.3. Regional distribution of antibiotic-resistant bacteria. **A**: Percentage of carbapenem-resistant *P. aeruginosa* isolates by country in the EU/EEA (16). **B**: Distribution of *P. aeruginosa* XDR phenotypes in Spanish regions, adapted from (24). **C**: Percentage of *S. pneumoniae* penicillin-resistant isolates (16). **D**: Percentage of carbapenem-resistant *Acinetobacter* sp. isolates (16).

Another interesting point of such study is that the percentage of resistant isolates is much higher in samples from ICU patients than in those from other hospital wards, as it would be expected, thus further complicating the management of already problematic patients (24). Another trend that is replicated in Spain is the wide inter-regional variation of the resistance prevalence (**Figure 1.3B**), which implies that measures to counter AMR must be guided by local epidemiological data, rather than by general trends. Data from a Spanish hospital also underlines the negative impact of AMR on patient's outcome,

[1] The common definition of MDR is "non-susceptibility to at least one agent in at least three antibiotic classes", while XDR profile is "non-susceptibility to at least one agent in all but one or two antibiotic classes". A related concept, PDR indicates that the strain is resistant to all of the antimicrobial agents considered (23).

rendering that *P. aeruginosa* infections by an MDR strain increase long-term mortality risk by more than 30% (25). While among G+ pathogens the classical main source of concern is MRSA, *S. pneumoniae* still poses a major epidemiological challenge due to a relatively high incidence among risk groups (20-30% in persons aged 65 or more by 2015 (26)) and to serotype replacement. The latter phenomenon is due to the nature of the currently used pneumococcal vaccines, which are made up of capsular polysaccharides from a series of prevalent serotypes. The fact that only a certain set of the more than 98 pneumococcal serotypes are included in these vaccines leads to the evolutionary selection of non-vaccine serotypes, that tend to epidemiologically replace those against which vaccines provide protection (27). Adding to those issues, there is a relevant percentage of penicillin or macrolide-resistant strains which, in the case of Spain, is in the range of 10–25% (**Figure 1.3C**). Although the more recent trend is towards a slow decrease in the resistance prevalence, a wide historical series shows a general increase (from the ≈10% resistance to macrolides in the early 1990s up to the ≈30% in the 2000s) that must be surveilled and tackled (28). This is even more so if we consider the variability in resistant phenotypes among serotypes and the phenomenon of serotype replacement (29, 30). Anyway, by all means, the higher AMR epidemiological representative in Spain is *Acinetobacter* sp., with a current carbapenem resistance percentage >50% (**Figure 1.3D**). The time series from the late 1990s to nowadays is also worrying, with some hospitals reporting up to 80-90% of strains resistant to a somewhat wide range of antibiotics (31, 32). This reveals an alarming situation that also merits close monitoring.

Higher AMR rates have been associated with the misuse and abuse of antibiotics both at an individual and population level (33, 34). To worsen the problem, the new antibiotics pipeline has dried out since the 1980s (35), rendering that there are no new drugs to make up for the ones to which bacteria are becoming resistant. Therefore, antimicrobial stewardship[2] programs have been proposed and implemented to try and control antibiotic use and lengthen the effective life of currently available drugs (36). It might be a temptation to link the apparent decrease in resistance among some bacterial populations to a rationalization of antibiotics use. However, the highly heterogeneous implementation of such interventions, together with the great difficulties of their evaluation (37), discourages considering that the current measures are working. The enormous variability in AMR management and prevalence depending on bacterial species, antimicrobial group, and geographical region should also prevent any overly optimistic conclusion (16). For

[2] The term “stewardship” has proven to be rather difficult to translate, given the lack of equivalent terms in many languages. So, antibiotic stewardship programs have been labelled as “*bon usage des antibiotiques*” in French, “*rationaler Einsatz von Antibiotika*” in German or “programas de optimización de uso de antimicrobianos (PROA)” in Spanish.

example, as is the case for many other issues, low and middle-income countries have a disproportionately large burden due to AMR in comparison with high-income regions (19). In this case, multipronged approaches are especially needed. Too much focus in the healthcare sector, as it is usual for AMR management particularly in high-income countries, would require an important investment and a strong infrastructure that is lacking in poor economies (38). Particularly, improving sanitation is expected to have a very high impact in low to middle-income countries (19). On the other hand, weak governance, impervious social beliefs or low public awareness make technological approaches (introduction of novel antimicrobials, etc.) more viable as a "quick-fix" in these settings, although it would obviously require adequate funding (38).

Whichever the case, a general conclusion is that, since the surveyed bacterial populations thus far contain indisputably high AMR proportions, effective interventions should be enforced sooner rather than later. Or, as it has been said, being AMR a public health issue, in which individual actions affect broader populations, it "demands moral solutions because it represents situations in which each of us acting alone can produce a worse outcome than all of us working together" (39). Indeed, natural history is currently showing us the potential risks of ignoring the evolutionary forces operating on pathogenic microorganisms, and their close relationship to human action. In the meantime, academic, preclinical, and clinical pipelines are already in place to deliver antimicrobial alternatives as soon as possible.

2. RISE, FALL, AND RENAISSANCE OF PHAGE THERAPY

One of the oldest alternatives to antibiotics is phage therapy: using bacteriophages —viruses that infect bacteria— to control harming bacterial populations. The first phage-related observation is considered to be that of a naturally occurring 'phage therapy', since Ernest Hankin pointed out the bactericidal ability of river water against *Vibrio cholerae* in 1896, although whether such reference can be considered a phage observation or not is currently disputed (40). About the same time (1898), the Russian bacteriologist Nikolái Gamaleya observed a 'transmissible lytic agent' specifically active against *Bacillus subtilis* (41). Nonetheless, the formal discovery of phages, at least in the West, is considered to be related to the independent observations of Frederick Twort (1915) and Felix d'Herelle (1917) (42-44). While it is generally accepted that Twort was the first to suggest that a virus was indeed the agent that killed bacteria, it was d'Herelle who coined the term 'bacteriophage' (from the Greek, literally meaning 'bacteria eater'). He was also the first to propose using such infectious particles to treat bacterial infections, starting in 1919 by having dysentery patients ingest bacteriophage solutions, although he did not publish

his results until 1931 (45). Although many early studies following d'Herelle's contribution addressed the potential use of phage as therapeutics (46), since then, the history of phage therapy became clearly bifurcated. With the discovery of antibiotics (Fleming's seminal paper on penicillin is dated in 1929) (47) and their extensive therapeutic introduction (which began around 1942), the Western world disengaged from phage therapy research and commercial production of phages (**Figure 1.4**), which had already started both in Western Europe and the US around 1940. The reasons for this disengagement were probably an apparently superior efficacy of antibiotic therapy, together with their non-infective nature, but also a certain controversy about the robustness of the phage therapy trials back then (46).

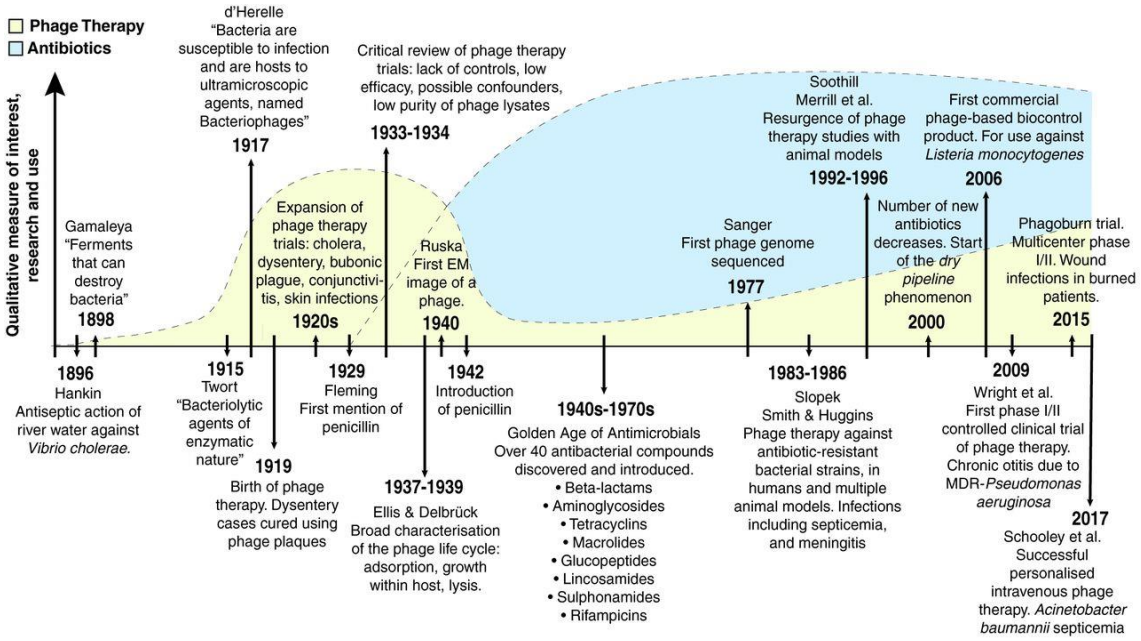


Figure 1.4. Timeline of the major events in phage therapy and antibiotics. Taken from (48).

Nonetheless, phages were continued to be used as a complement or alternative to antibiotics in Eastern Europe and, particularly, in the (former) Soviet Union. Two institutions maintained the phage therapy legacy in the East: the Eliava Institute of Bacteriophage, Microbiology and Virology, in Tbilisi, Georgia, and the Ludwik Hirszfeld Institute of Immunology and Experimental Therapy, in Wroclaw, Poland (49). As a conclusion thus far, the 'Golden Age' of antibiotic development (1940–1970) completely prevented the good fortune of phage therapy as a real alternative, especially in the West, and, while intense research still continued in Eastern Europe, several alleged problems of phage therapy and their research contributed to precluding its wide use as prophylactics or

therapy. Among such problems, Sulakvelidze *et al*. cite: i) a narrow host range of phages that may have contributed to produce negative results in trials due to a defective pre-screening process against the targeted bacterial strain; ii) insufficient purity of early therapeutic phage preparations (including the presence of inflammatory endotoxins); iii) low stability or viability of phage solutions; iv) lack of knowledge on the mechanism of action of phages (*e.g.*, temperate phages do not produce generalized lysis); v) exaggerated claims of effectiveness of commercial phage preparations; vi) failure to establish convincing scientific proof of efficacy, mainly due to studies lacking placebo controls or subjective interpretation when placebo controls were used (46).

Since those early reports, the landscape of antimicrobial therapy has changed a lot. With the 'dry pipeline' of new antibiotics and the rise of AMR, a true 'renaissance' of the phage therapy concept has occurred. The newly sparked interest in phages has granted new efforts to gather evidence supporting this alternative therapy. Although the first randomized and controlled trial of phage therapy was conducted back in 1966 (50), the first one to have an impact in the phage therapy renaissance era was published in 2009 (51). In that trial, the safety of a topical phage preparation was proven, and no differences were found between treatment groups in terms of efficacy. Similar results have been provided by a series of later trials examining different routes of administration (topical, oral, intravenous…) (48). Thus, safety can be considered proven not only by the classical evidence of almost 100 years of safe use of phage therapy in Eastern Europe but also by recent clinical trials that comply with current standards. In fact, the oldest reports of alarming side effects (*i.e.*, inflammatory or anaphylactic reactions) can be attributed to contaminants from early phage preparations (probably, lipopolysaccharides or other molecules from the bacterial host) (48). Other concerns of phage therapy, such as the involvement of phages in transmitting virulence factors or resistance genes can be also overcome by current technology (for example, by examining the phage genomes before their application and issuing some sort of 'genetic passport' for phages used for therapy, discarding temperate phages and using exclusively virulent ones). Nonetheless, since most proteins are still hypothetical, with no experimental data to clarify their function, this may pose an important challenge in practically implementing phage therapy.

Phage therapy trials have also experienced some drawbacks in this new renaissance era. The trial Phagoburn, for example, which ran between 2015 and 2017 to evaluate the topical treatment of burn wounds infected with *P. aeruginosa* with a 12 phages cocktail, failed to live up to expectations (52). While, again, they reported no side effects due to the phage product, and the cocktail decreased the bacterial load in the wounds, this decrease happened at a slower rate than with the standard-of-care treatment (thus superiority or non-inferiority was not proven). The main reasons for this failure have been

cited to be a diminishing in the phage titers of the cocktail after manufacturing (thus patients received less than the initially expected dose) and also that some bacterial strains from the patients were resistant to the phages in the cocktail (52). Other studies that took place between 2005 and 2017 on the safety and efficacy of an oral administration of phages to treat diarrheal disease failed to detect viral replication in the gut of healthy individuals nor patients, and no significant beneficial effects were found for phage therapy in comparison with standard-of-care (53). These issues, which have been mainly attributed to experimental design mistakes (54), have not precluded, however, the continuance of phage therapy trials with different approaches. For example, in 2018, a phage cocktail was tested as a prebiotic to treat gut dysbiosis with some success in reducing symptoms of intestinal inflammation (55). Other clinical trials are also underway, either in phase 1-2 (NCT02664740, NCT03808103) or entering phase 3 (NCT03140085).

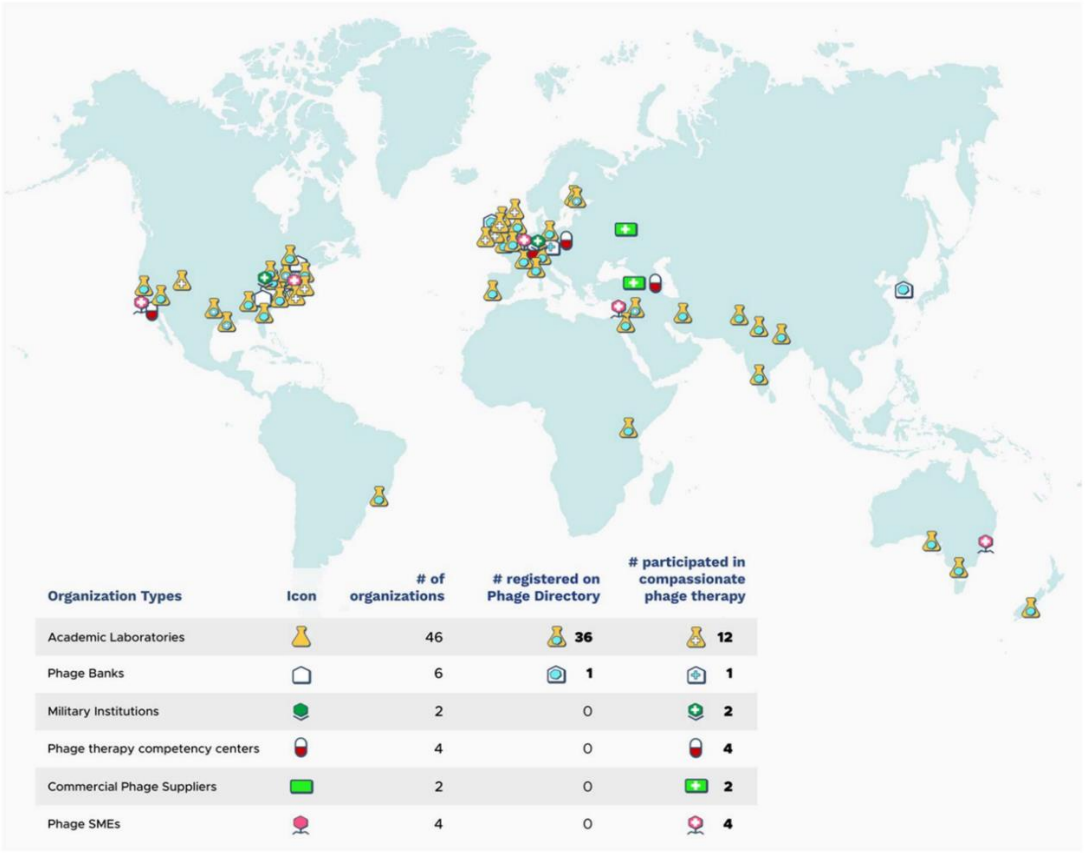


Figure 1.5. Geographic distribution of institutions that have participated in compassionate phage therapy cases or have demonstrated interest to do so. Taken from (56).

Nonetheless, the translation of phage therapy from the lab to the bedside must still walk a long way. The so-called 'compassionate use', however, can shorten such a

journey. It is known as compassionate use the application of not yet approved drugs or medical devices, without the coverage of a clinical trial, on a specific patient whose condition precludes the successful use of approved therapies (56). The advantages of compassionate phage therapy are: i) the possibility of immediate clinical use for a particular need or patient; ii) obtaining data to inform future work and iii) its applicability to virtually any form of phage therapy, provided that the conditions for compassionate use are fulfilled (48, 56). There have been many examples of successful compassionate phage therapy in literature, many of them encouraging the practical possibility of introducing phage therapy in clinical practice (**Figure 1.5**). However, the drawbacks of the recent clinical trials suggest that perhaps the approach should substantially vary to achieve some sort of universally recognized success.

To begin with, phages are known to attack specific bacterial hosts: often they are species-specific, sometimes even they only target a limited number of strains. Secondly, the nature of strictly virulent phages, which imposes bacterial host killing for their propagation, produces, as a consequence, an antagonistic co-evolution between host and parasite (54). In this reciprocal evolution, the host (bacterium) evolves to acquire resistance to phage infection, and, thus, this provokes the selection of phages able to overcome said resistance. This dynamic behavior makes phages indisputably different than other antimicrobials. Therefore, the current understanding of phage therapy is drifting away from the common chemotherapeutic framework of 'static' drugs (*i.e.*, antibiotics: one specific drug is approved and marketed for a given application). In this way, the proposed future for phage therapy works more like a 'personalized' or 'precision therapy', in which the infective bacteria isolated from the patient are screened against a 'phage library' from which the more suitable phages, *i.e.*, to which the bacteria are susceptible, are chosen for the treatment (54, 57). In fact, in most of the more recent successful case reports of phage therapy, phages were selected to target the patient's infecting strains and even phage mutants have been selected *in vitro* with enhanced therapeutic properties (58, 59). This new paradigm even gives room to some theoretical elaboration, such as the one presented by J. P. Pirnay: a world in which the infectious bacterium is automatically identified by genome analysis and a suited phage to treat it is then synthesized based on an artificial intelligence-driven platform (60). What this latter approach generally lacks nowadays is a consistent legal framework. The most advanced experience in this direction is that of Queen Astrid's Military Hospital in Belgium, favored by the acceptance of phage therapy as magistral preparations (61). Under European law, a magistral preparation is "any medicinal product prepared in a pharmacy in accordance with a medical prescription for an individual patient" (European Directive 2001/83/CE, Art. 3). Therefore, the proposal was to consider phages as active pharmaceutical ingredients of magistral

preparations to be delivered to specific patients under the direct responsibility of a medical doctor and a (hospital) pharmacist. Requirements for this framework were, basically, the production and external quality control of the phage active ingredients based on an *ad hoc* monograph (which, in this case, was jointly prepared by experts in the Queen Astrid military hospital, the national regulatory agency, and the Belgian Institute of Public Health). Under this regulation, phage therapy can be currently administered in Belgium. This experience proves that the use of phage therapy requires, first of all, an expert initiative (given the special character of phage as medicines), and then the awareness (and will to collaborate) of the competent authorities to set up a suitable regulatory framework, preferably within the limits of the current regulation. In Spain, the case for compassionate use or a magistral phage preparation is uncertain. The competent agency (Agencia Española de Medicamentos y Productos Sanitarios, AEMPS) has not issued an official opinion on phage therapy. In early 2020, a group of experts from the FAGOMA network (62), comprising researchers in phage and phage-derived products and their biomedical application (including the author of this thesis and his thesis supervisor), compiled a comprehensive dossier on the current phage therapy state-of-the-art for the AEMPS. The agency is currently reviewing it, and an official statement on the matter is still, supposedly, on its way.

Despite all these efforts, the industrial and financial ecosystem still prefers, nowadays, a phage cocktail-based approach, more in tune with the current pharmacoeconomic model (54). Given the difficulties that such 'universal' phage cocktails have encountered in their development, some other alternatives have arisen.

3. REPURPOSING PHAGE LYSINS AS ANTIMICROBIALS: ENZYBIOTICS

Not all phages lyse the host bacterial cell to release the progeny, but the ones that do so in the course of their infective cycle (**Figure 1.6**) have devoted an important number of genetic resources to the lysis itself. Lysis is accomplished by breaking the bacterial outer barriers. This is to say that, for lysis to occur, the cell wall must be broken. The bacterial cell wall is an essential and highly conserved structure that is found in a few different arrangements among bacteria. It always contains a layer of a conserved and complex polymer, the peptidoglycan, which is made up of a glycan strand (composed of alternating β-(1,4)-linked residues of *N*-acetylglucosamine [NAG] and *N*-acetylmuramic acid [MurNAc]) linked to a peptide subunit of three to five alternating L- and D- amino acids (aa) that stems out of the MurNAc.

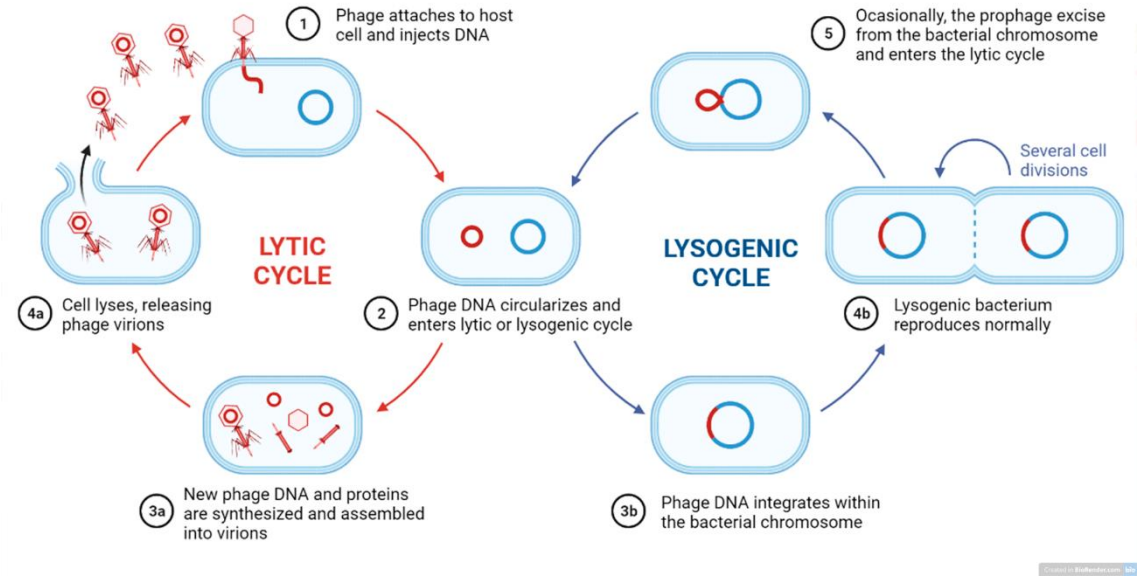


Figure 1.6. Phage infective cycles. Relevant steps of the lytic and lysogenic cycles of phages are explained.

The peptide subunits of different glycan strands are cross-linked to each other, on many occasions through an interpeptide bridge of variable length (**Figure 1.7A**). The specific composition and structure of peptidoglycan is a taxonomical trait, *i.e.*, phylogenetically close bacteria tend to display the same peptidoglycan variants (63). Among the many essential functions of the cell wall, it works as a limiting barrier that provides the cell its shape, facilitates and regulates the exchange of substances between the bacterium and the environment and contains the osmotic pressure due to the solute-crowded cytoplasm in comparison with the hypotonic outside (64). The capital importance of the cell wall in general and of peptidoglycan in particular makes it a most relevant target for antibacterial drugs. In fact, the first-ever discovered group of antibiotics, β-lactams, together with many others, target the molecular machinery that is responsible for synthesizing or remodeling cell wall components such as the peptidoglycan itself. This is also convenient since there is no eukaryotic counterpart to the bacterial cell wall, therefore antimicrobials that target such element and its biosynthesis/remodeling machinery have a lesser chance to harm eukaryotic hosts (65). Phages have evolved strategies to break this barrier, mainly by disrupting the peptidoglycan structure. They do so by timely expressing a particular enzyme, termed (endo)lysin, which functions as a murein hydrolase, *i.e.*, by catalyzing the hydrolysis of specific bonds within the peptidoglycan mesh.

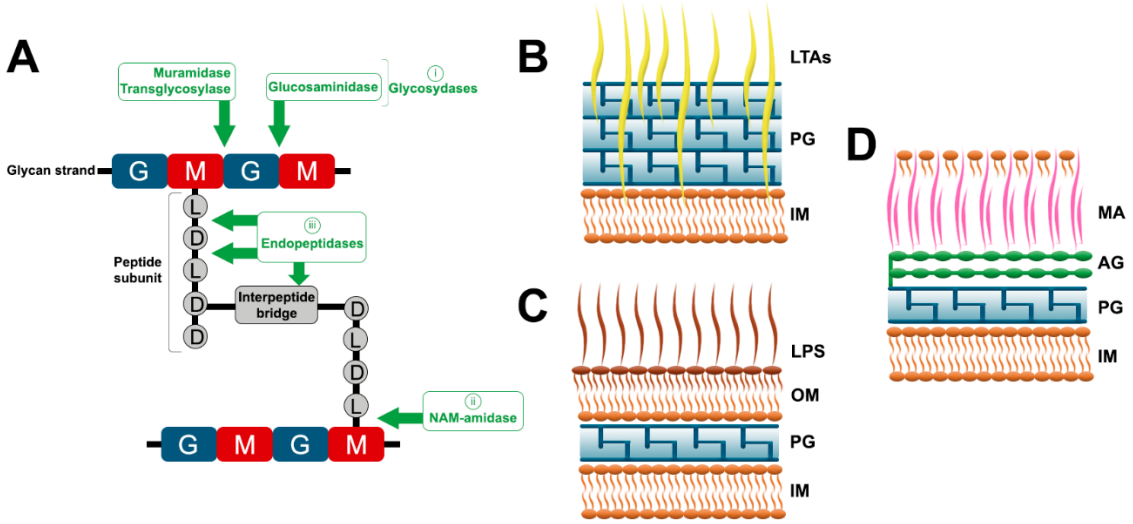


Figure 1.7. Peptidoglycan and cell wall structures. A: Schematic structure of peptidoglycan, including catalytic activities present among phage lysins with their cleavage sites. G = N-acetylglucosamine, M = N-acetylmuramic acid, L = L-aa, D = D-aa. **B**, **C, D:** Schematic representation of the cell wall architectures of G+ (**B**), G− (**C**), and *Mycobacterium* and relative bacteria (**D**). IM = inner membrane, PG = peptidoglycan, LTAs = (lipo)teichoic acids, OM = outer membrane, LPS = lipopolysaccharides, AG = arabinogalactan, MA = mycolic acids.

According to the bond they break, lysins can be classified into different catalytic activities (**Figure 1.7A**) (66). A general classification relies on the peptidoglycan element they attack. In this regard, they are (i) glycosidases, if they target the peptidoglycan glycan bonds, (ii) *N*-acetylmuramoyl-L-alanine amidases (NAM-amidases, EC 3.5.1.28), if they break the bond between the glycan strand and the peptide subunit, and (iii) (endo)peptidases if they target any point within the peptide subunit or the interpeptide bridge (EC 3.4.X.X). Then, within these classes, different specific catalytic activities can be distinguished. Among glycosidases, some break the β(1,4) glycosidic bond between MurNAc and NAG, on the reducing side of the former. Depending on the mechanism, these enzymes can be either lytic transglycosylases (EC 4.2.2.n1), if they operate by a non-hydrolytic mechanism that cleaves the glycosidic bond by an intramolecular reaction, resulting in an *N*-acetyl-1,6-anhydro-muramyl moiety; or *N*-acetylmuramidases (EC 3.2.1.17), generically termed just 'muramidases' or 'lysozymes'[3], if they are true murein hydrolases. Peptidoglycan glycosidases can also break the bond between NAG and MurNAc, on the reducing side of NAG, in which case they are *N*-acetylglucosaminidases (or just 'glucosaminidases', EC 3.2.1.52). Among murein peptidases, there are also a variety

[3] The term 'lysozyme' has provoked some equivocations in the literature. The generic 'lysozyme' has been used to name, in general, phage lysins, although its exact meaning is that of N-acetylmuramoyl-β-1-4-N-acetylglucosamine hydrolase. This probably dates back to the historic naming of such enzymes, as the so-called T7 'lysozyme', which is a misleading name, given that, actually, it is a NAM-amidase (67).

of specific activities. The most common activities cleaving within the peptide subunit would be L-alanyl-D-glutamate endopeptidases or γ-D-glutamyl-meso-diaminopimelate endopeptidases (EC 3.4.19.11), while there exists a somewhat wide range of interpeptide bridge endopeptidases, given the already variable nature of such peptidoglycan component (there are, for example, D-alanyl-glycyl endopeptidases, glycyl-glycyl endopeptidases, etc.).

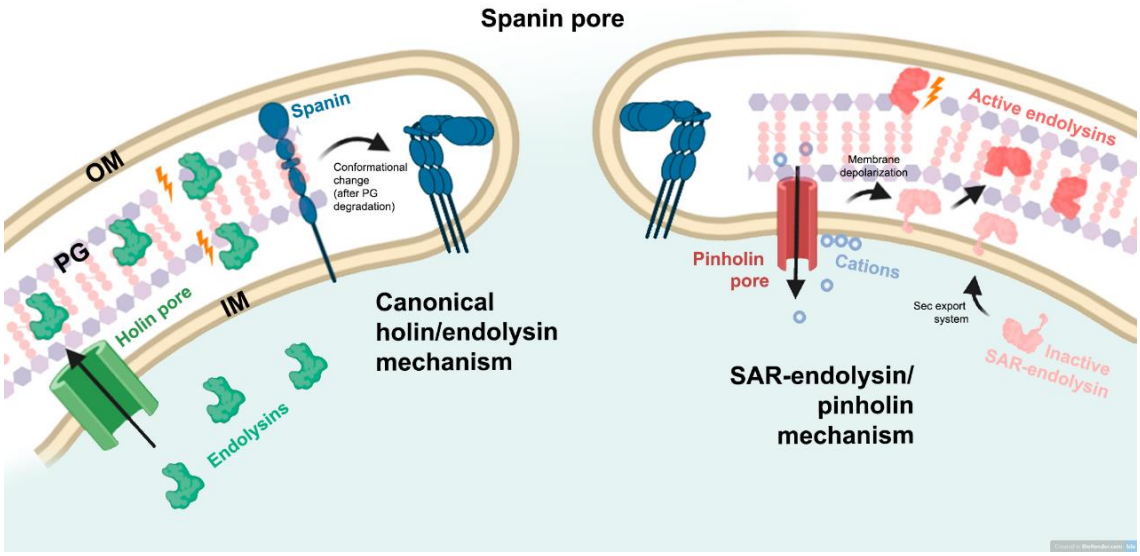


Figure 1.8. Different phage-encoded lysis mechanisms. PG = peptidoglycan; IM = inner membrane.

When the lysin is expressed, it reaches the periplasmic space and catalyzes the cleavage of certain peptidoglycan bonds. This disrupted peptidoglycan is no longer able to contain the osmotic pressure, and thus lysis occurs by osmotic shock. For the lysin to be able to reach the periplasm, additional phage gene products are required (**Figure 1.8**). In particular, the canonical phage-mediated lysis theory implicated a second protein, the holin, which accumulates within the cytoplasmic membrane in such an arrangement that produces non-specific pores through which lysins are leaked towards its polymeric target (68). A different strategy is that of the signal-anchor-release (SAR) endolysin (**Figure 1.8**). These lysins contain an N-terminal so-called SAR which engages the host *sec* mechanism that, therefore, exports the lysin, which then remains anchored to the cytoplasmic membrane. For the SAR-lysin to be released, a sort of holin intervenes, named pinholin, which creates thin pores in the membrane (with a nanometric diameter, in contrast with the micrometric range of the canonical holin channels). Such pores cause the membrane to depolarize, and such depolarization functions as a signal that causes the detachment and refolding of the SAR-lysin into a catalytically active conformation

(68, 69). Finally, there are some additional genes or lysis variants coupled to different arrangements of the cell wall. The G− cell wall differs from that of G+ in that the former contains just a thin peptidoglycan layer that is covered by an outer membrane (OM), while the latter has a thick peptidoglycan layer and, typically, an outer layer of (lipo)teichoic acids (**Figure 1.7BC**). The G− OM is composed, in its inner face, by phospholipids, and in the outer one, by lipopolysaccharides (LPS). Since these LPS contain negatively charged phosphate groups, the OM is electrostatically stabilized by bound divalent cations (Mg^{2+}, Ca^{2+}). Given the presence of the OM, the phages that infect G− bacteria have evolved a unique strategy to overcome such barrier: spanins (**Figure 1.8**). Spanins are proteins that work either as a single molecule or as a two-component complex to insert both in the inner and the outer membrane (by a transmembrane domain and a lipoylated end, respectively) and span the periplasm. The nonenzymatic action of spanins provokes the fusion of inner and outer membranes and the formation of lethal pores after lysin-mediated peptidoglycan degradation (70). It has been established that spanins are essential for effective lysis of G− bacteria (71). An additional element is present in phages infecting *Mycobacterium* and some other Corynebacteriales (such as *Rhodococcus* or *Corynebacterium*). Although originally considered G+, these bacteria are, as G−, diderm (*i.e.*, contain both an inner and an outer lipidic layer). In the case of Corynebacteriales, the outer layer is composed of an arabinogalactan moiety, covalently attached to the peptidoglycan and to an external coating of fatty acids, mainly mycolic acids (72) (**Figure 1.7D**). Due to this particular structure, the mycobacterial cell wall is sometimes referred to as a mycolyl-arabinogalactan-peptidoglycan complex. The presence of the mycolyl-arabinogalactan is an additional hurdle for bacterial lysis, therefore mycobacteriophages, corynephages, etc., usually encode an esterase activity that detaches the mycolic acids esterified to the arabinogalactan layer. This esterase activity is known as LysB (or lysin B) and cooperates with a canonical murein hydrolase (LysA or lysin A) to trigger bacterial lysis (73).

This lytic arsenal, evolved by phages to specifically and efficiently lyse (and thus kill) their bacterial hosts, has been proposed to be repurposed for antibacterial means. To tell the whole story, the antimicrobial potential of lysozymes has been long known (even since Fleming's reports of lysozyme in human secretions (74); or with many other reports or suggestions throughout phage research history). However, in our opinion, it was the seminal proposal by Vincent Fischetti *et al.* in 2001 (75, 76) that triggered the collective realization that a huge reservoir of genetic resources to fight harmful bacteria was out there to be profited (77). Since then, the exogenous use of recombinantly produced and purified lysins (and the scientific literature about it) has done nothing but grow. The concept underlying the so-called 'enzybiotics' (lysins repurposed as antimicrobials) is

remarkably simple: since the lysin is the one phage molecule responsible for host lysis ('from the inside'), let us use it 'from the outside', in a pure form, to kill infectious bacteria by inducing the breakage of their peptidoglycan. Since the introduction of the enzybiotic concept, many advantages have been pointed out advocating for their use, and very few drawbacks in comparison with antibiotics of whole virion phage therapy (**Table 1.1**). Among the advantages: i) a certain specificity towards the original bacterial host and some closely related bacteria, which would prevent normal microbiota to be harmed (78, 79), or, conversely, the possibility to have broad-range lysins, if needed (80); ii) a lower chance to provoke the appearance of *de novo* resistant bacteria, which is thought to be because of the essential nature of the highly conserved peptidoglycan (this is, changes in its structure would lead to a decreased fitness and/or virulence) (81, 82); iii) neither adverse immune responses nor production of neutralizing antibodies are expected, possibly due to the usual presence of phages —and their products— among the normal cohabitating microbial populations in humans (83). Besides, the particular mode-of-action of lysins enables them to remain unaffected by antibiotic resistance mechanisms (*i.e.*, they can be equally active against antibiotic-resistant strains), and they also display activity against metabolically inactive cells (*e.g.*, persister cells).

Table 1.1. Summary of some advantages and disadvantages of enzybiotics in comparison with antibiotics and phages. Cells are coloured to remark a negative feature (red), a positive one (green), or a somewhat neutral one (orange).

Feature	Antibiotics	Phages	Enzybiotics
Specificity	Low specificity	Highly specific (species or strain-specific)	Highly specific, tunable specificity
Resistance	Medium basal population resistance proportions, easily selected resistance genes	Easily evolved resistance, but compensated by reciprocal host-parasite co-evolution and high phage variability	Low chance of resistance, reduced fitness/virulence in resistant mutants
Side effects	Many known side effects (allergies, dysbiosis, etc.)	Absence of severe side effects	Absence of severe side effects
Pipeline	Slowed-down pipeline since the 1980s	High natural diversity, possibility to engineer *ad hoc* synthetic phages	High natural diversity, possibility to engineer *ad hoc* synthetic phages. Very easy to manage in the lab
Industrial production	Currently worldwide produced	May require specific infrastructure and expertise, especially under the personalized therapy paradigm	Technology in place (experience from other biomedical heterologous proteins)
Bedside handling	Very easy, consolidated standard-of-care	Easy administration, although may require joint expertise (or technological solutions) to couple diagnostic, prescription, and drug preparation	Easy, current protocols can be easily adjusted
Regulatory status	Approved and in use	Only compassionate use or personalized magistral preparations. Some infrastructure in place.	Ongoing clinical trials

This said, the actual behavior of lysins as antimicrobials will not be fully apprehended until their introduction in the clinic. For example, a recent review suggests that horizontal

gene transfer, a not easily reproduced evolutionary force in laboratory conditions, may spread lysin resistance from those intrinsically resistant species to the initially susceptible ones (84). On the other hand, the same work suggests that the higher cost of lysin-based therapy compared with antibiotics would probably restrict its misuse and thus lessen the concerns of forcing the evolution towards generalized resistance.

It is also worth mentioning that there are many examples in the literature of synergy[4] between enzybiotics and antibiotics, allowing the formulation of complementary combination therapies that could also include the current standard-of-care drugs (85), and they can also act synergistically with each other due to them breaking different peptidoglycan bonds (86). The process of killing bacteria using lysins is straightforward in the case of G+ bacteria, whose peptidoglycan is readily available from without. Thus, the therapeutic potential of lysins against G+ has been extensively proven both *in vitro* and *in vivo* for a wide variety of bacteria (**Table 1.2**). Another important benefit of enzybiotics is that they very well fit in the traditional new drugs regulatory path. As a consequence, there are many companies around the world investing in enzybiotic therapeutics[5], and even some clinical trials are underway (87) (NCT03089697), with at least one entering Phase III (NCT04160468).

[4] A synergistic activity is defined as the one that arises from the cooperative action of two or more agents in such a way that the result of the combined action is greater than the mere sum of the activities of each of the agents involved.

[5] Some of the companies currently engaged in developing enzybiotic products are: ContraFect (https://www.contrafect.com), iNtRON Biotechnology, Lysando (www.lysando.com), GangaGen (https://gangagen.com), Telum Therapeutics (https://telumtherapeutics.com), Micreos (https://www.micreos.com) or Bioharmony Therapeutics (https://bioharmonytherapeutics.com).

Table 1.2. Some examples of phage lysins experimentally proven as enzybiotics against G+ and *Mycobacterium*. Taken from (83).

Species	Lysin/phage	Susceptible bacteria tested	Methodology used		Acc. No.; comments	References
			In vitro	*In vivo*		
S. pneumoniae						
	Pal/Dp-1	Pneumococci and relatives	Biofilm; synergy with Cpl-1	Colonization and sepsis (mice)	O03979	(76, 88-90)
	Cpl-1/Cp-1	Pneumococci and relatives	Biofilm; synergy with Pal and antibiotics; cell culture	Colonization, otitis, pneumonia, sepsis (mice)	P15057	(88-97)
	LytA	Pneumococci and relatives	Biofilm	Sepsis (mice)	P06653; major autolysin	(90, 98)
	Cpl-7/Cp-7	Streptococci; other G+	Biofilm		P19385	(97, 99)
	Cpl-7S	Streptococci; other G+	Cell culture	Colonization (mice), pneumococcal infection (zebrafish)	Engineered protein	(97, 99)
	Cpl-711	Pneumococci and relatives	Biofilm, synergy with antibiotics; cell culture; synergy with PL3	Colonization and sepsis (mice), pneumococcal infection (zebrafish)	Chimera of Cpl-7 and Cpl-1	(80, 86, 97)
	PL3	Pneumococci and relatives	Biofilm; synergy with Cpl-711	Pneumococcal infection (zebrafish)	Chimera of Pal and LytA	(86, 100)
GAS						
	PlyC/C_1	GAS and other streptococci	Biofilm, cell culture (intracellular killing of GAS)	Colonization (mice)	J7M5V6	(75, 101, 102)
	PlyPy/MGAS315 prophage	GAS and other streptococci		Sepsis (mice)	AAM79913	(103)
S. agalactiae (GBS)						
	PlyGBS	GAS, GBS and other streptococci		Colonization (mice)	Q5MY96	(104, 105)

Table 1.2 (Cont.). Some examples of phage lysins experimentally proven as enzybiotics against G+. Taken from (83).

Species	Lysin/phage	Susceptible bacteria tested	Methodology used		Acc. No.; comments	References
			In vitro	*In vivo*		
S. aureus						
	Lysostaphin	Staphylococci	Biofilm, synergy with LysK, $CHAP_K$ and antibiotics, controlled release	Sepsis and colonization (mice, rats)	P10547; from *Staphylococcus simulans*	(81, 106-117)
	LysK/K	Staphylococci	Biofilm, complex with peptides		Q6Y7T6	(116, 118, 119)
	$CHAP_K$	Staphylococci	Biofilm, synergy with lysostaphin, controlled release	Colonization (mice)	*CHAP* domain of LysK	(117, 120-123)
	ClyS	Staphylococci	Synergy with oxacillin and vancomycin	Colonization and septicemia (mice)	Chimera of Twort phage lysin (O56788) and phiNM3 phage lysin (Q2FWV2)	(124)
	SAL-1/SAP-1	Staphylococci	Biofilm	Bacteremia (mice), toxicity and pharmacokinetics (rats, dogs, monkeys), pharmacokinetics and pharmacodynamics (healthy humans)	SAL200 is a drug formulation of SAL-1	(87, 125-128)
	P128	Staphylococci	Biofilm activity, cell culture, synergy with antibiotics	Colonization and sepsis (rats)	Chimera of Gp57 (Q6Y7R1) and lysostaphin. Under clinical testing	(129-136)
	LysGH15/GH15	Staphylococci	Biofilm	Sepsis and pneumonia (mice)	D6QY02	(137-141)
	CF-301 (PlySs2)/*S. suis* 9/1591 prophage	*S. aureus, S. pyogenes, S. pneumoniae*; other G+	Biofilm, synergy with antibiotics	Sepsis (mice)	M1NS67; under clinical testing	(82, 142, 143)
	ClyF	Staphylococci	Biofilm	Sepsis (mice)	Chimera of Ply187 (O56785) and PlySs2	(144)
Mycobacterium sp.						
	LysB/Ms6	Mycobacteria	Growth inhibition with surfactants		Q9ZX49; esterase	(145, 146)
	LysB/Bxz2	Mycobacteria	Growth inhibition with surfactants		Q9FZR9; esterase	(145)
	LysA/BTCU-1	Mycobacteria	Cell culture		O64203; intracellular killing of *M. smegmatis*	(147)
	LysB/BTCU-1	Mycobacteria	Cell culture		R9R591; intracellular killing of *M. smegmatis*; esterase	(147)

4. SECOND GENERATION ENZYBIOTICS: CHIMERIC LYSINS

Another huge advantage of phage lysins is their very well described modular architecture. This is to say that phage lysins are typically composed of several different functional modules, usually an enzymatically active domain (EAD), which catalyzes the cleavage of the peptidoglycan mesh, and a cell wall binding domain (CWBD), which serves to recognize and bind a target ligand within the cell wall, directing the lysin towards its substrate (78). This structure appears especially among lysins from phages that infect G+, whereas it is not so common among those of G− (which more typically bear monomodular lysins, harboring a single EAD). The modularity of phage lysins makes them amenable for synthetic biology strategies based on the construction of chimeric proteins, made up of modules from different origins. This was early shown by our laboratory as proof of the modular evolution of lytic enzymes from the system of *S. pneumoniae* and its phages (148). Recently, a periodization for lysins research has been proposed (149), according to which the first generation of lysin-based therapeutics was that of wild type lysins, then the second generation is that of lysins engineered (either by the combination of modules or by other strategies) to improve biochemical properties (bactericidal activity, stability…). The third generation, which will be reviewed later on, is that of engineered lysins with improved clinical properties (bioavailability, biocompatibility, and half-life) (**Figure 1.9**).

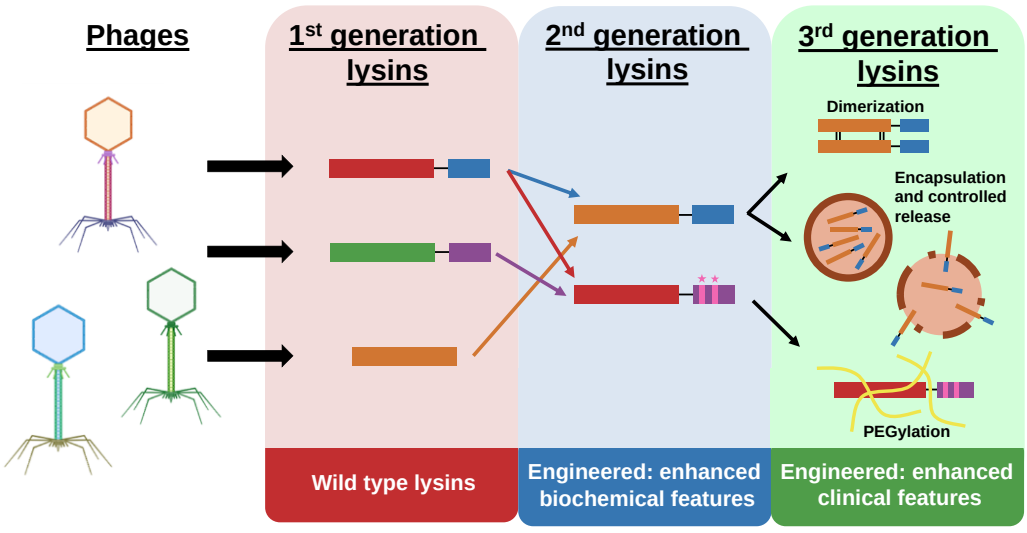


Figure 1.9. Periodization of lysin-based antimicrobials research.

In our laboratory, several chimeric lysins have been constructed within a second-generation framework, to either improve activity (Cpl-711, PL3) or to tune the activity range

(Csl2, **Figure 1.10**) (80, 100, 150). Cpl-711 was constructed by fusing the N-terminal EAD of Cpl-7, the lysin of pneumococcal phage Cp-7, which is a muramidase domain from the *Glyco_hydro_25* (PF01183) family, to the linker and C-terminal CWBD of Cpl-1, lysin of Cp-1 phage. Such CWBD (from family *CW_binding_1*[6], PF01473, made up of six almost perfect repeats) has the peculiarity that it specifically recognizes the choline residues found within the teichoic acids of pneumococci and some other related bacteria. Thus, such repeats are named "choline-binding repeats" (CBRs), and the CWBD has been also termed "choline-binding domain" (CBD). The CBDs are common elements of relevant surface proteins within the pneumococcal system, and thus the presence of choline in the pneumococcal teichoic acids is a required trait for it to maintain normal physiology (151). The original CWBD of Cpl-7, on the contrary, belongs to family *CW_7* (PF08230) and it recognizes a conserved and widespread peptidoglycan motif (152); therefore, it is not as specific as Cpl-1 or Cpl-711.

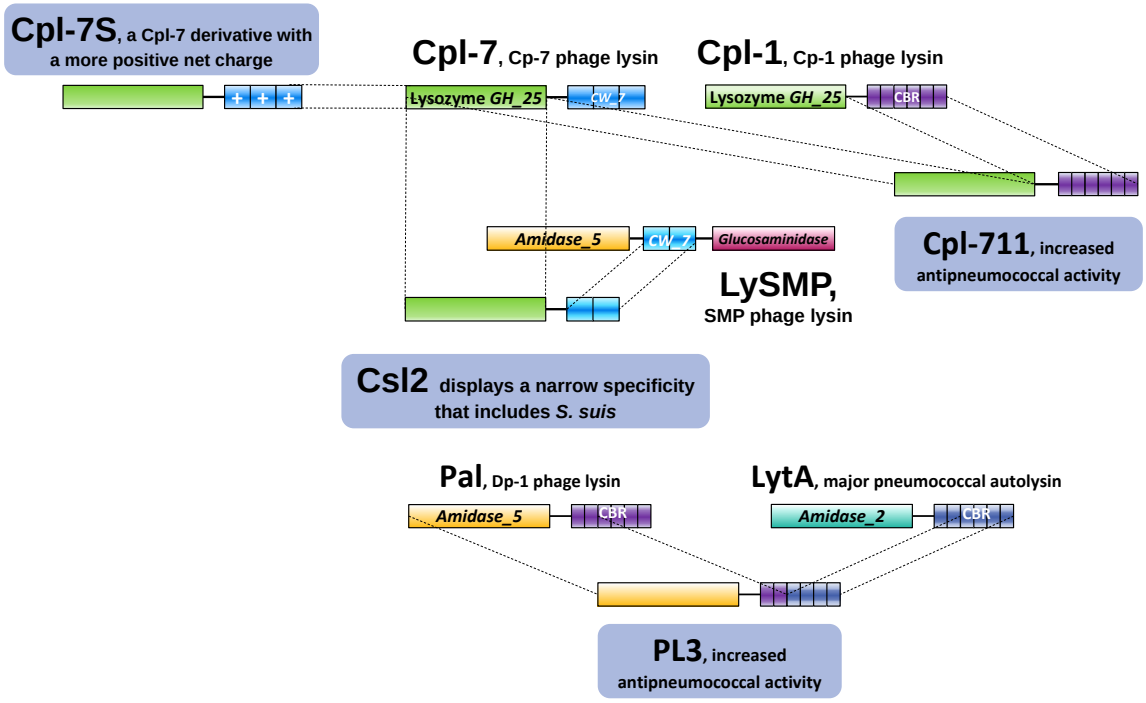


Figure 1.10. Engineered enzybiotics previously constructed at the Host-Parasite Interplay in Pneumococcal Infection group of CIB Margarita Salas. Chimeric lysins are those indicated in a blue-shaded frame.

[6] *CW_binding_1* family has been renamed as *Choline_bind_1* in the latest Pfam version (34.0). However, the name *CW_binding_1* was maintained in this thesis for coherence with the many results presented (for example, in RESULTS **Chapter 1**) based on a previous Pfam release.

The chimera Cpl-711 was found to be more efficient as an antimicrobial than both parental Cpl-7 and Cpl-1 enzymes (80, 86, 97). Since the EAD of Cpl-1 and the one of Cpl-711 is the only differing section between both enzymes, it was speculated that the improved activity of Cpl-711 relative to Cpl-1 was due to Cpl-7 EAD having evolved towards a higher catalytic efficiency given that the parental wild type enzyme contains a less specialized CWBD (80). Cpl-711 has been tested both *in vitro* and *in vivo* in a variety of settings, such as pneumococcal biofilms (80), to prevent pneumococcal carrier-state in an *in vivo* murine model (97), or in synergy with antibiotics (85). PL3 was a similar case of artificial evolution of lysins by a rational combination of domains. In this case, the EAD of Pal, lysin of the pneumococcal phage Dp-1, which is a NAM-amidase belonging to the *Amidase_5* family (PF05382), was fused to the last four C-terminal CBRs of *S. pneumoniae* major autolysin LytA, preserving the linker and first two CBRs of Pal. PL3 was equally superior to both parental enzymes in terms of antimicrobial activity (86, 100). A synergistic action between both chimeric enzymes, Cpl-711 and PL3, has been observed, proposed to be because they cleave different peptidoglycan bonds. Such synergy has enabled to lower the dose of enzymes needed to achieve an *in vivo* therapeutic effect down to minimum amounts (86). Csl2 was constructed to check whether slightly different versions of CWBDs from *CW_7* had different bacterial specificities. Thus, Csl2 is a version of Cpl-7 but bearing the CWBD of lysin LySMP from the *Streptococcus suis* SMP phage. Csl2 was shown to have a lytic preference for *S. suis* strains in comparison with the parental Cpl-7 (150). Engineering by point mutation was applied in the case of lysin Cpl-7S to invert the net charge of its CWBD (making it more positive), on the basis that a higher net charge would allow better interaction with the negatively charged bacterial surface (153). Using this approach, a broad-range lysin was obtained with a more potent bactericidal effect than the parental Cpl-7 (99).

Currently, chimeric lysins generation has been elevated to the next level by applying a custom, high throughput cloning platform based on type IIs restriction enzymes termed ‘VersaTile’, developed by the group of Yves Briers at Ghent University (154). The main advantage of these kind of strategy is that it enables to massively construct and then test many modules and chimeric variants at once. Therefore, many interesting conclusions can be reached, such as rational ‘construction rules’ for further engineering (155), or variants enriched against a certain target or bearing specific traits of interest (154). With these novel tools, a remarkable advance in the synthetic biology of modular lysins is warranted for the next few years.

5. ENZYBIOTICS AND GRAM-NEGATIVE BACTERIA

Second-generation lysins have been especially useful to overcome the OM (156). Indeed, for some time since the foundational proposal of lysins as a new generation of antimicrobials for the post-antibiotic era, G− bacteria had been deemed refractory to lysins activity when added from without the cell. The explanation relied on the presence of the G− OM, which would act as a permeability barrier to prevent the lysin from reaching its target, the bacterial peptidoglycan. Many initial attempts at enabling lysins activity from without against G− bacteria used co-administered OM-permeabilizing molecules (*e.g.*, EDTA, organic acids, essential oils, etc., **Figure 1.11A**) with good outcomes in terms of bactericidal activity (99, 157, 158). However, although these combinations may be readily applicable in fields such as food preservation, the co-administration of OM-permeabilizers could prove challenging or even impracticable for *in vivo* treatments. Due to this potential difficulty, later approaches focused on the engineering of the enzymes themselves (**Figure 1.11B**). The common approach is the fusion of the enzyme to membrane-interacting parts of a diverse nature, giving rise to some lysin-based neologisms: i) artilysins, lysins fused to OM-permeabilizing antimicrobial peptides (AMPs) (159), ii) lysocins, fusions with bacteriocin elements (160), iii) innolysins, containing phage receptor binding proteins (161). Irrespective of the name (*i.e.*, of the identity of the additional membrane-active part), all such chimeras should be considered under the same anti-G− paradigm.

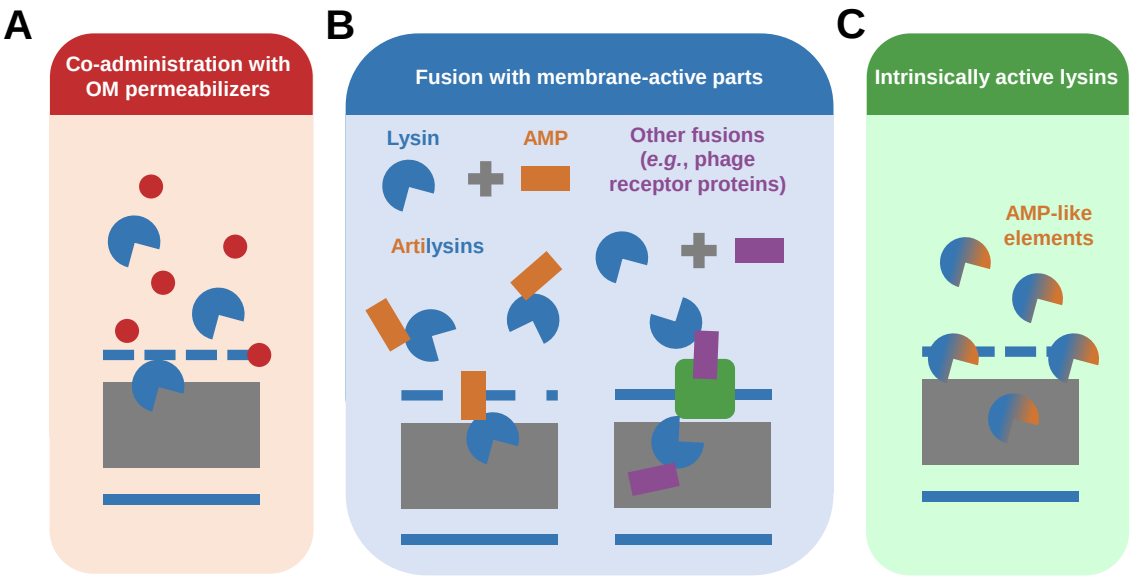


Figure 1.11. Strategies to render G− bacteria susceptible to exogenously added lysins. A: Use of wild-type lysins in combination with OM permeabilizing agents (such as EDTA, essential oils, or organic acids). **B**: Generation of synthetic lysins containing membrane-active parts (such as AMPs, receptor binding proteins, or bacteriocin elements). **C**: Intrinsically active enzymes from without the G− cell.

Although the first attempts to test the antimicrobial activity of wild type lysins against G− bacteria failed to convey as exciting results as those rendered by their G+ counterparts, there already were several notions in the literature suggesting an intrinsic ability of some phage lysins to interact with the G− surface with a bactericidal effect, mostly from the laboratory of Prof. Tanji Yasunori in the early 2000s (162-164). Such reports pointed out to non-enzymatic elements (normally located at C-terminal region) within the very proteins, capable of killing bacteria via membrane damage in a similar way as AMPs do, and as previously hinted for common lysozymes (165, 166) and later also for those of phage origin (167). More recent reports have since also joined such evidence, suggesting that the AMP-like subdomains represent a relatively common trait, and even proposing the use of peptides based on these elements as AMPs or parts for further development of antimicrobials (168-171). It has been suggested that such AMP-like elements might cooperate to host lysis by providing an additional affinity towards the cell wall, because of their high net charge (172-174). Such AMP-like elements could be either an alternative to the CWBDs found in multimodular G+ phage lysins for substrate binding or additional aid to G− lysis due to their ability to interact with membranes (172). However, it has not been yet properly examined how widespread this trait would actually be, and, therefore, its true functional and evolutionary implications are largely unknown. Either as wild-type enzymes co-administered or not with permeabilizing agents or as engineered ones, there already exists a remarkable body of evidence that advocates for the possibility of using enzybiotics also against G− bacteria (**Table 1.3**).

On another subject, AMPs have also long been considered as a possible alternative to antibiotics. AMPs are small proteinaceous molecules with activity against microbial pathogens. The main mechanism by which AMPs have been said to kill microorganisms is membrane disruption, although some other modes of action have also been proposed, such as intracellular components targeting, interference with cell wall synthesis, or displacement of membrane proteins (175). Among some of the cited advantages of AMPs are a slower emergence of resistance, good antibiofilm activity, and immunomodulatory properties (176). Some AMPs are a part of the innate immune response, but the general AMP reservoir is broad, as it includes many sources among all of life's domains, as well as, being proteins, AMPs can be relatively easily designed, tuned, and synthesized. The latest literature implies that AMPs can also be derived from phage lysins, particularly from those of phages that infect G− bacteria (167, 169).

Table 1.3. Some examples of phage lysins experimentally proven as enzybiotics against Gram-negative bacteria. Taken from (83).

Species	Lysin/phage	Susceptible bacteria tested	Methodology used	Acc. No.; comments	References
P. aeruginosa					
	Lys1521/*Bacillus amyloliquefaciens* phage	G–	Activity on intact bacteria	Q94ML9	(162-164)
	EL188/EL	G–	Activity on permeabilized bacteria	CAG27282	(157, 177)
	KZ144/fKZ	G–	Activity on permeabilized bacteria	AAL83045	(177, 178)
	OBPgp279/OBP	G–	Activity on intact bacteria	YP_004958186	(179)
	Art-175	G–	Activity on intact bacteria	Chimera of KZ144 and SMAP-29 peptide	(180, 181)
	LysPA26/JD010	G–	Activity on intact bacteria, biofilm	A0A1V0EFL1	(170)
A. baumannii					
	LysAB2/FAB2	G– and *S. aureus*	Activity on intact bacteria *In vivo*: sepsis (mice)	F1BCP4	(171, 182)
	LysABP-01/ØABP-01	G–	Activity on intact bacteria; synergy with colistin	KF548002	(183)
	PlyAB1/Abp1	*A. baumannii*	Activity on intact bacteria	YP_008058242	(184)
	PlyF307/RL-2015	*A. baumannii*; other G–	Activity on intact bacteria, biofilm *In vivo*: sepsis (mice)	AJG41873	(169, 185)
	LysAB3/*A. baumannii* ATCC 17978 prophage	*A. baumannii*	Activity on intact bacteria	ABO12027	(186)
	LysAB4/*A. baumannii* ATCC 17978 prophage	*A. baumannii*	Activity on intact bacteria	CP000521	(186)
E. coli					
	Lysep3/Ep3	*E. coli*, *P. aeruginosa*	Activity on permeabilized bacteria	A0A088FRS5	(187)
	Lysep3-D8	G–, *Streptococcus* sp.	Activity on intact bacteria	Chimera of Lysep3 and Lys1521 (Q94ML9)	(188)
	Colicin-lysep3	*E. coli*	Activity on intact bacteria *In vivo*: intestinal infection	Chimera of Lysep3 and colicin A (Q47108)	(189)
	EndoT5/T5	*E. coli*	Activity on permeabilized bacteria	Q6QGP7	(190)
	PlyE146/*E. coli* 8.0569 prophage	G–	Activity on intact bacteria	EKK47578	(191)
K. pneumoniae					
	K11gp3.5/K11	G–	Activity on permeabilized bacteria	B3VCZ3	(192)
	KP32gp15/KP32	G–	Activity on permeabilized bacteria	D1L2U8	(192)
	KP27 lysin/KP27	G–	Activity on permeabilized bacteria; cell culture	K7NPX3	(193)
Citrobacter freundii					
	CfP1 lysin/CfP1	*Citrobacter* sp.	Activity on intact bacteria	A0A1B1IXL3	(194)
S. maltophilia					
	P28	G– and some G+	Activity on intact bacteria	Lytic enzyme from a bacteriocin system	(195)
Burkholderia sp.					
	AP3gp15/AP3	G–	Activity on permeabilized bacteria	A0A1S5NV50	(173)

Since surface-related interactions are the fundamental mechanism underlying AMPs activity (**Figure 1.12**), some physicochemical properties that enable AMPs to closely interact with biological membranes are reportedly widespread among them, such as a high cationic charge or high hydrophobic moment (HM)[7] (197). The specific properties of AMPs have allowed proposing data mining or *in silico* optimization strategies to detect or improve AMPs (198, 199). However, although it is safe to say that there is a set of physicochemical properties common among AMPs, it is also now commonly acknowledged that there may be some AMP features not so easily recognizable which can also be crucial for their function (176, 197).

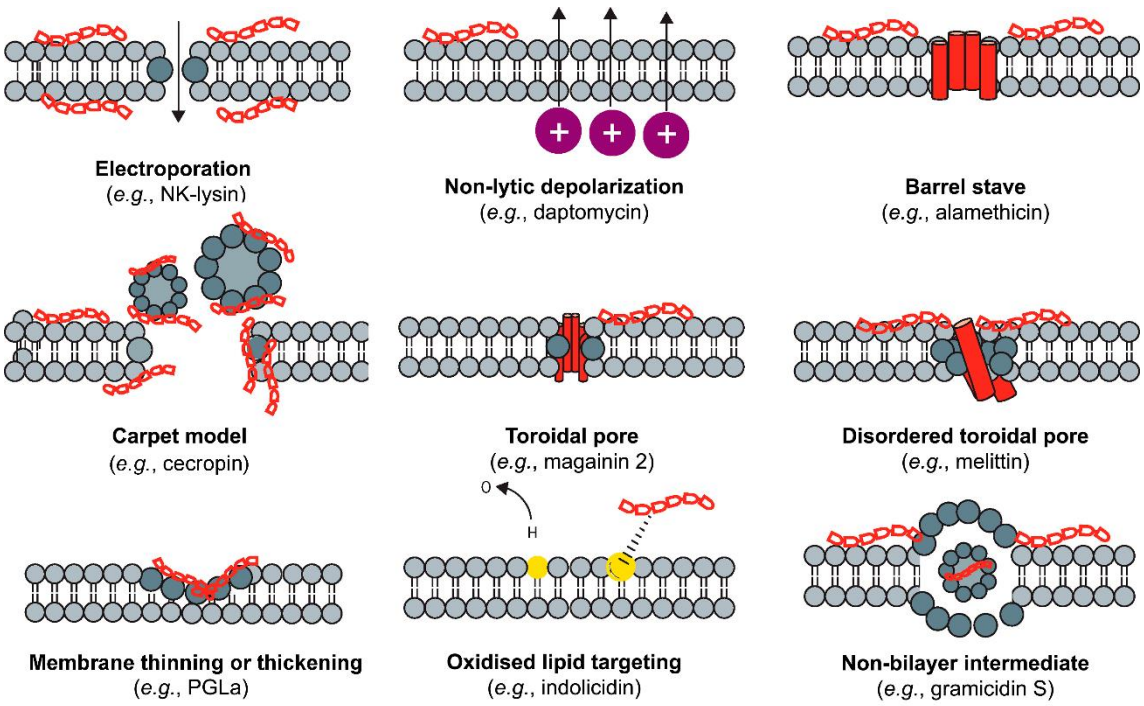


Figure 1.12. Some membrane-related mechanisms of action of AMPs. Adapted from (176).

AMPs are currently used as active ingredients of commercial pesticides (200) and can be safely used as antimicrobials for aquaculture or livestock feeding (201), or as food preservatives (202). However, several factors have hampered the translation of AMPs towards an actual clinical application so far. To begin with, AMP production presents many challenges. Chemical synthesis is rather expensive, and biotechnological production is hampered by the fact that being antimicrobial agents in nature, on many occasions AMPs are not well tolerated by microbial hosts, and thus production is not cost-effective.

[7] The hydrophobic moment is a vector variable proposed by D. Eisenberg that measures the amphiphilic nature of a given 3D peptide conformation (196).

However, recent experiences suggest that biotechnological production can be adapted to actually produce AMPs with high yields at affordable prices, for example by using alternative heterologous hosts or expressing peptides fused to a carrier protein (203, 204). Other causes of failure of AMPs reaching the market for clinical applications are instability of peptide formulations, low *in vivo* half-life mainly due to degradation by proteases, or confounding biological effects on the host immune system (176). Nevertheless, perhaps the major drawback is the systemic or local toxicity effects described for many membrane-disrupting AMPs. Many efforts have been conducted towards reducing toxicity and increasing selectivity against pathogenic bacteria. Some engineering efforts have focused on the AMPs physicochemical properties, and some rational rules have been provided, for example: i) increased net charge enhances antimicrobial activity, but excessive charge can induce hemolytic effects (205); ii) the positioning of positive charges, rather than its number, is what mainly affects the therapeutic index of AMPs (206), *e.g.*, a smaller angle subtended by the positively charged residues improves selectivity against bacteria and enhances permeabilization (207); iii) tuning amphipathicity also has consequences on the balance between antibacterial activity and cytotoxicity — decreasing helicity by introducing D-enantiomers lowers hemolytic activity (208); iv) acetylation of N-terminus or amidation of C-terminus protect AMPs from degradation, although acetylation diminishes net charge, while amidation favors a positive net charge (176). Other strategies rely on the formulation of combined synergistic therapies with currently standard-of-care antibiotics to lower the effective concentration of peptides. Synergy has been proven for a variety of peptide-antibiotic combinations, although the specific mechanisms (or the potential consequences of the AMP-antibiotics interaction for cytotoxicity) have not been clarified. A commonly adduced mechanism of synergy is the increased permeability induced by AMPs that may allow antibiotics with intracellular targets to better reach them (209). Despite potential drawbacks, clinical research on AMPs is currently still underway, with at least three recent AMP-based clinical trials that have successfully entered phase III (NCT03409679, NCT01597505, and NCT01598311). This milestone underlines the feasibility of AMPs-based antimicrobial therapies.

6. THIRD GENERATION ENZYBIOTICS: TOWARDS AN IMPROVED APPLICABILITY

As previously explained (see **Section 4**), the name ‘third generation lysins’ refers to those lysins or lysin-based products or formulations devised to enhance the antimicrobial performance in an *in vivo* clinical setting (149). As such, the fundamental issues to be addressed under clinical conditions can be summarized as: i) immune responses to the

lysin upon systemic administration (to date, some mild inflammatory reactions and the elicitation of slightly lysin-neutralizing antibodies have been described) (89, 96); ii) availability of the active compound at the specific site of infection; iii) *in vivo* half-life, which is rather short (20–60 min) for systemically administered lysins (87, 91), presumably due to protease-mediated proteolysis. As seen in **Table 1.4**, many different strategies have been assayed thus far with a diversity of purposes. Encapsulation and controlled release of lysins have already been explored as means to improve the *in vivo* half-life due to the sustained release of the cargo. With this aim, different nanoscale materials are now being considered as carriers. This interest is prompted also by features such as the high surface area to volume ratio of nanomaterials, and the unique chemical and physical properties of the polymers used to produce them.

Table 1.4. Some examples of third-generation engineered lysins. Adapted from (149).

Lysin/s	Application	Approach	Reference
Cpl-1	Pulmonary	Aerosolization	(96)
Cpl-1	Pulmonary	Loading in chitosan nanoparticles	(210)
Cpl-1, Lysostaphin	Systemic	Dimerisation of lysin to prolong half-life	(211, 212)
$LysK_{CHAP}$	Topical	Co-formulation with lysostaphin in a thermoresponsive PNIPAM gel	(117)
LysK, Lysostaphin	Systemic	Fusion to an albumin binding domain to prolong half-life	(213, 214)
Lysostaphin	Topical	Formulation in chitosan-protein sponges	(215)
Lysostaphin	Musculoskeletal	Formulation in a PEG-hydrogel	(216)
Lysostaphin	Topical	Formulation in an antimicrobial chitosan gel	(217)
Lysostaphin	Topical	Formulation in nano-emulgels	(218)
Lysostaphin	Topical	Coating on liposomes containing standard-of-care antibiotics	(219)
Lysostaphin	Topical	Covalent linkage to cellulose fibers	(220)
Lysostaphin	Catheter	Coating on plastic surfaces (polystyrene and FEP)	(221, 222)
Lysostaphin	Systemic	Glycosylation to prolong half-life	(223)
Lysostaphin, Cpl-1, Hen egg-white lysozyme	Systemic	PEGylation to prolong half-life	(111, 224, 225)
Lysostaphin, JDlys	Topical	Fusion to a peptide to target intracellular *S. aureus*	(226, 227)
PlySs2, SAL-1	Implant/catheter	Immobilisation on self-assembling spider silk proteins	(228, 229)
T4 lysozyme	Topical	Fusion to cellulose-binding module for immobilisation on bandages	(230)
T4 lysozyme, Hen egg-white lysozyme	Topical	Immobilisation on cellulose nanoparticles	(231)

In particular, polymeric nanoparticles (NPs) are quite versatile for engaging in biomimetic modular designs that may emulate specific bacterial traits. A perfect example of these biomimetic designs are dendrimeric NPs bearing choline (or choline analogue) residues, thus resembling the already mentioned architecture of pneumococcal cell wall (232). These choline residues act as binding ligands for the so-called choline binding

proteins (CBPs), a family of surface proteins to which bacterial and phage peptidoglycan hydrolases in the pneumococcal system typically belong. Because of the specificity of such a trait, choline and CBPs have emerged as sources or targets for novel antipneumococcal molecule design (151). CBPs are able to bind pneumococcal choline residues within the cell wall using CBDs, but they also recognize structural analogue molecules, such as diethylaminoethanol (DEAE) (233). This property has been exploited for the design of different antipneumococcal therapeutics based on choline or choline analogues either as free molecules or grafted onto dendrimeric particles (232, 234-236). Through competition with the choline residues in the teichoic acids, choline analogues are able to sequester CBPs and avoid pneumococcal autolysis and daughter cells separation, thus inducing chain formation (234). Moreover, the multivalent effect achieved by cell wall-mimicking dendrimers, containing several copies of choline on their surface, exponentially increases the affinity of the ligand for the CBPs and, therefore, induces chaining in a concentration-efficient manner (232). Such chained pneumococcal cells have been shown to be better removed by phagocytosis, thus posing an alternative, non-lytic approach to tackle pneumococcal infections (235). As a conclusion, there is evidence in the literature to propose a polymeric material containing either choline or choline analogues to specifically bind CBPs, including lytic enzymes bearing CBDs.

There are additional advantages to this kind of strategy when bio-based polymers are employed as a scaffold. For example, chitosan is a β(1,4)-linked linear copolymer of 2-amino-2-deoxy-D-glucopyranose and 2-acetamido-2-deoxy-D-glucopyranose, obtained by *N*-deacetylation of chitin (**Figure 1.13**). Chitin can be easily obtained from shrimp and crab shells, common by-products of the food industry, thus making chitosan preparation a sustainable process in which residues are valorized (237). Turning residues into added-value products, such as biomedical devices, is a must if we are to close the loop of a circular economy, in which waste production is minimized (238). Of note, chitosan bears some properties that make it a good candidate for biomedical and pharmaceutical applications, including biocompatibility, biodegradability, and mucoadhesiveness (239). In the context of respiratory diseases, chitosan has already been used to deliver a variety of drugs into the respiratory tract (240) and, particularly, at least one proof-of-concept of chitosan NPs as a carrier for lysins into the lungs has been reported (210).

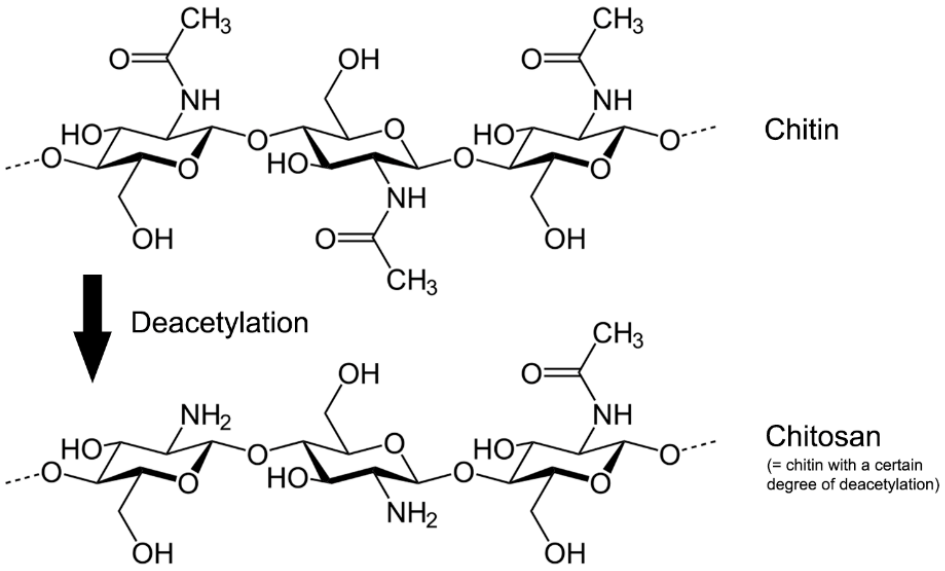


Figure 1.13. Chemical structure of chitin and chitosan.

On the other side, some authors addressed different chemical modifications of chitosan to improve its applicability and antifungal effectiveness, including derivatization with DEAE (241).

II. OBJECTIVES

Antimicrobial resistance is a problem that must be immediately tackled, being the development of new antimicrobials one of the most urgent courses of action. One of the main efforts currently carried out in this regard relies on phages, the natural parasites of bacteria. Lysins are the main phage-encoded molecules directly responsible for bacterial lysis upon infection. Therefore, lysins have been proposed as antibacterial agents when applied exogenously in a purified form, thus also named 'enzybiotics'. While Gram-negative bacteria have been traditionally thought to be refractory to the exogenous action of enzybiotics due to the presence of an outer membrane, there are reports on phage lysins intrinsically active against Gram-negative bacteria from without. This intrinsic activity is suggested to be related to the presence of antimicrobial peptide-like motifs within many of the lysins from phages that infect Gram-negative bacteria. However, the evolutionary reach of this kind of motifs is still largely unknown. Many of the most dangerous bacterial pathogens that are now becoming increasingly drug-resistant are Gram-negative. Such bacteria as *Pseudomonas aeruginosa* or *Acinetobacter baumannii* are responsible for high rates of mortality, mainly due to the establishment of nosocomial respiratory infections difficult to treat with standard-of-care antibiotics. On the other hand, the actual applicability of enzybiotics (and their derivatives) may depend on the development of specific formulations or administration strategies with enhanced clinically relevant properties (such as an increased bioavailability). In this regard, encapsulation and controlled release of enzybiotics is a promising approach to enhance their *in vivo* half-life. Moreover, using sustainable biopolymers, such as chitosan, as a scaffold contributes to waste valorization. Choline-binding domains from the pneumococcal system can be used as an affinity tag to specifically bind choline or choline analogues grafted onto the desired polymer, mimicking the structure and function of the pneumococcal cell wall. Taking into account all of the above stated, the proposed objectives for this thesis were:

1. Constructing a wide database of phage lysin sequences both for analyzing evolutionarily relevant structural patterns and as a tool for screening experimentally interesting lysins.
2. Selecting an enzybiotic candidate intrinsically active against the Gram-negative pathogen *P. aeruginosa* and characterizing its mode-of-action.
3. Designing and characterizing viable antimicrobial agents against *P. aeruginosa* and other relevant respiratory pathogens based on the selected enzybiotic candidate.
4. Developing a chitosan-based biomimetic system for the immobilization of enzybiotics based on the pneumococcal cell wall.

III. MATERIALS AND METHODS

1. BIOINFORMATIC AND STATISTICAL ANALYSES

1.1. Construction of a phage lysin sequences database

A search was conducted in the National Center for Biotechnology Information (NCBI) nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore) for complete genomes from phages infecting a series of bacterial taxa of interest. Such collection was further subjected to a screening protocol looking for genes annotated as lysins. Key terms for the inclusion within the lysin products subset were 'lysin', 'lysozyme', 'murein', 'amidase', 'cell wall hydrolase', 'peptidase' or 'peptidoglycan', while 'structural', 'tail', 'holin', 'baseplate' or 'virion protein' were used as exclusion criteria to avoid misidentifications. Other information contained in the GenBank genomic files (bacterial host, aa sequence, source phage denomination, and protein/genome unique identifiers) were also included in the lysin sequences database.

1.2. Curation of the lysins database

The previously gathered sequences were then subjected to a curation process that involved:

1) A sequence length cut-off based on the aa length distribution of the sequences. Only sequences between 50 and 550 aa (≈95% of sequences) were kept.
2) A sequence identity cut-off to avoid redundant entries. CD-HIT algorithm was used with default parameters and a 98% identity cutoff (242).
3) A functional domains search using PfamScan (243). All sequences for which no relevant hits were found were ruled out. A relevant hit was defined as the prediction of any kind of Pfam (PF) domain whose proposed or proved function would be functional within a phage lysin.
4) Finally, a literature-based careful manual curation was performed, resulting in several re-assignations of bacterial host genera and other amendments.

The complete, curated lysins collection, hereinafter named $\mathbb{S}^{LYS}$ is available at DIGITAL.CSIC (http://hdl.handle.net/10261/221469) (244).

1.3. Sequence-based calculation of physicochemical properties of proteins

The R package 'Peptides' was employed to calculate physicochemical properties based on the lysin sequences in our database (245). Dawson's pKa scale was used for prediction of net charge assuming pH = 7.0 (246); hydrophobicity scale was that proposed by Kyte and Doolittle (247) and average HM was calculated as previously

proposed (248) with a specified rotational angle of 100°, the one recommended for α-helix structures, and a window of 11 aa. Such properties were predicted in the whole sequences, in sequence quartiles (contiguous fragments of sequences that account in length each for a quarter of the whole sequence), or in peptides of 11 aa length to provide either a global vision or more local information.

1.4. Random Forest prediction of the Gram group of bacterial hosts

A random forest model was implemented to check the ability of physicochemical properties to predict lysin sequences as from a G+ or G– infecting phage. R package 'caret' (249) was employed for creating, fitting, and testing the random forest, and further analyses on the model (Receiver Operating Characteristic [ROC] curve, etc.) were performed using packages 'pROC' (250) and 'randomForest' (251). Net charge per residue (NCPR), hydrophobicity, HM, and aliphatic index were the variables included into the model. $\mathbb{S}^{LYS}$ was randomly partitioned in a training subset (75% of all entries) and a testing subset. The training subset was used to fit the random forest parameters (namely, the randomly selected variables for each node, which was fixed in 4) by a 5-fold cross-validation algorithm with 3 repeats. Then the constructed random forest was validated using the previously defined testing subset.

1.5. Construction of a training-testing dataset for AMPs prediction

A peptide sequence dataset of experimentally recognized AMPs and non-AMPs was constructed to generate a simple predictive algorithm able to distinguish between AMPs and non-AMPs. AMPs were obtained from the Antimicrobial Peptide Database (current URL: https://wangapd3.com/; original URL: http://aps.unmc.edu/AP; last time accessed October 2020) (252). A total of 3167 entries were retrieved and, of them, only those under 55 aa were considered for our database (2863 AMPs). Non-AMPs or, more specifically, 'random' peptides were obtained by sampling peptides from an aa sequence randomly generated using RandSeq (253), set up with the average aa composition computed from Swiss-Prot (254). The length of each random peptide was determined by sampling from a simulated normal distribution with the same mean (27.5) and standard deviation (11.0) as those of the length distribution of the cut-off AMPs collection. The final database ($\mathbb{S}^{AMP}$) was constructed by combining 2000 sampled elements from each set (*i.e.*, 2000 AMPs and 2000 non-AMPs). Sequence-based physicochemical properties (NCPR, hydrophobicity, HM, and aliphatic index) were calculated as described in **Section 1.3**. The final dataset can be accessed as Table S1 in the Supplementary Materials

of Vázquez *et al.*, 2021 (255). The previously constructed database $\mathbb{S}^{LYS}$ was used to analyse the differential presence of predicted AMP-like subdomains at the C-terminal end of lysins. For mining membrane-active enzybiotic candidates, a subset of $\mathbb{S}^{LYS}$ named $\mathbb{S}^{PSE}$ was obtained, containing only the aa sequences of the monomodular lysins from phages whose host bacteria were annotated as *Pseudomonas* (80 elements).

A *k*-Nearest Neighbours (*k*NN) algorithm was used to predict AMPs. R package 'caret' was employed for fitting, testing, and applying the *k*NN, based on the random partition of $\mathbb{S}^{AMP}$into a training set ($\mathbb{S}^{TRAIN}$, 75% of elements) and a testing set ($\mathbb{S}^{TEST}$). Normalized NCPR, hydrophobicity, average HM, and aliphatic index were used as the descriptor variables to calculate the distances within the *k*NN model. The *k*NN parameters (namely, the *k* number of neighbours, which was fixed at *k* = 30) were fitted with $\mathbb{S}^{TRAIN}$ by 10-fold cross-validation with 3 repeats. The *k*NN was furthermore evaluated using $\mathbb{S}^{TEST}$ as an 'independent' dataset (whose elements were not considered during training). The goodness of the model was assessed as explained in **Section 1.4**, calculating standard parameters such as sensitivity, specificity, accuracy, or the area under the ROC curve (AUC-ROC). For predicting AMP-like regions at the C-terminal end of the elements in $\mathbb{S}^{LYS}$and $\mathbb{S}^{PSE}$, the aforementioned physicochemical properties were only considered at the region between coordinates 0.75 × length of the sequence and 0.9 × length, according to previous results suggesting that AMP-like regions are not located at the C-terminal apex but rather immediately before (248).

1.6. Data visualization and statistical analysis of computational data

Default methods for data representation implemented in 'ggplot2' R package (256) such as kernel density estimation or Generalized Additive Model (GAM) smoothing were used for visualization of computational data analyses or, when noted, for summarising complex experimental data. Datasets derived from physicochemical computations of lysins sequences were generally deemed non-normal and heteroskedastic according to Shapiro-Wilk normality test and Breusch-Pagan test as performed in R. Thus, for performing comparisons among our data, robust statistical methods were used (257). Specifically, a generalization of Welch's *t*-test with trimmed means (default trimming level γ = 0.20) was used for multiple comparisons, adjusting *p*-values with the Bonferroni method. Effect sizes were calculated according to Wilcox and Tian's ζ (258). A general interpretation for ζ is given in the previous reference, being values of around 0.10 a small effect size, around 0.30 a medium effect and 0.50, and above a large one. All robust methods were used from the implementation in R Package 'WRS2' (259).

1.7. General use software (I): bioinformatic analysis of sequences

A variety of online available programs were used to get information on the genes, proteins, or genomes investigated in this thesis. For the general analysis of DNA sequences or the design of synthetic sequences (chimeric genes, oligonucleotides, etc.) the Geneious suite (Biomatters) was used. Multiple sequence alignments were performed using Clustal Omega as implemented at the EMBL-EBI (243) and visualized using JalView (260). Structural and functional features of aa sequences were probed using BLAST (261), HMMER (262), and PF (263) to detect sequence-similar entries, homologous proteins or to classify potential functional domains within the proteins. Further information was obtained using TMHMM (264), SignalP (265), Phobius (266), LipoP (267), JPred (268), HeliQuest (269), and EMBOSS charge (270) to predict, respectively: transmembrane helices and protein topology, signal peptides, both topology and signal peptides (Phobius), lipoylation or signal peptides, secondary structure, physicochemical properties of aa helices, and local charge plot of a protein sequence. Relevant parameters (molecular weight, isoelectric point, molar extinction coefficient) for protein management (estimation of concentration, etc.) were obtained from the ExPASy ProtParam tool (271). Dali server (272) was used for the pairwise comparison of three-dimensional structures.

Sequence Similarity Networks (SSNs) were sometimes generated for visually assessing the similarity clustering of sequence sets. For this purpose, the Enzyme Similarity Tool from the Enzyme Function Initiative server (EFI-EST) was employed (273). Briefly, this tool performs a local alignment from which every possible pair of sequences receives a score similar to the *E*-value obtained from a typical BLAST analysis. A threshold score value was selected for each SSN so that below such threshold, sequence pairs were considered non-similar and, therefore, the pair would not be connected in the resulting representation. Scores were selected so that sequence pairs whose similarity was below 30-40% were deemed non-similar. The SSN graphs were produced Cytoscape 3 with yFiles organic layout (274).

1.8. General use software (II): analysis and representation of experimental data

Several programs were used to manage experimental data. Spreadsheet software, such as Excel, or statistical language R within the RStudio environment, were used for data treatment and datasets construction. Data were represented either using the R package ‘ggplot2’ or GraphPad 6.0. Statistical analyses were performed either in R or using GraphPad InStat. Unless otherwise stated, quantitative differences between

experimental conditions were analysed to determine whether such differences were statistically significant under the assumption of a signification level α = 0.05 by using ANOVA. The *post hoc* tests of Bonferroni, Tukey, or Dunnett were further used for multiple comparisons.

2. BACTERIAL STRAINS, PLASMIDS, AND OLIGONUCLEOTIDES

Bacterial strains, plasmids, and oligonucleotides used in the course of this thesis are detailed in **Tables 3.1** to **3.3**. Bacterial strains used for susceptibility testing at the S. Albertí laboratory (UIB) are published in: (275-278).

Table 3.1. Bacterial strains used throughout this work.

Species	Strain	Description	Reference or source[a]
Escherichia coli strains for plasmid maintenance and protein expression			
E. coli	DH10B	Strain for plasmid maintenance	CIB
	BL21(DE3)	Strain for protein expression	CIB
Strains for antimicrobials testing			
P. aeruginosa	PAO1	Standard *P. aeruginosa* laboratory strain	CIB
	109.1	Clinical strain	HUB
	126.1	Clinical strain	HUB
	2-006	Clinical strain	HUB
	39.5	Clinical strain	HUB
	57.1	Clinical strain	HUB
	68.1	Clinical strain	HUB
M. catarrhalis	MC-RYC-1	Clinical strain	HRYC
	MC-RYC-2	Clinical strain	HRYC
	MC-RYC-3	Clinical strain	HRYC
A. baumannii	AB-RYC-2	Clinical strain	HRYC
	AB-RYC-3	Clinical strain	HRYC
Acinetobacter pittii	AB-RYC-1	Clinical strain	HRYC
K. pneumoniae	KP-RYC-1	Clinical strain	HRYC
	KP-RYC-2	Clinical strain	HRYC
	KP-RYC-3	Clinical strain	HRYC
E. coli	EC-RYC-1	Clinical strain	HRYC
	EC-RYC-2	Clinical strain	HRYC
	EC-RYC-3	Clinical strain	HRYC
S. aureus	ATCC 12600	Type strain	CECT
S. pyogenes	ATCC 12344	Type strain	CECT
Streptococcus milleri group (SMG)	C5-C20	Clinical strain	UCM
S. pneumoniae	R6	Standard *S. pneumoniae* nonencapsulated laboratory strain	CIB
	D39	Serotype 2 strain, R6 isogenic	CIB

[a] CIB, Centro de Investigaciones Biológicas Margarita Salas; HUB, Hospital Universitario de Bellvitge (Hospitalet de Llobregat), donated by Carmen Ardanuy; HRYC, Hospital Ramón y Cajal (Madrid), donated by Rosa del Campo; CECT, Colección Española de Cultivos Tipo; UCM, Universidad Complutense de Madrid, donated by Juan Miguel Rodríguez.

Table 3.2. Plasmids used throughout this work.

Plasmid	Description	Reference or source
pET-28a(+)	Expression vector. Enables 6×His tag fusion at N-terminal end. KAN^R	Novagen
pUC-PA87	Derived from pUC57. Contains synthetic gene *pae87*. AP^R	GenScript, This work
pET-PA87	Derived from pET-28a(+) and pUCPA87, overexpresses gene *pae87* for production of protein Pae87, fused to a 6×His tag at N-terminal end. KAN^R	This work
pT7-PA87	Derived from pT7-7 and pUC-PA87, overexpresses gene *pae87* for production of protein Pae87 without a 6×His tag. AP^R	This work
pET-PA87-E29	Derived from pET-PA87, overexpresses gene *pae87-e29* which encodes protein Pae87 with mutation E29A. KAN^R	This work
pET-PA87-E46	Derived from pET-PA87, overexpresses gene *pae87-e46* which encodes protein Pae87 with mutation E46A. KAN^R	This work
pET-PA87-E2946	Derived from pET-PA87-E29, overexpresses gene *pae87-e2946* which encodes protein Pae87 with mutations E29A and E46A. KAN^R	This work
pET-PA87-F	Derived from pET-28a(+), overexpresses gene *pae87-f* for production of chimeric protein Pae87-F, composed of Pae87 fused to domain F at C-terminal end. KAN^R	GenScript, This work
pET-PA87-RF	Derived from pET-28a(+), overexpresses gene *pae87-rf* for production of protein Pae87-RF, similar to Pae87-F but with F domain sequence randomized. KAN^R	GenScript, This work
pTRD762	Derived from pT7-7, overexpresses chimeric antipneumococcal enzybiotic Cpl-711. AP^R	(80)
pTRD750	Derived from pT7-7, expresses enzybiotic Cpl-7S. AP^R	(99)
pET-A7S	Derived from pET-29a(+), overexpresses gene *cpl-a7s* for production of chimeric protein Cpl-A7S, composed of Cpl-7S fused to LytA CBD at C-terminal end. KAN^R	GenScript, This work

Table 3.3. Oligonucleotides used throughout this work.

Oligonucleotide	Sequence (5'→ 3')[a]	Annealing site
T7p	TAATACGACTCACTATAGG	T7 promoter (forward)
T7t	GCTAGTTATTGCTCAGCGG	T7 terminator (reverse)
pae87_f	CTAAGGTACCCATATGGCTCTGACCGAGCAAGACTTCC	5' of *pae87* (forward)
pae87_3'	TACAAAGCTTATTTGAAGGATTGATAGGCTTCTGCCAG	3' of *pae87* (reverse)
e29a_f		Triplet coding for E29 of *pae87*, for mutation (forward)
e29a_r		Triplet coding for E29 of *pae87*, for mutation (reverse)
e46a_f	TTCTGTTCGCACGCCACTGG	Triplet coding for E46 of *pae87*, for mutation (forward)
e46a_r	TGGCGTGCGAACAGAATTTTCGG	Triplet coding for E46 of *pae87*, for mutation (reverse)

[a] Restriction enzyme recognition sites are underlined and mutated bases with respect to the wild-type sequence are highlighted in grey.

3. CULTURE CONDITIONS AND MEDIA

All solutions and media were autoclaved at 120 °C and 1 atm for 20 min or filtered sterilized through 0.22 µm pore Millipore filters.

For long-term storage, bacterial strains were frozen in the corresponding culture medium with 10% (v/v) glycerol and kept at −80 °C. All G− bacteria (*P. aeruginosa*, *E. coli*, *M. catarrhalis*, *A. baumannii*, etc.) were routinely grown in Lysogeny Broth (LB) at 37 °C with shaking at 200 rpm. G+ bacteria (*S. aureus*, *S. pyogenes*, SMG) were cultured in

Brain Heart Infusion (BHI) at 37 °C and without shaking. *S. pneumoniae* was grown in C medium adjusted at pH 8.0 (279) supplemented with 0.08% yeast extract (C+Y) in the same conditions as other G+. Culture growth was monitored by turbidimetry at 550-600 nm (OD_{550} or OD_{600}) using a Thermo Scientific Helios Epsilon spectrophotometer.

Solid culture was carried out either on LB plates (LB with 1.5% agar), for G− strains, or in blood agar (trypticase soy agar with 5% defibrinated sheep blood) for G+. Antibiotics as positive selection markers were used at 100 µg/ml (AP) or 5 µg/ml (KAN).

4. DNA TECHNIQUES

The DNA techniques used throughout this paper were essentially those described in (280).

4.1. *E. coli* cells transformation

Competent *E. coli* cells were prepared using the Hanahan rubidium chloride method (281). Transformation of pure plasmid or ligation products into previously prepared competent cells was accomplished by applying a heat shock (37 °C, 3 min) on the ice water pre-cooled cells mixed with the DNA. Heat shock-transformed cells were further incubated for 1 h at 37 °C after adding four volumes of LB medium and were finally plated on LB agar plates containing the appropriate antibiotic for selection.

4.2. Plasmids isolation

Plasmids were isolated from 5-10 ml stationary phase cultures of the corresponding *E. coli* strain bearing the plasmid of interest. NZYMiniprep kit, based on the alkaline lysis protocol, was generally used for plasmid purification according to the manufacturer's instructions (NZYTech).

4.3. DNA gel electrophoresis

Agarose gels (0.7% or 1.5%, w/v) were prepared in TAE buffer (40 mM Tris-HCl, 20 mM acetic acid, 2 mM EDTA, pH 8.1). DNA samples were mixed with 1/4 volume of loading buffer (30% [w/v] Ficoll 400, 0.2% [w/v] bromophenol blue, 0.2% [w/v] xylene cyanol and 40 mM EDTA) and then applied onto the gel. Electrophoreses were run at constant voltage (100 V) during 60-90 min and then gels were stained with GreenSafe (NZYTech). DNA bands were observed under UV light.

DNA electrophoresis was used throughout this work either analytically, for DNA visualization in middle steps of recombinant DNA construction processes (PCR products

detection, transformants/recombinants screening, restriction enzymes digestion checks, etc.); or preparatively, for obtaining enough quantities of purified fragments of DNA. In the latter case, the NZYGelpure kit (NZYTech) was used for DNA purification.

4.4. PCR DNA amplification and mutation

PCR was used for screening of *E. coli* colonies, to test for the presence of recombinant transformants using the appropriate primers (typically T7p and T7t). In this case, colony biomass was used as a template for amplification. Otherwise, plasmid DNA was used as a template (50-100 ng per reaction) for site-directed mutagenesis, which was carried out by overlap extension PCR (see **Section 4.6**). PCR was conducted using either Taq (for screening purposes, TaKaRa) or NZYProof polymerase (NZYTech), for high fidelity amplification in PCR point mutation protocol, according to the manufacturers' instructions. Reaction mix would usually contain 1.5 mM $MgCl_2$, primers at 0.5 µM, and 0.25 mM dNTPs. PCR products were cleaned up with the NZYGelpure kit (NZYTech).

4.5. DNA sequencing

Sequencing was routinely used to check the constructed DNA molecules. DNA sequencing was performed by the Sanger method, using an automated sequencer ABI Prism 3700 (Applied Biosystems) at the sequencing service of the Centro de Investigaciones Biológicas Margarita Salas (Secugen). ddNTPs-terminated amplification reactions were performed in a thermal cycler using BigDye Terminator v3.1 (Applied Biosystems).

4.6. Synthetic genes construction and cloning

The synthetic genes *pae87*, *pae87-f*, *pae87-rf*, *cpl-a7s*, and the DNA fragment encoding domain F were purchased from GeneScript. DNA and aa sequences are available at **Annex 2**. Except for the first one, the rest were provided cloned into expression vector pET28-a(+) between restriction sites NdeI and HindIII, in frame with an N-terminal 6×His tag (**Table 3.2**). Gene *pae87* was provided cloned into a pUC57 vector, and it was afterwards subjected to restriction enzyme digestion (using NdeI and HindIII) and ligation (with T4 DNA ligase) into pET28-a(+).

For the construction of *pae87* mutants (E29A, E46A, and E29A/E46A), an overlap extension PCR protocol was performed. Briefly, plasmid pET-PA87 was used as a template for two separate PCR reactions per mutation. Each of those reactions amplified a fragment of gene *pae87* in such a way that both fragments shared an overlapping section

(≈20 nucleotides [nt]) in which the mutated bases were. A third PCR reaction was then performed using the external, *pae87* flanking primers pae87_f and pae87_3' (**Table 3.3**). Optionally, an intermediate PCR reaction was performed using overlapping fragments as their own 3'-OH priming ends prior to adding the external *pae87* oligonucleotide primers and amplifying the whole mutated gene. Cycling conditions for the PCR followed the recommendations of the manufacturer for NZYProof polymerase. Final PCR products were cleaned up with NZYGelpure kit, digested with NdeI and HindIII, cloned into a pre-digested pET-28a(+) vector, and transformed into *E. coli* DH10B. Colonies were screened by PCR and vectors putatively bearing the mutated gene were then checked by sequencing.

5. PROTEIN TECHNIQUES

5.1. Proteins production

Typically, all proteins were obtained from 1 l cultures of the corresponding producing *E. coli* strain in LB medium containing the corresponding selection antibiotic. Such cultures were inoculated with 10 ml of a stationary phase pre-culture and were incubated up to an OD_{600} ≈ 0.8. Cultures were then induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and further incubated overnight at room temperature (≈ 20 °C). Then biomass was collected by centrifugation (12,000 × *g*, 20 min, 4 °C) and resuspended in an appropriate buffer (binding buffer, see **Section 5.2**) up to 20-30 ml total volume. Cells were disrupted either using a French Press at 8.3 MPa or sonication (five 30 s pulses in a Branson 450 sonifier, interspersed with 30 s incubations on ice). Cell disruption was checked by microscopic observation (see **Section 6.8**), and cell debris was separated by centrifugation (18,000 × *g*, 20 min, 4 °C). Supernatants were collected and DNA was precipitated with streptomycin sulphate (0.6 mg/mg protein as quantified by the Bradford assay [NZYBradford reagent, NZYTech]) at 4 °C and centrifuged again in the same conditions to remove the DNA precipitate. Finally, the clarified supernatant was subjected to the purification process.

5.2. Proteins purification and quantification

Protein purification was accomplished by chromatographical procedures, with additional steps in some cases. Three distinct processes were used:

1) For CBPs, such as Cpl-711 or Cpl-A7S, the collected biomass was resuspended in a binding buffer composed of 20 mM sodium phosphate buffer (NaPiB), pH 6.0, 1 M NaCl. The clarified supernatants containing the protein of interest were

then circulated through a DEAE-Sepharose column (HiTrap DEAE FF 5 ml, GE Healthcare) using an automated liquid chromatography system (ÄKTA Start, GE Healthcare). DEAE behaves in this case as an affinity ligand for CBPs, since it is a structural analogue of choline, the natural ligand of such proteins (233, 282). The column was then washed with the binding buffer until A_{280} baseline was recovered and, afterwards, bound CBPs were eluted with an eluting buffer composed of 20 mM NaPiB, pH 6.0, 1 M NaCl, 2% choline. In this case, choline acts as a competitive ligand for the DEAE-bound CBPs. Eluted purified fractions were dialyzed against storage buffer (20 mM NaPiB, pH 6.0, 100 mM NaCl), and kept at −20 °C.

2) Non-CBP, His-tagged proteins (*e.g.*, Pae87 and all of its derivatives) were purified by immobilized metal ion affinity chromatography using a HisTrap FF 5 ml column (GE Healthcare) loaded with Ni^{2+}. The chromatographic procedure was performed essentially as described in the previous paragraph, but using 20 mM NaPiB, pH 7.4, 0.3 M NaCl, 40 mM imidazole as a binding buffer and 20 mM NaPiB, pH 7.4, 0.3 M NaCl, 0.5 M imidazole as an elution buffer. Proteins were dialyzed against a storage buffer composed of 20 mM NaPiB, pH 7.4, 100 mM NaCl and kept at 4 °C for no more than a month.
3) Finally, non-CBPs nor His-tagged proteins (Cpl-7S, Pae87 as expressed from pT7-PA87) were purified by anion exchange chromatography preceded by ammonium sulphate fractional precipitation, as described in (283). Cpl-7S precipitated at $(NH_4)_2SO_4$ concentration ≈ 20−35% (w/v) and was then dialyzed against 20 mM NaPiB, pH 6.0. Afterwards, the dialyzed protein solution was loaded onto a DEAE-Sepharose column. Then a NaCl gradient was applied up to 0.5 M. Cpl-7S eluted approximately at 0.3 M NaCl, while Pae87 was not retained in the column at such pH. The buffer of the purification fractions was finally exchanged to 20 mM NaPiB, pH 6.0, 100 mM NaCl and kept at −20 °C.

All purified proteins were checked by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE, see **Section 5.3**), and concentration was estimated by A_{280} measurement and applying Lambert-Beer law with predicted molar extinction coefficients and molecular weights displayed in **Table 3.4**.

Table 3.4. Protein parameters as predicted by ProtParam.

Protein	Molecular mass (kDa)	Number of aa	p*I*	Molar extinction coefficient (M^{-1} cm^{-1})
Pae87	23.05	206	9.11	32555
Pae87-His	20.88	186	8.95	32555
E29A	22.99	206	9.24	32555
E46A	22.99	206	9.24	32555
E29A E46A	22.93	206	9.35	32555
Pae87F	27.95	250	9.12	37150
Pae87RF	27.95	250	9.12	37150
Cpl-711	39.12	339	4.63	126865
Cpl-7S	38.58	342	4.86	68425
Cpl-A7S	56.87	505	4.93	149785

5.3. SDS-PAGE

SDS-PAGE was used to monitor protein purification and to check protein integrity in terms of its predicted molecular weight, and it was conducted according to the protocol described by Laemmli (284). 12.5% or 15% SDS-polyacrylamide gels were used as prepared in 100 × 75 × 1 mm plates, and samples were loaded onto them after 10 min treatment at 98 °C in the presence of one volume of 2× Laemmli buffer (125 mM Tris-HCl pH 6.8, 20% [v/v] glycerol, 10% [v/v] β-mercaptoethanol, 0.004% [w/v] bromophenol blue). Then, electrophoresis was conducted at a constant voltage (100-150 V) with the gels immersed into 25 mM Tris-HCl, 192 mM glycine, and 0.1% (w/v) SDS. After the run, gels were stained with BlueSafe (NZYTech) for 15-30 min. Pre-stained NZYColour Protein Marker II from NZYTech was used to estimate the molecular weight of the samples.

5.4. Fluorochrome labelling of proteins

An N-hydroxysuccinimidyl (NHS) ester of Alexa488 fluorochrome (NHS-Alexa488, ThermoFisher) was used to fluorescently label purified proteins. A labelling reaction was set up by incubating the dye together with the protein at a 1:5 (protein:dye) ratio in 20 mM NaPiB, pH 7.4, 100 mM NaCl for 1 h at room temperature. At this pH, N-terminal amino groups would be more reactive than lateral chains amino groups, and therefore the dye would preferentially be incorporated at N-terminal. The reaction was stopped by adding 10% volume of 1 M Tris-HCl, pH 7.4. Free dye was then separated using a HiTrap Desalting FF 5 ml column on an ÄKTA Start liquid chromatography system with 20 mM NaPiB, pH 7.4, 100 mM NaCl as a mobile phase. The much smaller molecule of the dye was longer retained within the column, while the dyed protein eluted earlier. The degree of labelling of each elution sample was calculated by estimating the protein and dye

concentrations measuring, respectively, A_{280} and A_{495} and applying the Lambert-Beer law with the respective molar extinction coefficients. A_{280} was corrected with a factor of 0.11 (corresponding to A_{280}/A_{495} for the dye). Only samples with a labelling degree of ≈ 1 were used for experiments.

5.5. Synthesis and quantification of synthetic peptides

Peptides P87 and P88 were synthesized and provided by GenScript as a freeze-dried powder. They were dissolved in water and concentration was estimated by measuring A_{280}, with a molar extinction coefficient of 5500 M^{-1} cm^{-1}.

6. FUNCTIONAL ASSAYS

6.1. Peptidoglycan purification and muralytic activity assay

A dye-release assay to measure *P. aeruginosa* PAO1 peptidoglycan degradation was set up using Remazol Brilliant Blue (RBB) to label purified peptidoglycan (285, 286). To purify PAO1 peptidoglycan, a culture of at least 1 l of PAO1 in LB was grown up to OD_{600} ≈ 0.8-1.0. Cells were harvested by centrifugation (4,000 × *g*, 15 min, 4 °C) and resuspended in 20 ml phosphate buffer saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) per culture litre. Cells were permeabilized by adding 80 ml/l of culture of 5% SDS and boiling for 30 min with vigorous shaking. Biomass was further incubated in the presence of SDS overnight at room temperature. Next, cell debris was collected by ultracentrifugation (100,000 × *g*, 1 h, 20 °C) and washed with distilled water. The resuspended biomass was subjected to dialysis against distilled water for 24 to 72 h to wash out as much SDS as possible. Then the samples were ultracentrifuged and washed again as many times as necessary to remove all SDS (typically 1-3 times more, SDS was considered to be mostly washed when samples no longer generated a thick foam upon resuspension). Two further purification steps were applied: 1) removal of nucleic acids by treating with RNase and DNase (50 µg/ml each) for 30 min at 37 °C; 2) proteins degradation with 300 µg/ml trypsin overnight at 37 °C. Finally, the purified peptidoglycan was ultracentrifuged again, the supernatant was removed and the pellet was dried for 24-48 h at 37 °C to determine the dry weight yield of the process. Typically, some 12 mg of purified PAO1 sacculi were obtained per litre of culture.

Purified sacculi were dyed by resuspending them in a freshly prepared 0.02 M RBB solution containing 0.2 M NaOH. Incubation in this solution was prolonged for about 6 h at 37 °C with shaking and then overnight at 4 °C. After staining, several ultracentrifugation and washing steps with distilled water were conducted until supernatants were clear

(usually, 3-4 washing steps). The resuspension water volume of the final pellets was adjusted for an $A_{595} \approx 1.5$ of the supernatant of an overnight digestion (see below) of the RBB-sacculi with a known active enzyme (Pae87). In this way, the usual yield of dyed substrate per liter of culture was around 2 ml at 5-6 mg/ml.

For the dye release assay, 100 µl of the RBB stained sacculi were centrifuged (12,000 × *g*, 20 min, room temperature) and the supernatant was discarded. Then, the pelleted sacculi were resuspended in 100 µl of a solution of an appropriate reaction buffer (for Pae87 and its derivatives, 20 mM NaPiB, pH 6.0, 150 mM NaCl) containing the desired concentration of a putatively muralytic enzyme or just buffer for the blank control. Then the samples were incubated for 10 min at 37 °C and reactions were stopped by incubating further 5 min at 95 °C. Samples were then centrifuged (12,000 × *g*, 20 min, room temperature) and the A_{595} of supernatants was measured in a VersaMax multi-well plate spectrophotometer.

6.2. Analysis of degradation products

To analyze the degradation products that resulted from Pae87 activity, a similar protocol to that described in (287) was followed. Data acquisition and analysis were performed with the collaboration of Virginia Rivero-Buceta (Polymer Biotechnology group, CIB-CSIC). Briefly, 100 µl of the *P. aeruginosa* PAO1 peptidoglycan purified as indicated in **Section 6.1** were centrifuged (12,000 × g, 20 min, room temperature) and resuspended again in the same volume of a suitable reaction buffer (20 mM NaPiB, pH 6.5, 100 mM NaCl for Pae87 or 50 mM NaPiB, pH 4.9, for the positive control cellosyl). Then, 10 µg of the corresponding enzyme (or the corresponding volume of water for the negative control) were added and incubated overnight at 37 °C. Reactions were stopped by incubating at 98 °C for 5 min. The tubes then were centrifuged (12,000 × *g*, 20 min, room temperature) and the supernatants were collected. 0.5 M borate buffer, pH 9.0 was added to adjust the samples' pH to 8.5-9.0 and 10 µl of freshly prepared 2 M $NaBH_4$ were added to reduce the sample at room temperature for 30 min. Next, sample pH was adjusted to 2.0-4.0 with 25% orthophosphoric acid.

Then the soluble muropeptides of each sample were separated by reverse-phase high-performance liquid chromatography (RP-HPLC) on a Kinetex C18 Column (1.7 µm, 100 Å, 150 x 2.1 mm, Phenomenex) coupled to an LXQ mass spectrometer equipped with a linear ion trap (Finnigan TM LXQTM, Thermo Scientific). Ionization was achieved by electrospray. Muropeptides were eluted at a flow rate of 0.4 ml/min with the elution gradient: t = 0 min, 95% A; t = 0.5 min, 93% A; t = 3 min, 82% A; t = 11 min, 50% A; t = 12 min, 50% A; t = 12.1 min, 95% A, t = 15 min, 95% A; being A: 0.1% formic acid in

water and B: 0.1% formic acid in 40% acetonitrile. The mass spectrometer was operated on a double play mode in which the instrument was set to acquire a full MS scan (150–2,000 m/z) and a collision-induced dissociation (CID) spectrum on selected ions from the full MS scan. Spectra were analyzed using Xcalibur software (Thermo Scientific).

6.3. Antibacterial activity assays (I): resting cells

Except when noted, all antibacterial assays were performed in 96-well plates. For testing antibacterial activity on resting cells, *i.e.*, in non-growing cells, overnight cultures of the bacterial strains of interest were used to inoculate fresh cultures. Such cultures were grown up to mid-to-late exponential phase as evaluated by turbidimetry. The OD_{600} at which biomass was harvested by centrifugation (3,000 × *g*, 10 min, 4 °C) varied according to the assayed bacteria, ranging between 0.3 and 0.6. The pelleted cells were resuspended in ½ volume of an appropriate buffer and plated onto a 96-well plate (100 µl per well). In parallel, a 'drugs plate' was prepared, adding 100 µl of buffer per well containing the desired concentration of the compounds to be assayed (or just buffer for the control). Then, 100 µl of these drug solutions were transferred to each well of the assay plate, which was then incubated typically at 37 °C for a given time. Turbidity could be monitored during the incubation time in a VersaMax multi-well plate reader. After incubation, samples were taken to plate 10-fold serial dilutions on the appropriate culture media and evaluate the viable cell counts. Unless otherwise stated, distilled water was used as the diluent. Morphology and viability of the treated bacterial cells could also be assayed by microscopy (see **Section 6.8**).

6.4. Antibacterial activity assays (II): growing cells

Antibacterial assays on growing cells were also done in 96-well plates. Each well was filled with 180 µl of a dilution from an overnight culture of the desired bacterium in the appropriate culture medium. OD was then monitored up to early exponential phase and then 20 µl of a solution containing the compounds to test (or just medium/buffer for the control) were added to each well. Growth (OD) was measured for the desired time (usually 16-24 h) at 37 °C.

6.5. Antibacterial activity assays (III): biofilms

P. aeruginosa biofilms were prepared by plating 200 µl/well of a 1:100 dilution of overnight cultures in trypticase soy broth (TSB) medium onto a 96-well polystyrene plate. Biofilm cultures were incubated overnight (16-20 h) at 37 °C without shaking. After

incubation time, OD_{600} was recorded in a VersaMax multi-well plate reader to evaluate total growth. Supernatants were carefully removed from each well and the biomass attached to the plate surface (biofilms) was washed thrice with 250 µl of sterile distilled water per well. For the disaggregation assay, 250 µl of buffer (20 mM NaPiB, pH 6.0, 150 mM NaCl) were added onto each well containing or not the corresponding amount of the antibacterial compound assayed. Incubation was then resumed at 37 °C for 2 h.

For quantifying the antibiofilm activity, after removing the supernatants and washing thrice with sterile distilled water, 200 µl of water and 50 µl of a 1% solution of crystal violet (CV) were added onto each well. Biofilms were stained for 15 min and then the liquid was carefully removed and biofilms were washed thrice with distilled water. Finally, 200 µl of ethanol were added to solubilize the stained biofilms, which were thoroughly scratched and dissolved into the ethanol by pipetting up and down several times. The amount of biofilm was then estimated by measuring A_{595}. Alternatively, some wells were reserved for measuring viable cells, in which case, biofilms were directly dissolved into 200 µl of buffer without prior staining, and 10-fold dilutions of each well were plated onto LB agar.

6.6. Antibacterial activity assays (IV): MICs and synergy

Minimum Inhibitory Concentrations (MICs) of antibiotics or AMPs were determined according to the Clinical and Laboratory Standards Institute (CLSI) guidelines for *P. aeruginosa* antimicrobial susceptibility testing (288). 5×10^5 to 10^6 CFU/ml inocula were plated onto each well of a 96-well plate containing equal volumes of serial 2-fold dilutions of the antimicrobials to be tested in cation-adjusted Mueller-Hinton broth. Plates were then incubated at 37 °C for 20-24 h and then examined both visually and spectrophotometrically (measuring OD_{600} in a microplate reader) to confirm growth or absence of it. When these observations were insufficient to conclude whether there was bacterial growth in any well, *p*-iodonitrotetrazolium was used as a vital stain, adding 20 µl 0.2 mg/ml *p*-iodonitrotetrazolium per 100 µl culture.

Interactions between antibiotics and peptides were tested by the checkerboard assay (86, 289). Such assays were performed in the same microtiter plates using the same medium and incubation conditions as described for MIC determinations. Compounds were usually tested in a range from 1/32 to 2× MIC. The fractional inhibitory concentration (FIC) for each well was calculated ($FIC_A = C_A/MIC_A$; where FIC_A is the fractional inhibitory concentration of compound A, C_A is the concentration of compound A in a given well and MIC_A is the MIC of compound A) and the sum of FIC or FIC index ($FICI = FIC_A + FIC_B$) was used to assess whether the combination was synergistic. Synergy was

defined, as usual, by a FICI ≤ 0.5. Checkerboard assays were complemented by final point viable cell determinations in selected combinations.

6.7. Membrane interaction assays

The physiological effects of the interaction of proteins or peptides with bacterial membranes were measured by using fluorescent probes, respectively, SYTOX Green, for measuring complete permeabilization (OM + inner membrane) or N-phenyl-1-naphthylamine (NPN), to measure OM permeabilization.

SYTOX assays were conducted with suspensions of PAO1 resting cells (≈ 10^7 CFU/ml) prepared from an actively growing culture (OD_{600} ≈ 0.4-0.6) pelleted (3,000 × *g*, 10 min, 20 °C) and resuspended in 20 mM NaPiB, pH 6.0, 150 mM NaCl, 100 mM sorbitol. The sorbitol was added as an additional osmoprotectant to maintain cellular integrity. 100 µl of this suspension were added to each well of a FluoroNunc 96-well plate together with 5 µM of the probe, and fluorescence was recorded in a Varioskan Flash microplate reader (Thermo Scientific) with an excitation wavelength of 485 nm and 520 nm for the fluorescence detection. Incubation was prolonged until baseline was reached and then 100 µl of the compounds assayed at the proper concentration were added to each well. Then, measurements were prolonged for 60 min at 37 °C. NPN assay was similar, but the bacterial suspension was obtained by centrifuging and resuspending in ½ volume of the same buffer a 10^8 CFU/ml culture (OD_{600} ≈ 0.5), and the final probe concentration was 10 µM. The fluorescence recording settings were 350 nm excitation, 420 nm emission. In both cases, non-dyed PAO1 cells treated with just buffer were incorporated as background measurement, and dyed cells treated with buffer as a negative control.

6.8. Microscopy assays

Morphological examination of bacterial cells was carried out with a phase-contrast Leica DM4000B microscope. Routinely, BacLight LIVE/DEAD kit was used to stain treated or control bacterial cells and check cell viability by differential red/green fluorescence in the same microscope equipped with an L5 (excitation bandpass 480-440), an N2.1 (515-560), and an A filter cube (340-380), and viewed either under an HC PL APO 63×/1.40–0.60 or an HC PL APO 100×/1.40 oil objective. The same equipment was used to observe NHS-Alexa488 labelled proteins (see **Section 5.4**). Propidium iodide (PI) was sometimes added at a final concentration of 20 µM for observing membrane-compromised cells. Confocal laser scanning microscopy (CLSM) was used on occasions to obtain better-resolved images. In such cases, samples were observed at a 63×

magnification using a Leica TCS-SP2-AOBS-UV CLSM equipped with an argon-ion laser. Images were analysed using LAS X software from Leica.

7. PHYSICOCHEMICAL AND STRUCTURAL ASSAYS

7.1. Circular dichroism

Circular dichroism spectra were acquired at 4 °C in 20 mM NaPiB, pH 6.5, 150 mM NaCl using a Jasco J700 spectropolarimeter equipped with a temperature-controlled holder. Far UV spectra were recorded from 260 to 200 nm in a 1 mm path length quartz cuvette. Each spectrum was obtained by averaging 5 accumulations collected at a scan rate of 50 nm/min and 2 s of response time. Buffer spectra were subtracted from protein or peptide spectra and molar ellipticity was calculated. Far UV spectra collected with different concentrations of 2,2,2-trifluoroethanol (TFE) were used to predict the ability of peptides to form secondary structures in the presence of biological membranes.

7.2. Protein crystallization and X-ray diffraction-based structural modelling

Crystallization and generation of the three-dimensional model of Pae87 based on experimental X-ray diffraction data were performed by Mateo Seoane-Blanco, at the Mark J. van Raaij laboratory in the Centro Nacional de Biotecnología (CNB-CSIC). Purified Pae87 was initially crystallized in a buffer composed of 20% (w/v) polyethylene glycol 8000, 0.1 M CHES, pH 9.5. The structure of this crystal was solved by molecular replacement using AP3gp15 structural model (PDB 5NM7, 50% identity) as a template. The resulting model had a resolution of 2.50 Å. Next, a second model including Pae87 ligand (PAO1 peptidoglycan solubilized by the enzyme itself before crystallization) was also obtained. In this case, crystallization conditions were 20% (w/v) polyethylene glycol, 5000 monomethyl ether, 0.1 M Tris base-HCl, pH 8.0, 8% ethylene glycol. The second model was solved by molecular replacement with the first one. It had a resolution of 1.22 Å, and thus it was the only one considered for further analysis since it had better resolution and did not substantially differ from the first one (RMSD = 0.4, Z-score = 32.3). The definitive Pae87 structural model presented an R-work/R-free of 0.142/0.176 and a molprobity score (290) of 1.12, in the 97th percentile of a set of structures with the same resolution.

8. NANOPARTICLES PRODUCTION AND ASSAYS

8.1. Chitosan derivatization and nanoparticles production

Chitosan and DEAE-chitosan (ChiDE) nanoparticles (NPs) were produced and provided by Javier Caro-León and María Rosa Aguilar, from the Instituto de Ciencia y Tecnología de Polímeros (ICTP-CSIC, Madrid). ChiDE was obtained by reacting a solution of Cs in hydrochloric acid with DEAE-Cl at pH 8.0 and 65 °C for 2 h. The product was then dialyzed against distilled water, then against NaOH 0.05 M, and finally against distilled water. The ChiDE was then recovered by lyophilization. Chitosan NPs (ChiNPs) and ChiDENPs were obtained by ionotropic gelation mediated by sodium tripolyphosphate (TPP) as a cross-linker agent. For producing DCsNPs loaded with the enzybiotic Cpl-711, equal volumes of Cpl-711 1.5 mg/ml and DCs 2 mg/ml pH 4.0 were mixed together with the cross-linker agent. Standard quality measurements (hydrodynamic diameter, polydispersity index [PDI]) were performed on each batch by dynamic light scattering (DLS).

8.2. Bioassays of NPs with pneumococcal cells

To check for the aggregating ability of NPs against R6 pneumococcal cells, resting cell suspensions were obtained in 20 mM NaPiB, pH 6.5, 150 mM NaCl, roughly as explained in **Section 6.3**. 900 µl of such suspensions were mixed with 100 µl of an NPs suspension at a maximum concentration of 0.5 mg/ml (to reach a final assay concentration of 50 µg/ml) or water as a control, in a 5 ml test tube. Tubes were subsequently incubated at 37 °C for 1 h measuring OD_{550} without shaking or resuspending sedimented cell aggregates. After incubation, samples were obtained for viable cell counts and viable cells observation by fluorescence microscopy (see **Sections 6.3** and **6.8**). Alternatively, the assay was conducted with R6 growing cells using the protocol described in **Section 6.4** with a maximum final concentration of NPs of 50 µg/ml. Samples were taken at mid-stationary phase for microscopical observation.

8.3. CBPs loading onto ChiNPs and ChiDENPs by surface adsorption

Two types of assays were devised to test the ability of NPs to bind CBPs. Firstly, an activity assay on R6 resting cells in which the compounds to test were the 0.5 mg/ml NPs suspensions together with 20 µg/ml of Cpl-711 for 30 min at room temperature filtrated or not through a 0.1 µm pore size syringe filter (Ministart High Flow PES, Sartorius) to remove most of the NPs. Secondly, Alexa488-labelled Cpl-711, Cpl-7S or Cpl-A7S at 20 µg/ml were mixed with NPs at 0.5 mg/ml and then observed under the fluorescence microscope or CLSM (**Section 6.8**).

8.4. *In vitro* release of Cpl-711 from ChiDENPs

The supernatant samples from the *in vitro* release experiment were provided by the ICTP group. Briefly, a sample of freeze-dried ChiDENPs-711 was dispersed in the same volume of PBS at 37 °C with stirring. Every hour, aliquots of the sample were centrifuged (24,000 × *g*, 45 min, 25 °C) and the supernatants were removed and replaced with fresh buffer. The presence of Cpl-711 in such supernatants was tested by SDS-PAGE and to ensure the functional integrity of the released Cpl-711, an activity assay on R6 resting cells was performed as previously described but adding 20 µl of the release sample onto 180 µl of the R6 suspension in 20 mM NaPiB, pH 6.5, 150 mM NaCl.

9. BIOCOMPATIBILITY ASSAYS

Cell culture assays were performed by Carmen Doñoro, from the Cell Cultures service of the Centro de Investigaciones Biológicas Margarita Salas (CIB-CSIC). Cytotoxicity was assessed using A549 adenocarcinoma human alveolar basal epithelial cells. The culture medium was Dulbecco's Modified Eagle's Medium with 10% fetal bovine serum. 80 µl of A549 cells were seeded at a density of 10^4 cells/ml in culture medium in a 96-well culture plate. Samples from the release experiment obtained at 1, 2, and 3 h were used to set up a drug plate in which 1/3 serial dilutions of each sample were prepared in PBS (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl). 20 µl of each dilution and sample were added onto the cells grown for 24 h at 37 °C (accounting for a final dilution of 1/5 of the samples with respect to the drug plate). Culture medium was used as a negative control. Incubation was prolonged for 4, 24, or 48 h at 37 °C. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) was used as an indicator of cell viability adding 20 µl per well from a 2.5 mg/ml stock. Afterwards, plates were incubated in the absence of light for 4 h, and MTT reaction was stopped with 100 µl of 10% SDS, 45% dimethylformamide, pH 5.5 per well. Absorbance at 570 nm and 690 nm was measured using a multiwell plate reader, and the percentage of cell viability was calculated with respect to the 100% signal (OD_{570} – OD_{690}) of the negative control.

IV. RESULTS

CHAPTER 1 − Architectural and physicochemical differences between lysins from phages infecting Gram-positive and Gram-negative bacteria

Results in this section are published as:

Roberto Vázquez*, Ernesto García, Pedro García. 2021. Sequence-function relationships in phage-encoded bacterial cell wall lytic enzymes and their implications for phage-derived products design. J Virol. doi: https://doi.org/10.1128/JVI.00321-21

1. INTRODUCTION

Current literature suggests that specific features, characteristic of lysins from phages infecting G−, are responsible for improved interaction with G− envelope. It could be possible to profit such features (AMP-like elements) for antimicrobial molecules design and, hence, are relevant both from an evolutionary and an applied perspective. Therefore, in this chapter, a wide database of phage lysin sequences was constructed, containing relevant information such as domain architectures, available data of the phages' host bacteria, and sequence-based calculated physicochemical properties. Based on such classifiers, a review on the differential appearance of certain features and their then possible link with an evolutionary function was conducted. In this way, the actual spread of AMP-like elements was uncovered and a data foundation for posterior work was produced.

2. CONSTRUCTION OF A PHAGE LYSINS DATABASE

The lysin sequences for our database (named as $\mathbb{S}^{LYS}$) were obtained from a set of 9,539 genomes retrieved from the NCBI database (accessed on April 2020). Gene annotations were used as criteria for preliminary inclusion into $\mathbb{S}^{LYS}$. So, for example, genes labelled as 'lysin', 'lysozyme', etc., were included, while such terms as 'structural', 'tail', etc., were exclusion criteria. After a curation process that included a length and identity cut-off, functional domains identification, and a literature-based review (for details see **Section 1.2** in MATERIALS AND METHODS), $\mathbb{S}^{LYS}$ final version, available at http://hdl.handle.net/10261/221469, contained 2,182 lysin sequences for which a total of 3,303 PF domains were predicted (**Table 4.1.1**). However, considering all the identical sequences, all the sequences of our dataset corresponded to 36,365 entries in the NCBI Reference Sequence database (RefSeq release 202). Each of our 2,182 sequences was associated with a bacterial genus corresponding to its described bacterial host, for which data on its Gram group and peptidoglycan chemotype was added (for the sake of this work, mycolic acid-containing diderm bacteria such as those belonging to the Corynebacteriales family were considered as G+). Only about 30% of the preliminarily included genes were maintained within $\mathbb{S}^{LYS}$ curation, with varying numbers among the 47 different genera but no apparent bias between G+ and G−. Of all the 2,182 sequences, 770 (35.29%) were assigned to G− bacterial hosts, while 1,412 (64.71%) were lysins from phages described as infecting G+ cells. Up to 45 different PF domain families were found within $\mathbb{S}^{LYS}$ (**Table 4.1.2**).

Table 4.1.1. **Traceability and yield of the curation process for $\mathbb{S}^{LYS}$ construction.**

Host genus	Gram	Outer architecture	Chemo-type (a)	Retrieved phage genomes	Genes detected by annotation	Proteins after length cut-off	Proteins after similarity cut-off	Proteins with at least one relevant PF hit	Proteins after genus reassignation	% Effective genes	PF hits	Average PF hits per protein
Achromobacter	–	Diderm (LPS)	A1γ	36	32	30	17	10	10	31.25%	13	1.30
Acidovorax	–	Diderm (LPS)	A1γ	3	4	4	3	2	2	50.00%	2	1.00
Acinetobacter	–	Diderm (LPS)	A1γ	139	152	138	61	50	50	32.89%	57	1.14
Aeromonas	–	Diderm (LPS)	A1γ	95	77	68	31	26	26	33.77%	27	1.04
Bacteroides	–	Diderm (LPS)	A1γ	5	6	6	4	3	3	50.00%	3	1.00
Burkholderia	–	Diderm (LPS)	A1γ	72	61	58	31	27	27	44.26%	36	1.33
Caulobacter	–	Diderm (LPS)	A1γ	42	58	55	19	17	17	29.31%	17	1.00
Cellulophaga	–	Diderm (LPS)	A1γ	52	38	35	10	9	9	23.68%	9	1.00
Cronobacter	–	Diderm (LPS)	A1γ	45	64	61	32	24	24	37.50%	26	1.08
Enterobacter	–	Diderm (LPS)	A1γ						1		1	1.00
Enterobacteria	–	Diderm (LPS)	A1γ	797	300	271	97	76	3	25.00%	3	1.00
Enterovibrio	–	Diderm (LPS)	A1γ						2		2	1.00
Escherichia	–	Diderm (LPS)	A1γ	684	745	682	173	150	211	20.13%	231	1.09
Fusobacterium	–	Diderm (LPS)	A1γ (b)	4	3	3	3	3	3	100.00%	3	1.00
Klebsiella	–	Diderm (LPS)	A1γ	266	259	241	96	66	67	25.48%	71	1.06
Microcystis	–	Diderm (LPS)	A1γ	8	8	8	6	5	5	62.50%	5	1.00
Pseudomonas	–	Diderm (LPS)	A1γ	625	485	455	132	89	89	18.35%	98	1.10
Ralstonia	–	Diderm (LPS)	A1γ	68	21	21	15	14	14	66.67%	15	1.07
Salmonella	–	Diderm (LPS)	A1γ	414	444	424	123	83	90	18.69%	93	1.03
Serratia	–	Diderm (LPS)	A1γ	31	40	32	20	13	13	32.50%	16	1.23
Shewanella	–	Diderm (LPS)	A1γ						1		1	1.00

Table 4.1.1 (cont.). **Traceability and yield of the curation process for $\mathbb{S}^{LYS}$ construction.**

Host genus	Gram	Outer architecture	Chemo-type	Retrieved phage genomes	Genes detected by annotation	Proteins after length cut-off	Proteins after similarity cut-off	Proteins with at least one relevant PF hit	Proteins after genus reassignation	% Effective genes	PF hits	Average PF hits per protein
Shigella	−	Diderm (LPS)	A1γ						3	-	3	1.00
Sphingobium	−	Diderm (LPS)	A1γ	2	2	2	1	1	1	50.00%	2	2.00
Sphingomonas	−	Diderm (glycosphin-golipid)	A1γ	7	2	2	1	1	1	50.00%	1	1.00
Stenotrophomonas	−	Diderm (LPS)	A1γ	28	25	20	8	6	6	24.00%	8	1.33
Vibrio	−	Diderm (LPS)	A1γ	578	225	216	103	79	76	35.11%	85	1.12
Yersinia	−	Diderm (LPS)	A1γ	70	62	57	18	16	16	25.81%	16	1.00
Arthrobacter	+	Monoderm	A3α (c)	316	400	375	124	103	103	25.75%	161	1.56
Bacillus	+	Monoderm	A1α	348	400	377	170	150	150	37.50%	286	1.91
Bifidobacterium (d)	+	Monoderm	A4α	10	2	2	1	1	1	50.00%	1	1.00
Clostridioides	+	Monoderm	A1γ						23		25	1.09
Clostridium	+	Monoderm	A1γ	109	174	163	67	53	30	30.46%	43	1.43
Corynebacterium	+	Diderm (mycolic acid)	A1γ	40	30	30	16	16	16	53.33%	24	1.50
Cutibacterium	+	Monoderm	A3γ						65		65	1.00
Enterococcus	+	Monoderm	A3α (*E. faecalis*, e) A4α (*E. faecium*)	127	150	150	64	51	50	34.00%	77	1.54
Lacticaseibacillus	+	Monoderm	A4α						13		28	2.15
Lactiplantibacillus	+	Monoderm	A4α						18		42	2.33
Lactobacillus	+	Monoderm	A4α	105	133	130	83	68	22	51.88%	41	1.86
Lactococcus	+	Monoderm	A3α	388	376	375	128	120	121	31.91%	144	1.19
Leuconostoc	+	Monoderm	A3α	35	45	45	17	11	11	24.44%	16	1.45
Levilactobacillus	+	Monoderm	A4α						9		14	1.56
Limosilactobacillus	+	Monoderm	A4α						6		14	2.33

Table 4.1.1 (cont.). Traceability and yield of the curation process for $\mathbb{S}^{LYS}$ construction.

Host genus	Gram	Outer architecture	Chemo-type	Retrieved phage genomes	Genes detected by annotation	Proteins after length cut-off	Proteins after similarity cut-off	Proteins with at least one relevant PF hit	Proteins after genus reassignation	% Effective Genes (f)	PF hits	Average PF hits per protein
Listeria	+	Monoderm	A1γ	75	87	72	19	17	17	19.54%	32	1.88
Mycobacterium	+	Diderm (mycolic acid)	A1γ	1990	3596	3456	584	299	299	8.31%	374	1.25
Propionibacterium	+	Monoderm	A3γ	202	214	214	89	70	5	32.71%	8	1.60
Rhodococcus	+	Diderm (mycolic acid)	A1γ	70	83	80	37	31	31	37.35%	70	2.26
Staphylococcus	+	Monoderm	A3α	349	456	377	93	79	79	17.32%	155	1.96
Streptococcus	+	Monoderm	A3α	1064	739	734	361	246	246	33.29%	688	2.80
Streptomyces	+	Monoderm	A3γ	240	208	207	104	97	97	46.63%	151	1.56
			Total	9539	10206	9676	2962	2182	2182	28.22%	3303	1.51
			Gram-negative	4071	3113	2889	1005	770	770	27.37%	844	1.1
			Gram-positive	5468	7093	6787	1957	1412	1412	28.68%	2459	1.74

(a) Chemotype assignations were based on (63).
(b) Contains lanthionine instead of m-DAP.
(c) Although there seems to be some variety in peptidoglycan structure among *Arthrobacter*, the most common chemotype was assigned.
(d) There are several chemotypes described within *Bifidobacterium*, but A4α was assigned to the only *Bifidobacterium* host entry in our database (*Bifidobacterium thermophilus*).
(e) Phage lysins with unassigned *Enterococcus* host species were assigned *E. faecalis* chemotype Aα3.
(f) % Effective genes is the proportion of the initially retrieved genes that were kept after all curation steps (*i.e.*, column 10 / column 6).
In the lowest three rows, summary numbers are shown: either the sum of each column values or weighted averages, in the case of the percentage of effective genes and PF hits per protein.

A literature review was conducted to assign a function to each of those domains. Most of them were deemed EADs (28 examples) while there were 15 types of CWBDs and 2 structural domains (that would probably be a part of virion-associated lysins). It is fair to warn that the screening method was biased by the annotation of the source genomes and by the current knowledge about the functional domains present in lysins (*i.e.*, by the present set of PF-identified lysin functional domains). Therefore, while the final curated data set will be useful to produce a landscape view of our current understanding of phage lysins, generalizations should be taken with care, since future works may clarify some still unknown parcels of lysins composition and function.

3. GENERAL DIFFERENCES AMONG G− AND G+ LYSINS

For most of the sequences within $\mathbb{S}^{LYS}$ (1,512; 69.3%), only one PF hit was predicted (**Figure 4.1.1**). However, as hinted before, there is an important difference between lysins from G+ infecting phages, which contained a single predicted domain in nearly 60% of the cases, and those from G−, with 90.6% of a single predicted domain (**Figure 4.1.1A**). On the other hand, just a few lysins contained four or more PF hits. However, all these figures do not necessarily reflect the actual structure of the corresponding proteins since: 1) there are domains that are made up of several repeats (up to 5), which are recognized as individual PF hits artificially inflating the 'number of domains'; 2) some sequences for which a single domain has been predicted may contain additional, still undescribed domains; or 3) some repeats or even full domains may not be appropriately predicted when there is enough evolutionary sequence divergence from the family model (as an example, the domain structure based on the 3D folding of pneumococcal autolysin LytA (291) shows six repeats at the CWBD, while a homology search only predicts five of them).

Consequently, to assess whether the G− lysins actually seem to be monomodular while for G+ the case is contrary, lysin lengths were examined. Lysins from G+ infecting phages were generally larger (distribution median = 317 aa) than those from G− infecting ones (median = 164 aa, **Figure 4.1.1C**). Complementing this observation, as shown in **Figure 4.1.1D**, EADs in G− lysins started approximately at the same point as those of G+ (*i.e.*, at the N-terminal end), but with a slight shift towards C-terminal in the case of G−. This small shift could be explained by the fact that some G− lysins have been described to have a CWBD at N-terminal, a case that rarely (if ever) has been observed for G+ ones (172).

Table 4.1.2. PF families detected within $\mathbb{S}^{LYS}$ and their proposed function.

Name	PF	Interpro	Domain Type	Reference	Activity	Comments
Amidase02_C	PF12123	IPR021976	CWBD	(292, 293)	Cell wall binding	Experimentally confirmed binding to *Bacillus* vegetative cells
Big_2	PF02368	IPR003343	CWBD	(294)	Cell wall binding	Similar to lectins; possibly binds carbohydrates
CW_7	PF08230	IPR013168	CWBD	(152, 283)	Cell wall binding	Shown to bind the *N*-acetyl-D-glucosaminyl-(β1,4)-*N*-acetylmuramyl-L-alanyl-D-isoglutamine moiety of peptidoglycan
CW_binding_1	PF01473	IPR018337	CWBD	(295)	Cell wall binding	Known to bind choline residues at the teichoic acids of some Mitis group streptococci
CW_binding_2	PF04122	IPR007253	CWBD	(296)	Cell wall binding	Clostridial cell surface motif; binds to surface glycans (specifically to highly conserved capsular polysaccharide PSII)
LGFP	PF08310	IPR013207	CWBD	(297)	Cell wall binding	Found in proteins from mycobacteria, next to esterase domains. Also found at C-terminal of glycoside hydrolase domains. Probable binding domain
LysM	PF01476	IPR018392	CWBD	(298)	Cell wall binding	Recognizes polysaccharides containing NAG residues, including peptidoglycan
PG_binding_1	PF01471	IPR002477	CWBD	(151)	Cell wall binding	Appears at N- or C-terminal of enzymes involved in cell wall degradation. Presumed to bind peptidoglycan
PG_binding_3	PF09374	IPR018537	CWBD	(299)	Cell wall binding	Potential peptidoglycan binding
PSA_CBD	PF18341	IPR041341	CWBD	(300, 301)	Cell wall binding	Binding to cell wall teichoic acids. Determines specific binding to different *Listeria* serovars
SH3_3	PF08239	IPR003646	CWBD		Cell wall binding	Putative binding domain
SH3_5	PF08460	IPR003646	CWBD	(302-304)	Cell wall binding	Peptidoglycan binding (in staphylococcal phages, full specificity determined by the presence of pentaglycine bridge; in *Lactobacillus* it has been suggested to bind peptidoglycan stressing participation of glucosamine and low specificity)
SPOR	PF05036	IPR007730	CWBD	(305, 306)	Cell wall binding	Peptidoglycan binding domain found in proteins involved in sporulation and cell division. Displays preferent binding to septal peptidoglycan lacking stem peptides
ZoocinA_TRD	PF16775	IPR031898	CWBD	(307)	Cell wall binding	Proposed to recognize A3α peptidoglycan, perhaps only with 2× alanine interpeptide bridges
DUF3597	PF12200	IPR022016	CWBD	(301)	Cell wall binding	Probable binding domain. Similar to *Listeria* phage lysin Ply118 CWBD responsible for specific binding to cell wall elements. Found at C-terminal of hydrolytic domains
3D	PF06725	IPR010611	EAD	(308)	Lytic transglycosylase	
Amidase_2	PF01510	IPR002502	EAD	(309, 310)	NAM-amidase	
Amidase_3	PF01520	IPR002508	EAD	(311)	NAM-amidase	
Amidase_5	PF05382	IPR0002508	EAD	(312)	NAM-amidase	
CHAP	PF05257	IPR007921	EAD	(313-316)	Amidase or peptidase	It has been described to have either an amidase activity or an endopeptidase activity (cutting between D-Ala and the poly-Gly interpeptide bridge of staphylococci)
Cutinase	PF01083	IPR000675	EAD	(317, 318)	Esterase	Present in fungal enzymes that detach acetyl side groups from polysaccharides
FSH1	PF03959	IPR005645	EAD	(319)	Esterase	
Glucosaminidase	PF01832	IPR002901	EAD	(320)	Glucosaminidase	
Glyco_hydro_108	PF05838	IPR008565	EAD	(321)	Muramidase	
Glyco_hydro_19	PF00182	IPR016283	EAD	(322)	Glucosaminidase (or muramidase)	Described as chitinase, probable glucosaminidase activity (or muramidase). Their only recognized function is as chitinases, which is hydrolysing bonds between NAG residues. Since such homopolymer does not exist amongst bacteria, it is speculated that *Glyco_hydro_19* in phage lysins has a glycosidase, murolytic activity.
Glyco_hydro_25	PF01183	IPR002053	EAD	(323)	Muramidase	
GPW_gp25	PF04965	IPR007048	EAD	(324)	Muramidase	Virion associated lysin domain

Table 4.1.2 (cont.). PF families detected within $\mathbb{S}^{LYS}$ and their proposed function.

Name	Pfam	Interpro	Domain Type	Reference	Activity	Notes
Hydrolase_2	PF07486	IPR011105	EAD	(325)	Lytic transglyco-sylase	
Muramidase	PF11860	IPR024408	EAD	(326)	Muramidase	
NLPC_P60	PF00877	IPR000064	EAD	(327, 328)	Peptidase	D-Glu-m-DAP endopeptidase
Peptidase_C39_2	PF13529	IPR039564	EAD	(73)	Peptidase	D-Glu-m-DAP endopeptidase
Peptidase_C93	PF06035	IPR010319	EAD	(329)	Peptidase	
Peptidase_M15_3	PF08291	IPR013230	EAD	(330)	Peptidase	
Peptidase_M15_4	PF13539	IPR039561	EAD	(73, 331)	Peptidase	D-Ala-m-DAP/m-DAP-m-DAP/L-Ala-D-Glu endopeptidase
Peptidase_M23	PF01551	IPR016047	EAD	(332-334)	Peptidase	Peptidase (cuts at Gly-Gly interpeptide bridge or L-Ala-D-Glu)
Pesticin	PF16754	IPR031922	EAD	(335)	Muramidase	
Phage_lysozyme	PF00959	IPR002196	EAD	(336)	Muramidase	
Phage_lysozyme2	PF18013	IPR041219	EAD	(337)	Muramidase	Virion associated lysin domain
Prok-JAB	PF14464	IPR028090	EAD	(338)	Peptidase	
SLT	PF01464	IPR008258	EAD	(339)	Lytic transglyco-sylase	
Transglycosylase	PF06737	IPR010618	EAD		Lytic transglyco-sylase	
PE-PPE	PF08237	IPR013228	EAD	(340)	Esterase	C-terminal to mycobacterial PE and PPE proteins, possible esterase based on active site homology
Prophage_tail	PF06605	IPR010572	EAD		Peptidase	Virion associated endopeptidase domain
Gp5_C	PF06715	IPR010609	Structural		Structural	
Gp5_OB	PF06714	IPR009590	Structural		Structural	

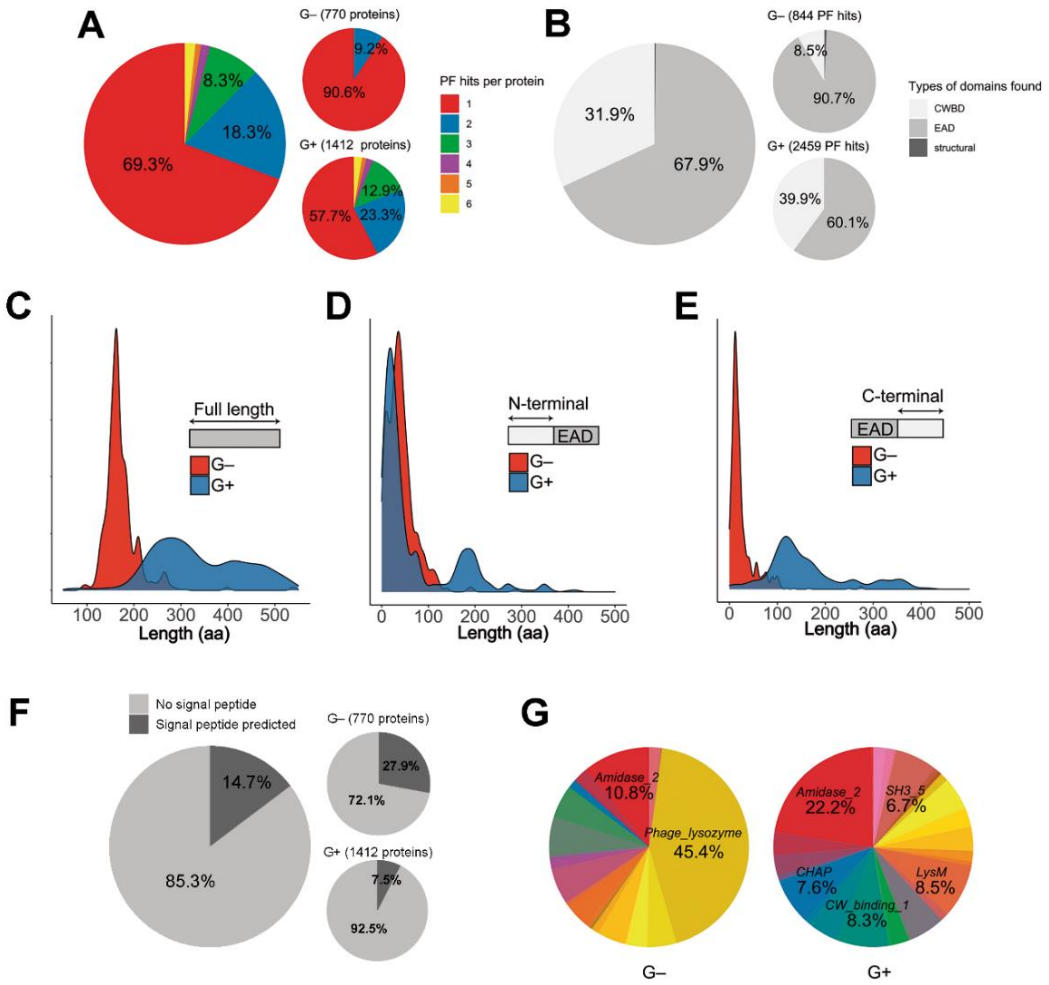


Figure 4.1.1. General properties of lysins from phages that infect G+ or G− bacteria. A: distribution of the number of PF hits predicted per protein. **B**: distribution of domain types. **C**: distribution of protein lengths. **D** and **E**: distributions of the number of aa before (**D**) or after (**E**) predicted EADs. **F**: Percentage of lysins with a predicted N-terminal signal peptide according to Phobius. **G**: PF domains variability (different colours stand for different PF domain families, corresponding to those shown in **Table 4.1.3**). In distribution charts (**B**, **C**, **D**) Y-axis shows an estimation of the distribution density.

Also, as shown in **Figure 4.1.1F**, there might be a high representation of SAR-lysins among G– phage lysins. The latter panel shows that a signal peptide was predicted by the Phobius server for 27.9% (215 lysins) of the G– phage lysins, while such elements were only predicted for a 7.5% (106 lysins) of the G+ ones. The prediction of a signal peptide is indicative of a SAR mechanism of lysins export into the periplasm (engaging the Sec secretion system), although signal peptide prediction algorithms may overlook some true SAR-signals mainly due to them lacking a cleavage site. Of note, **Figure 4.1.1D** also shows a secondary local maximum at around 200 aa, consistent with the not uncommon presence of CWBDs at a medial position (and, subsequently, the appearance of EADs immediately after the medial CWBD), or with a second, medial EAD predicted after an N-terminal one. Finally, **Figure 4.1.1E** illustrates that for most G− lysins there was little ‘additional space’ at the C-terminal end after the predicted EAD (in contrast with

G+, where there was a wide range of C-terminal lengths well above 100 aa). Hence, all these results strongly support the commonly sustained hypothesis that most lysins from phages that infect G− are monomodular, bearing a single EAD, while those from G+ are more variable in architecture, typically containing one or two EADs and a CWBD. Indeed, the percentage of CWBDs among the PF hits predicted in the G− group was much lower (8.5%) than that of G+ (39.9%, **Figure 4.1.1B**). Also, G+ presented a high diversity of different types of domains in contrast with the relatively lower diversity among the hereby analysed G− (**Figure 4.1.1G**). It is worth noting the high prevalence of *Phage_lysozyme* EADs in G−, which makes up to 45.4% of the total PF hits of G−, while *Amidase_2,* the most frequent EAD among G+, accounted only for 22.2% of total hits.

4. DIFFERENCES IN DOMAIN ARCHITECTURES ACROSS BACTERIAL HOST GROUPS

4.1. Domains distribution between G+ and G−

When checking the differential prevalence among G+ or G− of the different PF domain families detected within $\mathbb{S}^{LYS}$, it was noted that most domain types were preferred either in G+ or in G− (**Figure 4.1.2**). In other words, domain families that appear indistinctly in lysins from phages that infect G+ or G− seem scarce. Among EADs, *Amidase_5*, *Glyco_hydro_25*, *Peptidase_C39_2*, and *Transglycosylase* appeared exclusively in lysins from phages that infect G+. Other EADs such as *Amidase_2*, *Amidase_3*, *CHAP*, *Glucosaminidase*, *Peptidase_M15_4*, and *Peptidase_M23* were common in G+ but appeared also among G−.

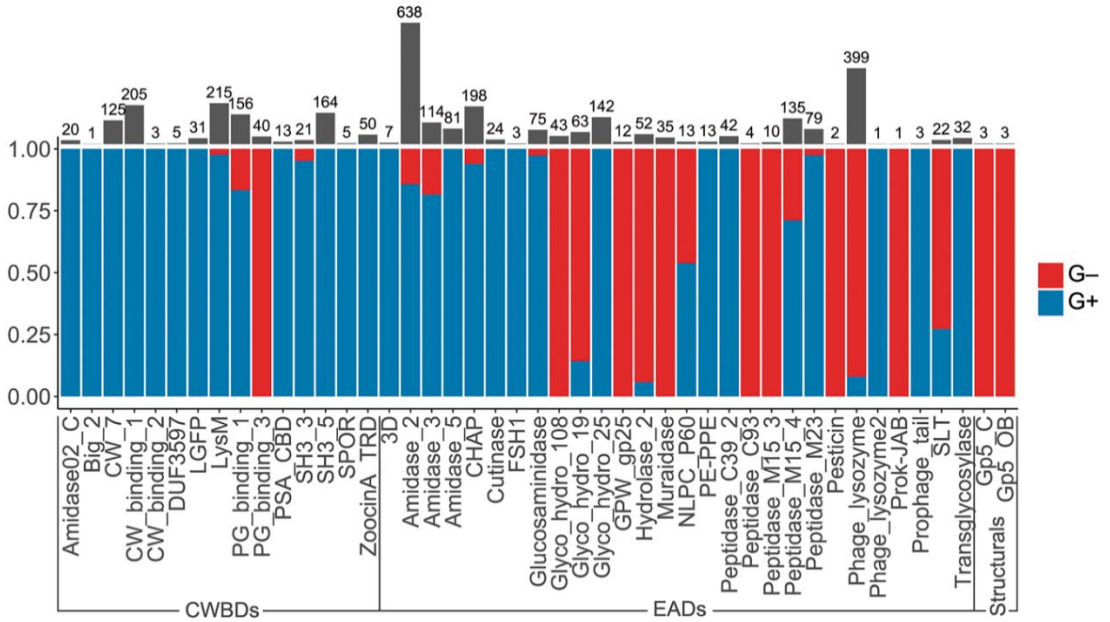


Figure 4.1.2. Differential distribution of PF hits among G− and G+ bacterial hosts. Y-axis shows the proportion of PF hits found in G+ within a given domain family. Grey bars and numbers above represent the total number of hits of each PF domain in $\mathbb{S}^{LYS}$.

The distribution was reversed for *Glyco_hydro_19*, *Hydrolase_2*, and *Phage_lysozyme*, predominant in G−. Only the EADs *Muramidase* and *Glyco_hydro_108* appeared exclusively in G− lysins, the latter often accompanied by the only G− exclusive CWBD in $\mathbb{S}^{LYS}$, *PG_binding_3*, making up the characteristic architecture [*Glyco_hydro_108*][*PG_binding_3*]. As for the other CWBDs, they were almost only predicted in lysins from G+ bacteria phages (some domains, such as *LysM*, *PG_binding_1* or *SH3_3* were also present in those from G− but not frequently).

4.2. Domains distribution across genera

There were also some trends uncovered in the PF hits distribution among host cell genera (**Figure 4.1.3**, **Figure 4.1.4**, **Figure 4.1.5**). We will review several genera of interest with enough representation in the database:

1) *Streptomyces* phage lysins seemed to prefer amidase EADs and, particularly, *Amidase_2* or *CHAP* (although *CHAP* domains have been reported to have a peptidase activity on occasions) (313, 314). CWBDs present were those from *CW_7*, *LysM*, or *PG_binding_1* families. It is worth noting that most of the *CW_7* occurrences in *Streptomyces* lysins have a predicted architecture [unknown][*CW_7*], in which the unknown N-terminal sections are quite similar between them (the lowest identity detected was 60%). This points out to a yet undescribed N-terminal EAD within such proteins as QGJ96544.1, QAY17123.1, or YP_009615350.1.
2) *Streptococcus* genus presents a wide variety of EADs and CWBDs. This may be due to the clinical relevance of several species within *Streptococcus*, which has led to an important research effort focuses on phages and lytic enzymes within this genus (it is, in fact, the genus with the higher number of lysin sequences in our data set, up to 688, with the second one, *Mycobacterium*, having just 374). This, in turn, may have resulted in a deeper characterization of the structure of the lysins within this group. There are, however, quite important architectural differences between lysins from phages that infect different streptococcal species that may be worth comment (**Figure 4.1.4**):
    a) Among the Mitis streptococci (*S. pneumoniae*, *Streptococcus mitis*, *Streptococcus oralis*), one of the most remarkable features is that all occurrences of *CW_binding_1* were in this group. This is easily explainable by the fact that these repeats have been shown to bind choline residues at the teichoic acids of the Mitis group cell wall (295). This confers to the *CW_binding_1*-bearing lysins of Mitis group a very high specificity against bacteria belonging to this

group. The most common architecture (**Figure 4.1.5**) resembles the major pneumococcal autolysin LytA ([*Amidase_2*] 5×[*CW_binding_1*]). Other common structures featuring *CW_binding_1* are [*Glyco_hydro_25*] 5×[*CW_binding_1*], [*Amidase_5*] 4×[*CW_binding_1*] or [*CHAP*] 5×[*CW_binding_1*]. Although the most common CWBD was *CW_binding_1* by far, some other CWBDs were also found, such as *SH3_5* or *CW_7*, which, incidentally, also appear to be quite spread among the other pseudotaxonomical groups of *Streptococcus*.

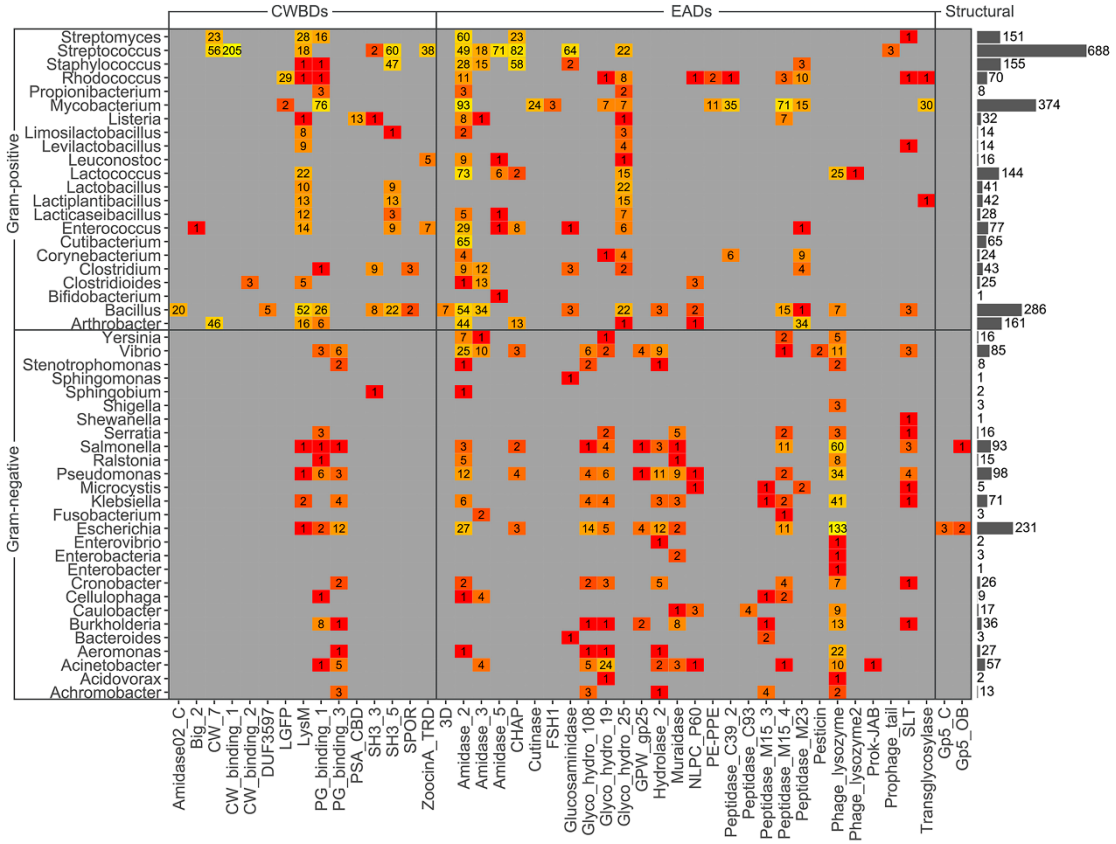


Figure 4.1.3. Heatmap of PF hits distribution across host bacterium genera. Numbers within each tile indicate the number of hits predicted for the corresponding taxon and PF family. The colour scale represents the number of hits from low (red) to high number (yellow). Grey bars at the right represent the total number of PF hits predicted within each genus.

b) It was common to find trimodular lysins among *Streptococcus suis* phages, bearing, apparently, two distinct EADs. The most common architecture was [*Amidase_5*] 2×[*CW_7*] [*Glucosaminidase*], together with [*Amidase_3*] [*LysM*] [*CHAP*]. Some other bimodular architectures with *SH3_5* were also found.

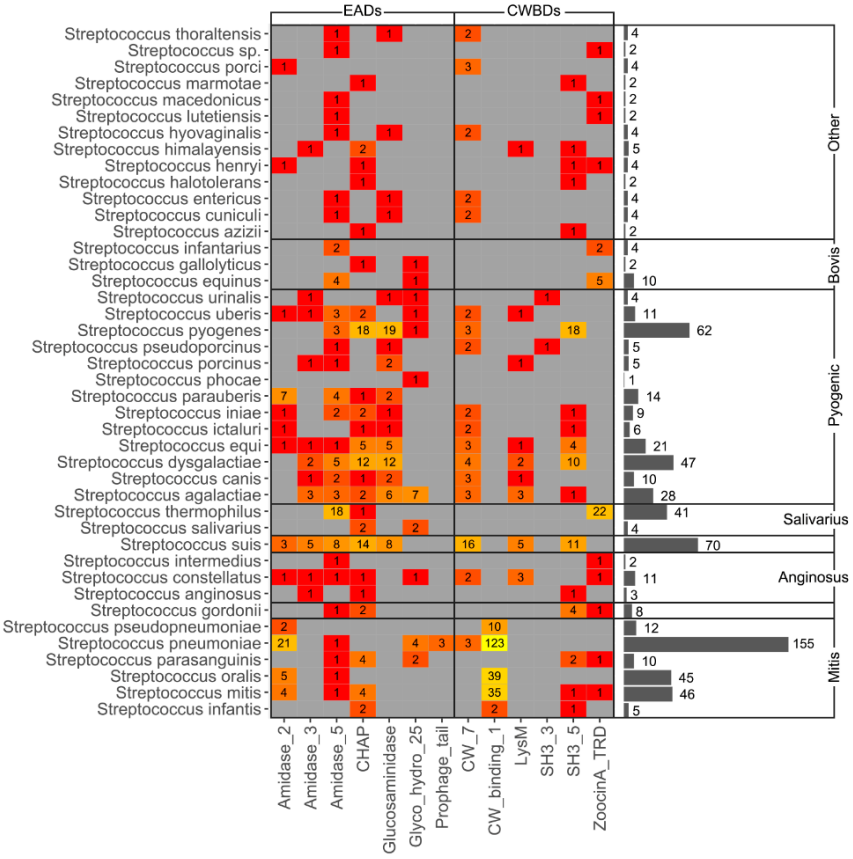


Figure 4.1.4. Heatmap of PF hits distribution. Numbers within each tile indicate the number of hits predicted for the corresponding taxon and PF family. The colour scale represents the number of hits from low (red) to high number (yellow). Grey bars at the right represent the total number of PF hits predicted within each genus.

c) The *Streptococcus thermophilus* phage lysins present in $\mathbb{S}^{LYS}$ contained *ZoocinA_TRD* as a CWBD, bearing the typical architecture [*Amidase_5*] [*ZoocinA_TRD*].

d) In the case of the Pyogenic group phage lysins, there was a certain variability. The trend, however, was towards trimodular, bicatalytic examples such as in *S. suis*, with the [*Glucosaminidase*] [*CHAP*] [*SH3_5*] architecture standing out. The same trimodular architectures that appeared in *S. suis* were also present among the Pyogenic group phage lysins.

3) Staphylococcal phage lysins presented less EADs variability, normally containing *Amidase_2*, *Amidase_3*, and/or *CHAP* domains, and *SHB3_5* seems to be the undisputedly preferred CWBD, as previously demonstrated (341). Staphylococcal *SH3_5* has been shown to bind the characteristic peptidoglycan type bearing a pentaglycine interpeptidic bridge (302). Common architectures were trimodular ([*CHAP*] [*Amidase_3*] [*SH3_5*], [*CHAP*] [*Amidase_2*] [*SH3_5*] or

[*Peptidase_M23*] [*Amidase_2*] [*SH3_5*]), but there were also some bimodular examples ([*CHAP*] [*SH3_5*]).

4) Most monocatalytic lysins from *Rhodococcus* phages contain an *Amidase_2* domain, either accompanied by *LGFP* repeats (a probable binding domain for mycolyl-arabinogalactan containing bacteria, see **Table 4.1.2**) or by an unknown domain (as in *YP_005087066.1*, *YP_009013817.1*, *YP_009017624.1*). There were also some bicatalytic architectures, mostly [*Peptidase_M23*] [*Glyco_hydro_25*] [unknown], in which the unknown C-terminal domain contained around 100 aa (see for example AWY04393.1, AWY04019.1, or AQP30934.1).
5) Among *Propionibacterium* and *Cutibacterium* phages, there was also a wide predominance of *Amidase_2* as an EAD, either accompanied by *PG_binding_1* or by a ≈110 aa unknown domain (*e.g.* YP_009151220.1, AGI12651.1, or YP_009148291.1).
6) *Mycobacterium*, or Corynebacteriales in general including thus *Corynebacterium* and *Rhodococcus* as well, are a very particular group regarding cell wall structure. Their cell wall is diderm and thus contains an 'outer membrane' composed of a variety of lipids among which outstand mycolic acids. Some of those mycolic acids are esterified with arabinogalactan, which is also covalently bound to the peptidoglycan (see **Figure 1.7**). To sufficiently disturb the micolyl-arabinogalactan-peptidoglycan complex and allow bacterial lysis, mycobacteriophages usually contain a pair of catalytically distinct lytic proteins, namely lysin A (a canonical endolysin that degrades peptidoglycan) and lysin B (a lipolytic enzyme that disrupts the mycolic acid layer by detaching it from the arabinogalactan). This is why PF domains putatively assigned an esterase activity (*Cutinase*, *FSH1*, *PE-PPE*, **Table 4.1.2**) were only predicted in *Mycobacterium* or *Rhodococcus* phage lysins, as proteins annotated as type B lysins. Among mycobacteriophages in $\mathbb{S}^{LYS}$, such lysins B normally contained one of the putative esterase domains and sometimes a single *PG_binding_1* repeat located at N-terminal as a CWBD. There are some type B lysins for which only a significant N-terminal CWBD (of *PG_binding_1* family) is found, but, at the C-terminus, there is still room for a catalytic domain. In some cases (*e.g.*: YP_009125503.1, YP_007869957.1), this C-terminal sequence resembles other esterase domains according to a HMMER homology search. There are also lysins B with a catalytic C-terminal domain and a non-assigned N-terminal end. Altogether this suggests the presence of CWBDs or EADs not present within the current PF framework. *PG_binding_1* also seemed to be the preferred CWBD for mycobacteriophage lysins A. EADs were, as it has been already pointed out (73), quite diverse. Mycobacteria stand out among other

bacteria hereby classified as G+ because of the variety of confirmed peptidase activities among their phage lysins (*Peptidase_C39*, *Peptidase_M15_4*, *Peptidase_M23*) accompanied, sometimes in the same protein, by different glycosidase activities (*Amidase_2*, *Glyco_hydro_19*, *Glyco_hydro_25*, *Transglycosylase*). It has been established (73) that the preferred lysin A structure is at least trimodular, with two EADs (typically a peptidase and a glycosidase/amidase) and a CWBD. Among this kind of architectures that was observed in $\mathbb{S}^{LYS}$, we may cite [*Peptidase_C39_2*] [*Amidase_2*] [*PG_binding_1*] or [*Peptidase_M23*] [*Amidase_2*] [*PG_binding_1*].

7) *Listeria* phage lysins also seem to contain a CWBD of its own, namely *PSA_CBD*, which appears at the C-terminal end of listeriophage lysins together with a peptidase or amidase EAD at N-terminal (rendering such structures as [*Amidase_2*] (1 or 2)× [*PSA_CBD*], [*Amidase_3*] 2×[*PSA_CBD*] or [*Peptidase_M15_4*] (1 or 2)× [*PSA_CBD*]).
8) *Lactococcus, Enterococcus*, and the genera in which lactobacilli have been recently reclassified (*Lactobacillus, Limosilactobacillus, Levilactobacillus, Lactoplantibacillus, Lacticaseibacillus*) (342) often displayed *Amidase_2* (preferential in the cocci) or *Glyco_hydro_25* as EADs. The unidentified C-terminal end of architectures that display an N-terminal *Amidase_2* within these groups of bacteria, again, evokes the plausible presence of a yet undescribed CWBD (*e.g.,* AGE60657.1, etc.). The most typical CWBD in this group was *LysM*. One or several LysM repeats at the C-terminal moiety often appear accompanying an N-terminal *Glyco_hydro_25*. SH3_5 also would work as a CWBD for some lactobacilli and enterococci phage lysins. A general conclusion is that, in these phages, lysins are commonly bimodular.
9) There was an important diversity of CWBDs among *Bacillus* phage lysins. Of the many families of CWBDs found, *Amidase02_C* and the putative CWBD *DUF3597* were exclusively observed in *Bacillus* phage lysins. Also, *SPOR* CWBD, which was also present in lysins from phages of clostridia, is known to be present in proteins involved in spore formation and maintenance (**Table 4.1.2**), so it is evolutionary plausible that it specifically appears in lysins from phages that infect sporulating bacteria such as *Bacillus* and clostridia. As for EADs in *Bacillus* phage lysins, there was a rather high variability as well but also presented some interesting features. For example, along with phages from Corynebacteriales and clostridia, *Bacillus* phage lysins were the only ones from G+ in which peptidase domains other than CHAP are somewhat widespread. Such peptidase-bearing lysins from *Bacillus* phages displayed a characteristic bimodular architecture

([*Peptidase_M15_4*] (1 or 2)× [*PG_binding_1*]). Very few *Bacillus* phage lysins presented a multicatalytic architecture (in fact, only two examples were found, [*NLPC_P60*] [*SLT*] and [*Peptidase_M23*] [*SLT*]). Most of the *Bacillus* lysins, then, display the traditional [EAD] [CBWD] architecture, with varying types of domains.

10) As previously concluded, the norm among the G– bacteria-infecting phages in $\mathbb{S}^{LYS}$ is to code for monomodular lysins, rendering that CWBDs were not quite widespread among these phages. However, there was a PF domain with an assigned CWBD function (namely, *PG_binding_3*) that was only present in G– bacteria within our dataset. Presumably, then, it might either specifically recognize a G– trait or be in some manner especially functional within the G– architecture. Other CWBDs that could occasionally be found were *PG_binding_1* and *LysM*. A differential trait of some G– lysins was the occurrence of N-terminal CWBDs together with an EAD at the C-terminal position. In this regard, the architecture [*PG_binding_1*] [*Muramidase*] was quite common. However, the most widely spread architecture among G– lysins was monomodular, with a single *Phage_lysozyme* domain. This was the most common architecture and family amongst G– bacteria by far, accounting for half [50.8%] of the identified G– lysins in our database, although it also appears, less commonly, in phages infecting some G+ bacteria (some bacilli, lactococci, and lactobacilli). There was still, however, a variety of EADs within G– lysins, among which we may cite *Muramidase, Amidase_2*, *Amidase_3*, *Hydrolase_2*, or *Peptidase_M15_4*, which, in turn, shows that G– phage lysins pose a certain diversity of catalytic activities, although muramidases dominate quantitatively.

Many of the aforementioned remarks are summarised in **Figure 4.1.5** and **Table 4.1.3** to facilitate interpretation.

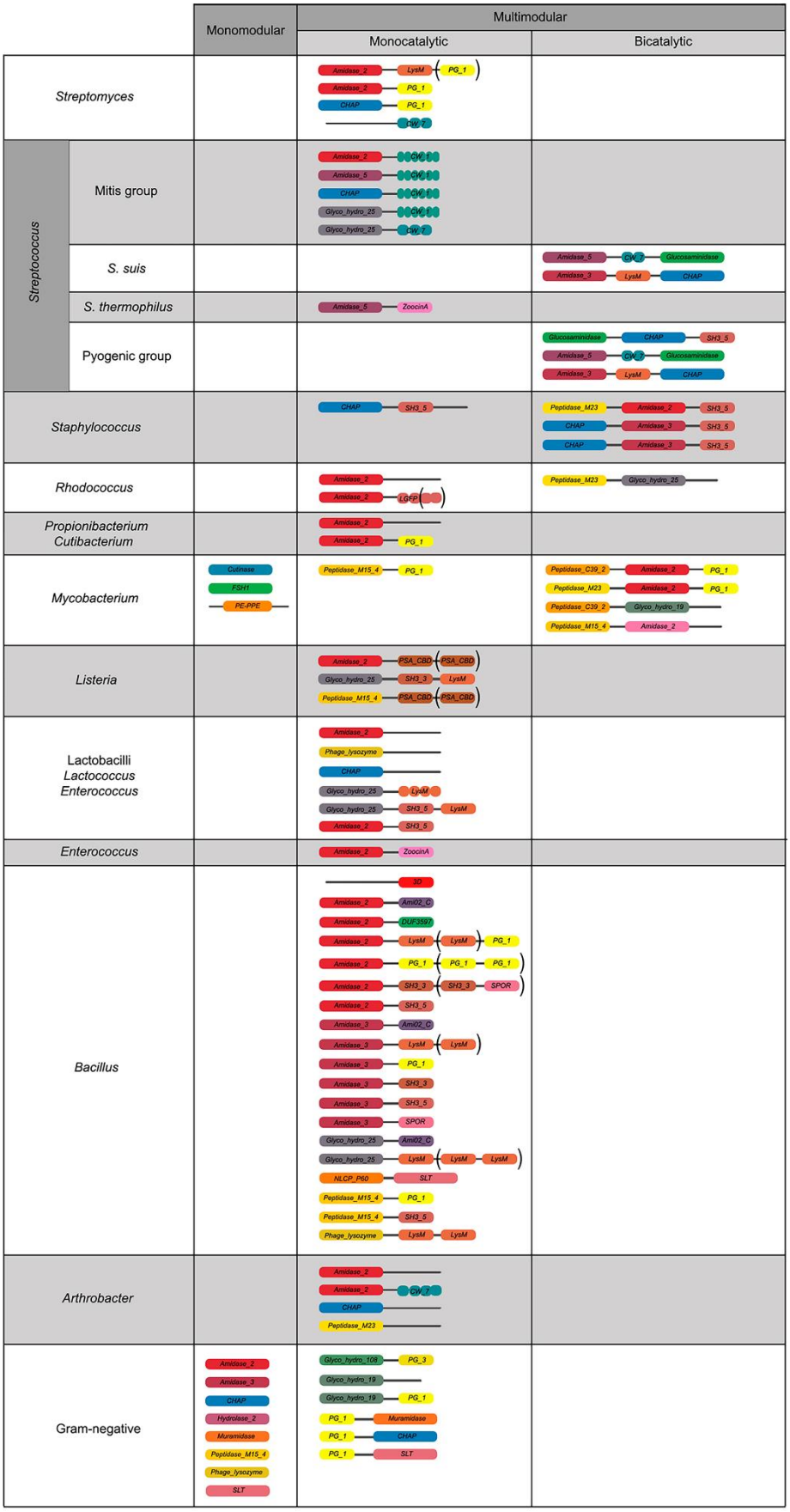


Figure 4.1.5. Relevant architectures observed in lysins from phages infecting different taxonomic groups of bacteria. Different colours mean different domains; brackets denote domains that appear only in some representatives of the depicted architecture.

Table 4.1.3. Summary table of the distribution of PF hits of phage lysins from G+ and G−[a].

Domain	Domain type	G+ (%)	G− (%)	Total	Encoded by phages of:
3D	EAD	7		7	
Amidase_2	EAD	547 (85.7)	91 (14.3)	638	Widely distributed in G+
Amidase_3	EAD	93 (81.6)	21 (1.4)	114	*Bacillus, Streptococcus, Clostridium, Staphylococcus*
Amidase_5	EAD	81 (100)		81	*Streptococcus, Lactococcus*
CHAP	EAD	186 (93.9)	12(6.1)	198	*Streptococcus, Staphylococcus, Strep-tomyces, Arthrobacter*
Cutinase	EAD	24		24	
FSH1	EAD	3		3	
Glucosaminidase	EAD	73 (97.3)	2 (2.7)	75	*Streptococcus*
Glyco_hydro_19	EAD	9 (14.3)	54 (85.7)	63	*Acinetobacter* and other genera
Glyco_hydro_25	EAD	142 (100)		142	Lactobacilli, *Bacillus, Streptococcus*
Glyco_hydro_108	EAD		43 (100)	43	Widely distributed in G−
GPW_gp25	EAD		12	12	
Hydrolase_2	EAD	3 (5.8)	49 (94.2)	52	*Escherichia, Pseudomonas, Vibrio*
Muramidase	EAD		35 (100)	35	*Pseudomonas, Burkholderia*
NLPC_P60	EAD	7	6	13	
PE-PPE	EAD	13		13	
Peptidase_C39_2	EAD	42 (100)		42	*Mycobacterium*
Peptidase_C93	EAD		4	4	
Peptidase_M15_3	EAD		10	10	
Peptidase_M15_4	EAD	96 (71.1)	39 (28.9)	135	*Mycobacterium, Bacillus*
Peptidase_M23	EAD	77 (97.5)	2 (2.5)	79	*Arthrobacter, Mycobacterium, Rhodo-coccus*
Pesticin	EAD		2	2	
Phage_lysozyme	EAD	32 (8.0)	366 (92.0)	398	Widely distributed in G−
Phage_lysozyme2	EAD	1		1	
Prok-JAB	EAD		1	1	
Prophage_tail	EAD	3		3	
SLT	EAD	6	16	22	
Transglycosylase	EAD	32 (100)		32	*Mycobacterium*
Amidase02_C	CWBD	20		20	
Big_2	CWBD	1		1	
CW_7	CWBD	125 (100)		125	*Streptococcus, Arthrobacter, Strepto-myces*
CW_binding_1	CWBD (repeat)	205 (100)		205	*Streptococcus*
CW_binding_2	CWBD (repeat)	3		3	
DUF3597	CWBD	5		5	
LGFP	CWBD (repeat)	31 (100)		31	*Rhodococcus*
LysM	CWBD	210 (97.7)	5 (2.3)	215	Widely distributed in G+
PG_binding_1	CWBD	130 (83.3)	26 (16.7)	156	*Mycobacterium, Bacillus, Streptomyces*
PG_binding_3	CWBD		40 (100)	40	Widely distributed in G−
PSA_CBD	CWBD	13		13	
SH3_3	CWBD	20	1	21	
SH3_5	CWBD	164 (100)		164	*Streptococcus, Staphylococcus*, lacto-bacilli, *Bacillus*
SPOR	CWBD	5		5	
ZoocinA_TRD	CWBD	50 (100)		50	*Streptococcus, Enterococcus*
Gp5_C	Structural		3	3	
Gp5_OB	Structural		3	3	

[a]Percentages and further remarks are only shown for domains represented by, at least, 30 hits.

4.3. Domains distribution across peptidoglycan chemotypes

To try to shed some more light on the potential evolutionary reasons for the differential distribution of EAD and CWBD families between the different bacterial groups, such distribution was examined regarding the different peptidoglycan chemotypes [as defined

by (63)] present in $\mathbb{S}^{LYS}$ (**Table 4.1.1**). Briefly, such peptidoglycan structure classification hierarchically relies on 1) the site of cross-linkage of the peptide subunit of the peptidoglycan, which defines the type (either A or B); 2) the nature of the cross-link, which defines the subtype (1, 2 or 3); and 3) the specific residue at position 3 within such peptide subunit, which makes up the specific peptidoglycan variant denoted by a Greek letter (**Figure 4.1.6A**).

Classification of PF hits by chemotypes did not provide a general explanation for CWBDs specificity better than other genus-specific traits (**Figure 4.1.6B**), such as the aforementioned presence of choline for the Mitis group streptococci that correlated with *CW_binding_1*. Some CWBDs did appear, however, only in restricted chemotypes.

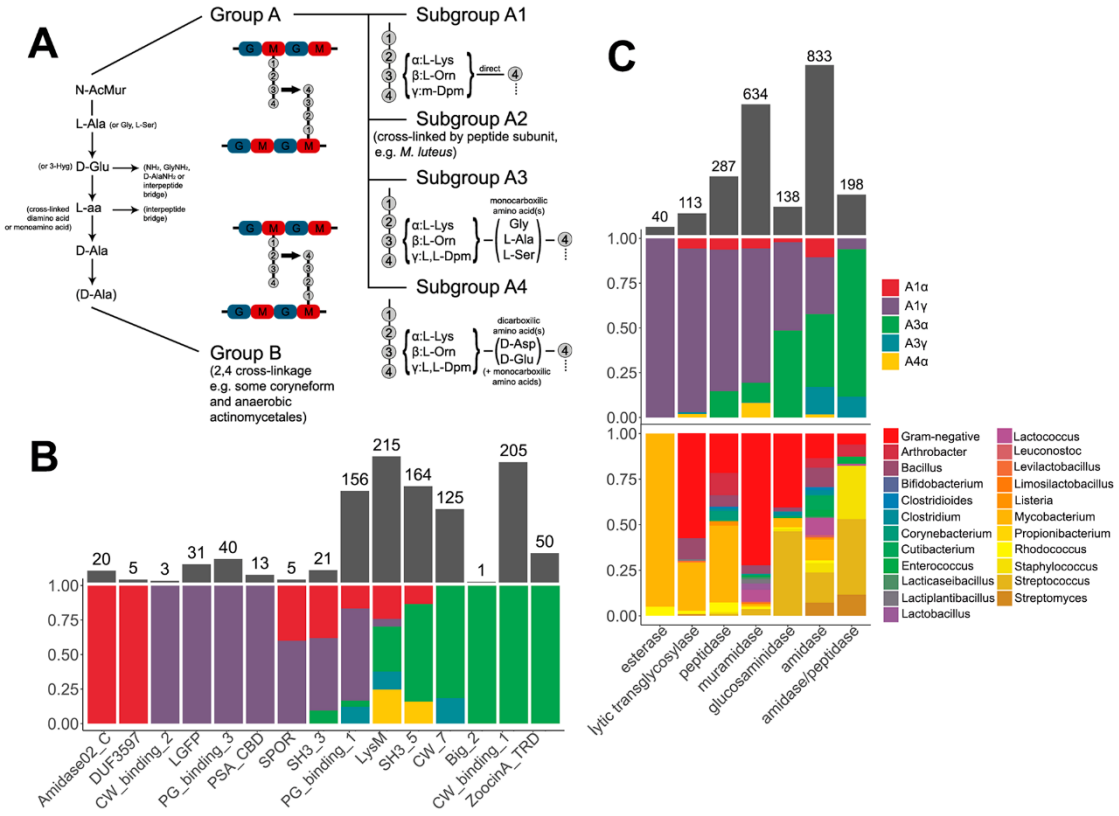


Figure 4.1.6. Differential distribution of CWBDs and catalytic activities across peptidoglycan chemotypes and taxonomic groups of bacterial hosts. A: Schematic representation of the relevant peptidoglycan chemotypes present for the bacterial hosts in S^{LYS}. **B:** Distribution of CWBD PF hits among chemotypes. **C**: Distribution of catalytic activities of EAD PF hits among chemotypes and taxonomic groups.

Such is the case of *Amidase02_C* or *PG_binding_3*, for example, which were only observed in lysins from phages that infect chemotype A1α and A1γ bacteria, respectively (since, as explained before, the former was only found in *Bacillus* phages and the latter in G− phages). Although current evidence does not allow drawing determinant

relationships between CWBD families and peptidoglycan composition in general, in some cases there is evidence supporting this. For example, *ZoocinA_TRD* is associated with A3α and it is known to bind precisely this peptidoglycan variant with two alanine residues at the cross-link (307). However, the causal link here goes from the published evidence backwards, *i.e.*, chemotype does not function as a good *a priori* predictor of the CWBD families' function. This is more clearly evident if we consider that some CWBD types appeared widespread among many different chemotypes, such as *LysM* and *SH3_5*. To check whether this apparent 'promiscuity' may be linked to the presence of subfamilies with potentially different ligands or if it could rather be a true promiscuous binding, SSNs were constructed with the PF hits of *LysM* and *SH3_5* (**Figure 4.1.7**). The *LysM* SSNs did not show prominent similarity clusters either classified by taxon or by chemotype of the bacterial host.

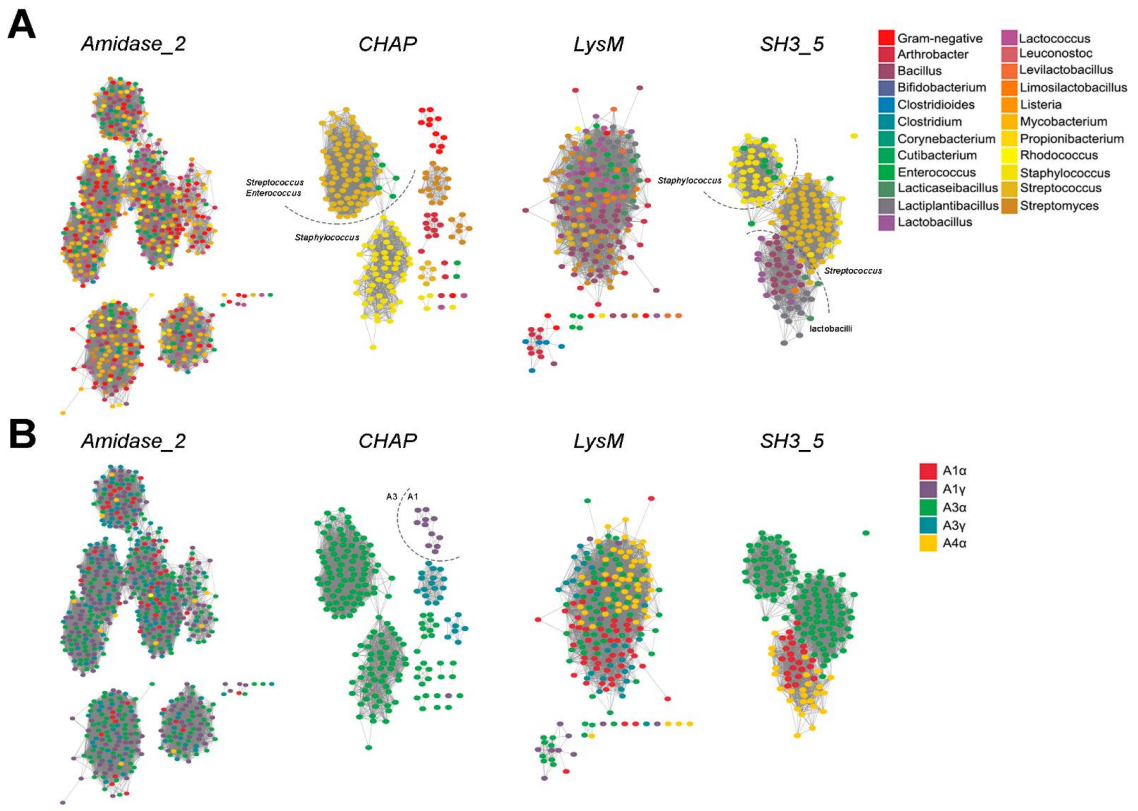


Figure 4.1.7. SSNs of the PF hits in $\mathbb{S}^{LYS}$ corresponding to different domain families. **A**: Classification of sequences according to the taxonomic group of the corresponding bacterial host. **B**: Classification by peptidoglycan chemotype of the host. Each node represents a single sequence. Dashed lines separate recognizable similarity clusters. Edge similarity cut-off was ≈ 40% for *CHAP*, *LysM*, and *SH3_5* and ≈ 30% for *Amidase_2*.

This observation suggests that *LysM* could be a truly 'universal' CWBD that may thus bind to a conserved cell wall ligand. The rather generic description of *LysM* ligands in the literature (as 'NAG-containing polysaccharides') supports this hypothesis (298). *SH3_5*, however, displayed at least two different sequence similarity groups that correlated well with different taxonomic groups. In this way, a cluster for staphylococcal phage domains was separated from another that comprised streptococci and lactobacilli, while lactobacilli phage domains also clustered together and apart from the ones from streptococcal phages. This strongly suggests the existence of different 'subfamilies' within *SH3_5* that would account for its apparent wide presence among different bacterial host groups. In fact, literature reflects that, while lytic enzymes with predicted *SH3_5* domains typically recognize polysaccharides (and peptidoglycan itself in particular) there seem to be different specializations. For example, the CWBD of the *Lactiplantibacillus plantarum* major autolysin Acm2 bound many different peptidoglycans with low affinity, being glucosamine the minimal binding motif (304). On the other hand, *SH3_5* domains from staphylolytic enzymes specifically bind crosslinked peptidoglycans (like the A3α peptidoglycan of *Staphylococcus* and *Streptococcus*), and the nature of the crosslink itself determines the affinity of such CWBDs for the peptidoglycan (302, 303).

The distribution of catalytic activities of the EAD PF hits in $\mathbb{S}^{LYS}$among chemotypes was more informative (**Figure 4.1.6C**). First of all, NAM-amidases were the most represented type of domains and also those that appeared among more different taxonomic groups and chemotypes, even more so than lysozymes. Indeed, *Amidase_2*, the most abundant PF domain in $\mathbb{S}^{LYS}$ (with 638 hits, **Table 4.1.3**), appeared both in lysins from G+ and G− phages. The SSN in **Figure 4.1.7** shows, however, that although *Amidase_2* seems a rather diverse group, with various observable similarity clusters, none of such clusters correlate with any of the classifiers of the bacterial hosts tested. Therefore, it cannot be posed any specialization hypothesis for *Amidase_2* at this point. Regarding muramidases, they were quite common among chemotype A1γ, probably in correlation with the widespread presence of *Phage_lysozyme* domains in G− infecting phage lysins. Glucosaminidases appeared evenly in lysins from phages that infect bacteria with peptidoglycans A1 and A3, but whereas in G+ (all A3s and a few A1s) glucosaminidase activity was mainly attributed to *Glucosaminidase* PF hits, the only domain family among G– putatively assigned a glucosaminidase activity was *Glyco_hydro_19*. Even more interesting, PF domains with peptidase activity were much more common in lysins from phages that infect subtype A1 bacteria. Thus, peptidases were not uncommon among G– and were also present in phages from A1 G+ (Corynebacteriales, *Listeria*, clostridia, or *Bacillus*). On the other hand, amidases/peptidases (which is the label given to *CHAP* domains) were much more prevalent among phages with A3 host bacteria, and only

seldom present in A1 (namely some G−). This suggests the presence of two different peptidase family clusters, one specialized in A1 and the other in A3 chemotype peptidoglycans, which would make sense given that the complexity of the peptide moieties is the main difference between these two subtypes. Adding to this conclusion, the *CHAP* SSN (**Figure 4.1.7**) did show clustering of the few *CHAP* examples in lysins from A1 phages, besides an apparent differentiation of *Staphylococcus* and *Streptococcus*/*Enterococcus*.

5. PHYSICOCHEMICAL DIFFERENCES BETWEEN PHAGE LYSINS FROM GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA

Thus far, a correlation of certain architectural features with different bacterial groups has been demonstrated, and some insights into the possible nature of these specializations have been provided in an evolutionary context. One of our prior hypotheses in this work was to provide evidence on the appearance of certain computable physicochemical properties in lysins from G−, thus linking one of such observable specializations with a possible role in the function of the respective lysins. To check the reach of these presumed AMP-like elements, first, a custom random forest predictive algorithm based on sequence-calculated physicochemical properties (NCPR, hydrophobicity, HM, aliphatic index) was set up (**Figure 4.1.8ABC**). The resulting algorithm yielded rather good quality parameters (*e.g.*, ROC-AUC = 0.897; testing set prediction accuracy = 87.9%, **Figure 4.1.8AB**) given the simplicity of the proposal. According to the subsequent analysis of the random forest model (**Figure 4.1.8C**), NCPR was the most relevant variable to distinguish between G+ and G− host bacteria, followed by HM and aliphatic index and, finally, hydrophobicity. The quality of the prediction could be much improved in a variety of manners, for example, including additional descriptors such as protein length, which is a quite good discriminator between lysins of G+ and G− (**Figure 4.1.1**). A second model built including such variables as protein length, the number of PF hits per protein and the presence or absence of a signal peptide achieved, in fact, a 96.15% accuracy and much better classification performance (**Figure 4.1.8DE**). As expected, in this case, the variable with a greater impact in classification was the protein length (**Figure 4.1.8F**) which, given the inevitably incomplete character of the PF predicted domains, is a better account of the architecture of the proteins (*i.e.* of the presence of just one or many domains, which is arguably the main difference between the two classes considered). However, the aim here was not to perform an accurate prediction of the bacterial host Gram group, but rather to assess how well a set of relevant physicochemical features could predict the Gram group. Indeed, the results indicate that there must be a certain physicochemical difference between G+ and G− lysins.

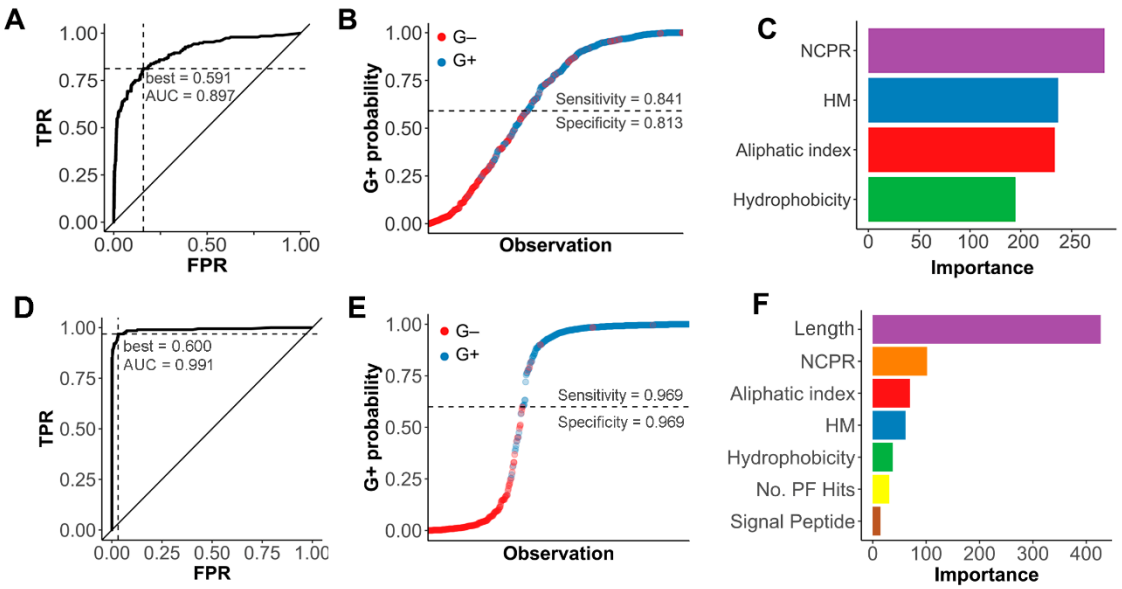


Figure 4.1.8. Random forest prediction and classification of lysin sequences according to their bacterial host Gram group. Two models are presented, the first one based on just lysins physicochemical properties (**A, B, C**) and the second one including also other variables relative to lysin architecture (**D, E, F**). **A, D**: ROC curves of the random forest predictive models (TPR: true positive rate, FPR: false positive rate). ROC best points of positive group (G+) probability for outcome maximization are presented, as well as the AUCs. **B, E**: Random forest castings of bacterial host Gram group on the testing subset of lysin sequences. The dashed lines represent the G+ probability threshold for classification based on the respective ROC best points. **C, F**: Importance (*i.e.*, mean Gini index decrease for each variable) of each of the four descriptors used for classification within each model.

This difference was further examined by comparing the distribution of said properties between lysins from G+ and G− phages by performing a series of analyses (**Figure 4.1.9**). The distribution of NCPR presented significantly higher values in G− than in G+ ($p \leq 0.0001$; ES = 0.66, **Figure 4.1.9A**). Moreover, the average prediction of local net charge (**Figure 4.1.9B**) indicated that such difference is mainly located at the C-terminal part of lysins from G−. A statistical comparison of NCPR at the last quarter of the proteins (Q4) between G+ and G− also confirmed that the main difference in net charge was located at this end (**Figure 4.1.9C**). Although the difference between the G+ and G− bacterial host groups was deemed statistically significant in every sequence quartile ($p \leq 0.0001$), the effect size was only 'moderate' at Q1-Q3 (ranging between 0.24 and 0.38), but it was higher at the final quartile (ES = 0.52). Hydrophobicity was also generally higher in G− lysins, but the difference regarding G+ was smaller ($p \leq 0.0001$; ES = 0.36). G– lysins had a more hydrophobic N-terminal part, which could be correlated to the previously discussed greater abundance of predicted signal peptides among the G– group (**Figure 4.1.1F**).

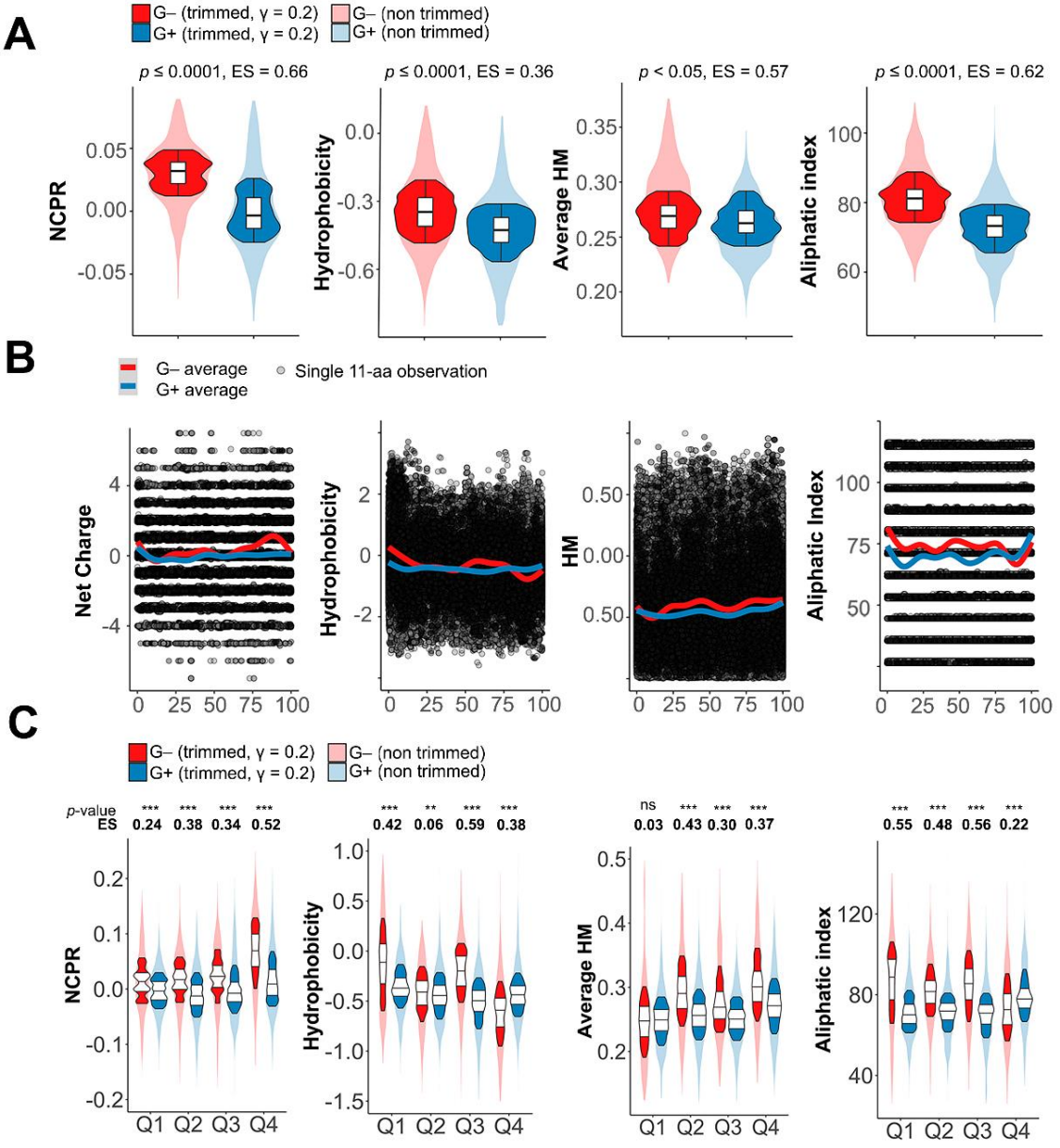


Figure 4.1.9. Differential physicochemical properties distribution among G+ and G− phage lysins. A: Distribution of net properties calculated along the whole protein sequences of lysins from phages infecting G− or G+. **B**: Local computation of physicochemical properties. Each dot represents the particular value calculated for an 11-aa window in a given lysin. Continuous lines are average tendencies based on either all G− or all G+ data points. **C**: Distribution of different properties at quartiles of lysin sequences. Asterisks indicate *p*-values (** ≤ 0.01, *** ≤ 0.001) obtained from the Yuen-Welch test for trimmed means with a trimming level of γ = 0.2; ES indicates the Wilcox and Tian's ζ measurement of effect size, as explained in **MATERIALS AND METHODS Section 1.6**.

The tendency was reversed at the C-terminal moiety, something that can be explained by the relative abundance of positively charged residues shown before for G–. It is at the third quartile (Q3), immediately before the high positive net charge patch described above, where the difference was statistically more relevant ($p \leq 0.001$; ES = 0.59). These

observations could also reflect the lysins' architecture: we have already established that the ones from G– infecting phages are generally monomodular and the G+ ones are multimodular. The multimodular architecture of the latter implies longer sequences and the presence of inter-domain linkers, which, being more exposed, tend to have more polar or charged residues (343). These circumstances would support an altogether lower hydrophobicity in lysins from G+ infecting phages, together with the fact that the G–, containing a single domain, would mostly tend to be globular, and thus comprise less exposed portions and relatively more hydrophobic residues. Perhaps linked to these phenomena, there was also found a statistically significant difference in the amphipathicity, measured by the average HM distributions, and in the aliphatic index. The G– presented greater average HM values along the whole protein sequence but the N-terminal end; while the aliphatic index was also consistently higher in G– than in G+ except at C-terminal (coincidental with the high net charge patch). Precisely, it is the presence of a high NCPR coupled with a still higher HM that evokes a similarity with AMPs. The lower hydrophobicity at Q4 of the G− group can be easily explained by the presence of positively charged residues, but the high HM suggests that nonpolar residues remain at such regions, potentially allocated in amphipathic 3D structures.

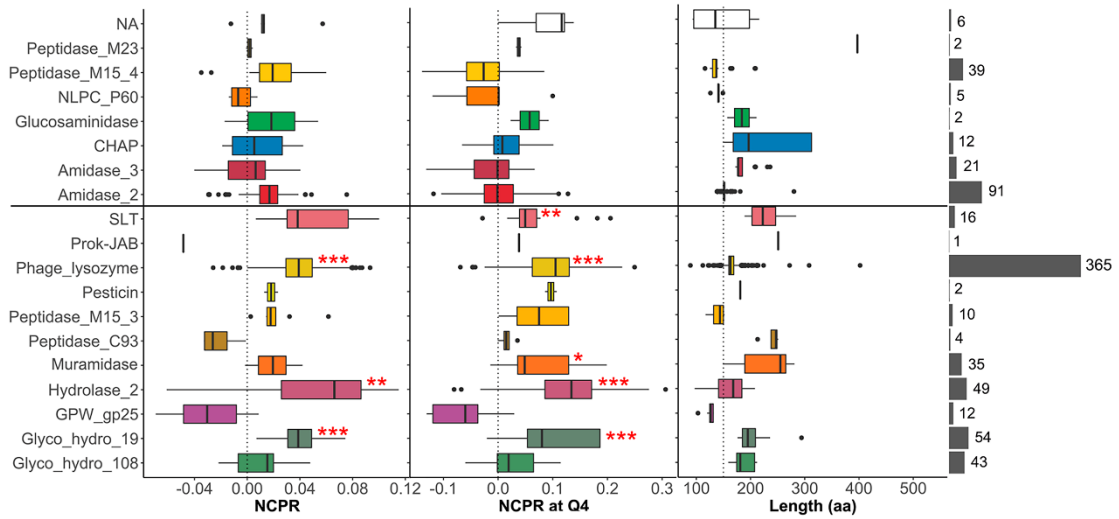


Figure 4.1.10. Net charge distribution of lysins from G− infecting phages classified according to the predicted EAD. The rightmost grey bars depict the number of lysins classified into each EAD group (lysins within the NA group are those for which an EAD was not assigned). All groups were compared with the distribution of the *Amidase_2* domain, as a highly represented, near-neutral control using Welch's test on trimmed means with *post hoc* Bonferroni correction (*, p-value ≤ 0.05; **, p-value ≤ 0.01; ***, p-value ≤ 0.001). The horizontal line separates EADs preferentially found in G+ (top) and those preferential for G−.

A closer examination of NCPR of lysins from G– infecting phages indicated that the high positive charge patch trait seems specific to some EAD domain families (**Figure**

4.1.10). Specifically, a statistically significant higher NCPR was found in lysins bearing *Phage_lysozyme*, *Hydrolase_2*, and *Glyco_hydro_19* domain families. When the analysis was restricted to the C-terminal end, the NCPR difference was observed for the same domains mentioned above but also in *SLT* and *Muramidase*. The average local net charge tendency for each EAD group found in G− lysins confirmed that a local positive net charge peak appears in the protein region immediately before the C-terminal end for all of the aforementioned domains (**Figure 4.1.11**). Interestingly, all of those domains were preferentially present in lysins from G− phages, according to **Figure 4.1.2**. This observation provides a basis to argue an evolutionary tendency in some G– infecting phages towards developing AMP-like subdomains at the C-terminal moiety of their lysins.

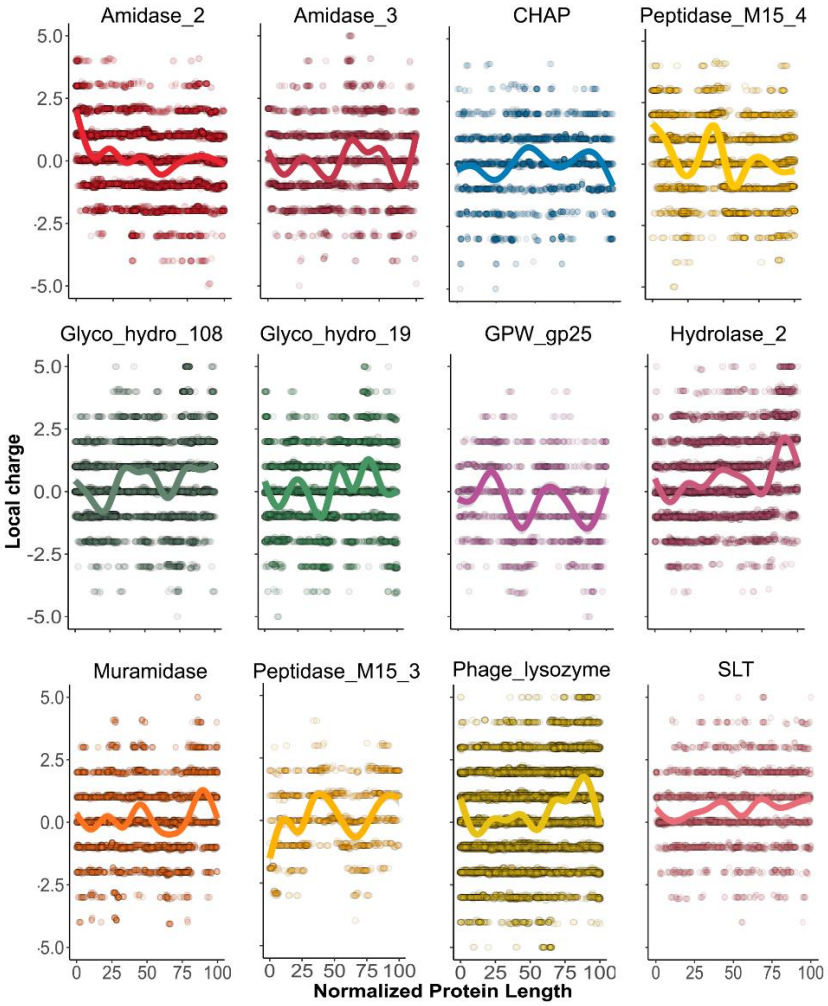


Figure 4.1.11. Local computation of physicochemical properties in lysins from G− phages classified according to EAD predictions. Each dot represents the particular value predicted for an 11-aa window from a given lysin. Continuous lines are average tendency lines.

IV. RESULTS

CHAPTER 2 – Use of physicochemical properties calculation for the mining of enzybiotic candidates. Characterization of enzybiotic candidate Pae87

Part of the results in this chapter are published in:

Roberto Vázquez, Sofía Blanco-Gañán, Susana Ruiz, Pedro García*. 2021. Mining of Gram-negative surface-active enzybiotic candidates by sequence-based calculation of physicochemical properties. Front. Microbiol. doi: 10.3389/fmicb.2021.660403.

Other results are included in the following manuscript in preparation:

Roberto Vázquez, Mateo Seoane-Blanco, Virginia Rivero-Buceta, Susana Ruiz, Mark J. van Raaij, Pedro García*. 2021. *Pseudomonas aeruginosa* phage JG004 lysozyme (Pae87) contains a bacterial surface-active antimicrobial peptide-like region and a possible independent substrate binding subdomain.

1. INTRODUCTION

Since many G− phage lysins bear AMP-like elements and, intrinsically, such domains should be able to interact with the G− bacterial surface, and even disrupt it, a bioinformatic search for AMP-like domains was set up for uncovering potential enzybiotic candidates. To this end, a simple *k*NN model was built to distinguish AMPs from regular, random peptides. Such a predictive model was applied to a database subset consisting of the monomodular *Pseudomonas*-infecting phage lysins in $\mathbb{S}^{LYS}$ (see the previous chapter). A further screening procedure was applied based on current literature and, in this way, a small collection of enzybiotic candidates was obtained. Enzybiotic candidate Pae87 was overexpressed, purified, and subsequently characterized. General evidence is thereby provided on Pae87 antimicrobial potential and, more specifically, its mode-of-action, including its ability to interact with G− bacterial surface. A three-dimensional structural model of Pae87 was obtained, and its potential AMP-like subdomain (named P87) was defined and tested.

2. DIFFERENTIATING ANTIMICROBIAL PEPTIDES BY THEIR PHYSICOCHEMICAL PROPERTIES

As explained in MATERIALS AND METHODS (**Section 1.5**), a database named $\mathbb{S}^{AMP}$, comprising aa sequences of AMPs and randomly generated peptides was constructed. Physicochemical properties (NCPR, hydrophobicity, average HM, aliphatic index) were calculated based on such sequences. As pointed out in previous literature (344, 345), AMPs in $\mathbb{S}^{AMP}$ displayed a statistically significant higher NCPR, hydrophobicity and HM (*p*-values < 0.0001; respective effect sizes 0.74, 0.37, 0.57, **Figure 4.2.1A**). No significant differences, however, were found in the aliphatic index, unlike what observed in the previous chapter for the mostly monomodular lysins of G−. Aliphatic aa (specifically Ile, Val, and Leu) are known to play a role in the folding and stabilization of globular proteins (346, 347), and thus it is understandable that they may be overrepresented in the globular G− lysins with respect to the multimodular G+ ones (**Figure 4.1.9**). Multimodular lysins necessarily contain linkers, which are known to be rich in polar residues and proline, rather than nonpolar or aliphatic ones (343). Although AMPs stand out by a somewhat higher hydrophobicity than randomly picked peptides, such hydrophobicity will not be necessarily contributed by aliphatic aa, which are not needed for globular stabilization or folding of a tertiary structure in short peptides such as AMPs in general.

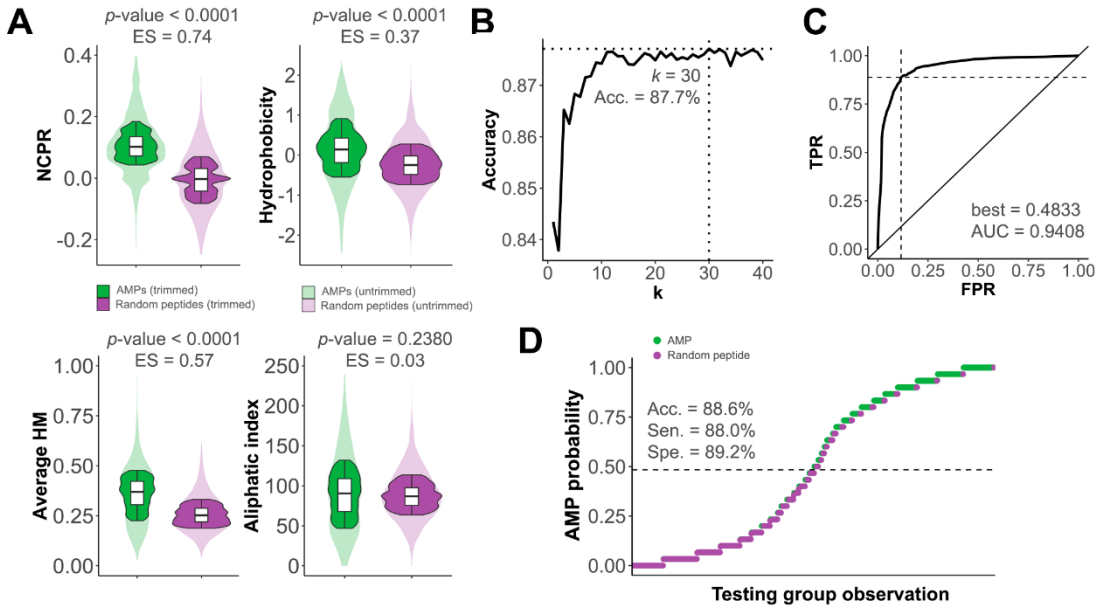


Figure 4.2.1. Characteristic features of AMPs and *k*NN predictive model construction. A: Distribution of sequence-based physicochemical features in $\mathbb{S}^{AMP}$ comparing between AMPs and non-AMPs. Original and trimmed (trimming level γ = 0.2) distributions are shown. Yuen-Welch test was applied to analyse significant differences and effect size was calculated according to Wilcox and Tian's ζ. **B**: parameter *k* fitting for the *k*NN model. **C**: ROC curve analysis of the *k*NN using the testing subset of $\mathbb{S}^{AMP}$. AUC and ROC-based best AMP probability threshold point for classification are provided. **D**: Classification of the testing subset using the *k*NN model. The dashed line represents the AMP probability threshold for classification based on the ROC best point. Sensitivity (Sen), Specificity (Spe) and Accuracy (Acc), are provided considering such classification threshold.

$\mathbb{S}^{AMP}$ was partitioned into a training and a testing subset (respectively 75% and 25% of the total entries, randomly sampled and classified into either of the subsets), and those subsets were used to build and evaluate a *k*NN model aimed to distinguish between AMPs and non-AMP peptides. Using the training subset, the model was fitted (with *k* = 30, optimal accuracy = 87.7%, **Figure 4.2.1B**) and subsequently evaluated with the testing subset (**Figure 4.2.1CD**). The ROC analysis yielded an AUC-ROC of 0.941 and optimal TPR and FPR of 0.888 and 0.116, respectively. The best cut-point of AMP probability for classification was established in 0.4833, based on this analysis. With this cut-off, classification quality metrics are shown in **Figure 4.2.1D**. Many efforts have been previously conducted to accurately predict AMPs, some with accuracies well above 90% (198, 348-351). However, these works aimed to produce depurated, highly reliable predictions of AMPs (often using resources out of scope here, such as larger descriptor sets or more sophisticated sequence metrics). Our purpose, however, was rather to get a simple and customized prediction in line with our analytic approach, and therefore the results yielded can be considered good enough to further proceed. It is more so if we consider that, rather than predicting AMPs themselves, our purpose was to generate a

'similarity metric' (which would be the AMP probability score computed by the kNN) based on physicochemical properties, with as less sequence bias as possible.

3. APPLICATION OF THE *k*NN TO THE PREDICTION OF AMP-LIKE ELEMENTS WITHIN LYSINS AND SCREENING OF ENZYBIOTIC CANDIDATES

The next step was applying the AMP-recognizing model to predict AMP-like peptides within the C-terminal end of the lysins in $\mathbb{S}^{LYS}$ (**Figure 4.2.2A**). Coherently with our observations in **Chapter 1**, most of the lysins for which an AMP was predicted at the C-terminal analysed region were G− (66.3%). Ultimately, this meant that AMPs were predicted for around 32% of all of the G− entries, in contrast with the G+ ones, for which only ≈ 10% contained a predicted C-terminal AMP.

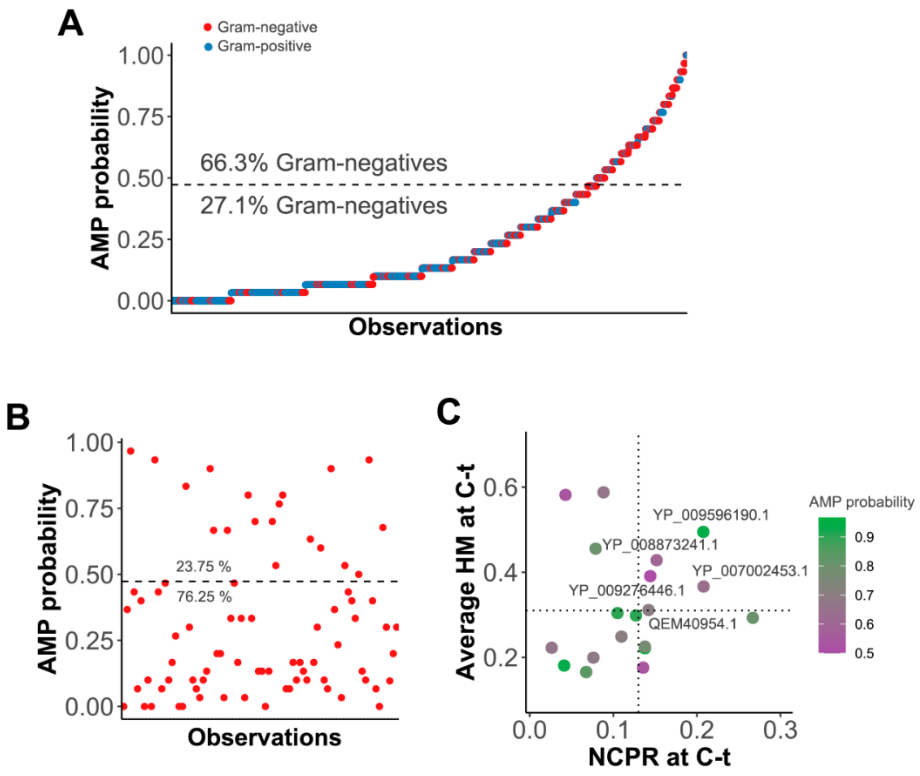


Figure 4.2.2. Prediction of AMP-like elements within lysin sequences. **A:** AMPs prediction at the C-terminal end of the lysins in $\mathbb{S}^{LYS}$, classified by the Gram group of their host bacteria. **B**: AMPs prediction at the C-terminal end of the lysins in $\mathbb{S}^{PSE}$. Dashed line represents the AMP probability cut-off for classification. **C**: Secondary screening of AMP-positive $\mathbb{S}^{PSE}$ entries by maximizing HM and NCPR. GenBank identifiers of the final candidates are shown. The dashed lines indicate the parameter cut-off values (HM ≥ 0.31 and NCPR ≥ 0.13). C-t means C-terminal region.

After this check, the *k*NN model was used to screen for enzybiotic candidates against bacterial pathogen *P. aeruginosa* using a subset of $\mathbb{S}^{LYS}$, named $\mathbb{S}^{PSE}$, only containing lysins whose bacterial host was labelled as *Pseudomonas* and whose architecture was monomodular. The 23.75% (19 examples) of the lysins in $\mathbb{S}^{PSE}$ were positive for an AMP-like element at the examined C-terminal region (**Figure 4.2.2B**). This does not rule out that any of the negative ones could bear AMP-like elements, but rather that there would be a low probability for it to be located exactly at the analysed coordinates within the respective sequences (between 0.75 × length of the protein and 0.9 × length, see MATERIALS AND METHODS **Section 1.5**). However, only the ones for which we found plausible AMPs at such region were kept for the sake of the screening, and a second selection step was applied to be left with the most promising ones for further testing (**Figure 4.2.2C**). This second screening was based on the C-terminal maximization of NCPR and HM. Cut-offs were defined by the lowest observed values of such variables in regions positively classified as AMPs among previous examples of enzybiotics intrinsically active against G− (**Table 4.2.1**, NCPR = 0.13, HM = 0.31).

Table 4.2.1. Previously published examples of lysins with intrinsic activity against Gram-negative bacteria from without.

GenBank ID	Phage	Bacterial host	PF architecture	aa	AMP score[a]	NCPR at C-t[b]	HM at C-t[b]	Reference
AJG41873.1	RL-2015	*A. baumannii*	*Phage_lysozyme*	146	0.7000	**0.1279**	0.4097	(169)
ARB16052.1	JD010	*P. aeruginosa*	*Phage_lysozyme*	145	0.1333	0.0844	0.1580	(170)
ALC76575.1	vB_AbaP_CEB1	*A. baumannii*	*Glyco_hydro_19*	185	0.6667	0. 2392	**0.3111**	(352)
AEV89716.1	OBP	*P. aeruginosa*	NA (bimodular)	328	0.0000	−0.0021	0.2372	(179)
YP_006383882.1	SPN9CC	*Salmonella enterica*	*Phage_lysozyme*	167	0.4333	0.1127	0.4908	(168)
ADX62345.1	φAB2	*A. baumannii*	*Glyco_hydro_19*	185	0.7333	0.2393	0.3236	(182)
AAK40280.1	Morita2001	*Bacillus amylolyquefaciens*	[*Phage_lysozyme*] 2× [*LysM*]	258	0.1333	0.1008	0.2575	(164)
YP_009285691.1	vB_CfrM_CfP1	*C. freundii*	*Peptidase_M15_4*	131	0.0000	−0.0028	0.2331	(194)

[a]AMP probability as computed for the peptide comprised between 0.75 × length and 0.9 × length coordinates. Values above the AMP probability threshold are underlined.
[b]C-t means C-terminal region. The NCPR and HM values used as a cut-off for the screening depicted in **Figure 4.2.2C** are shaded in black.

It is worth noting that only three of the eight examples provided in **Table 4.2.1** yielded a positive result in the kNN AMP classification. Again, this does not imply that the negative ones do not contain AMP-like regions, but rather that such regions were not detected by the model within the predefined coordinates. The latter explanation is supported by the profiles of local AMP predictions shown in **Figure 4.2.3**. In this figure, a plausible

accumulation of AMP-positive peptides can be observed in all of the lysins in **Table 4.2.1**, although in some of them those areas lie outside the [0.75, 0.9] region considered here for the screening.

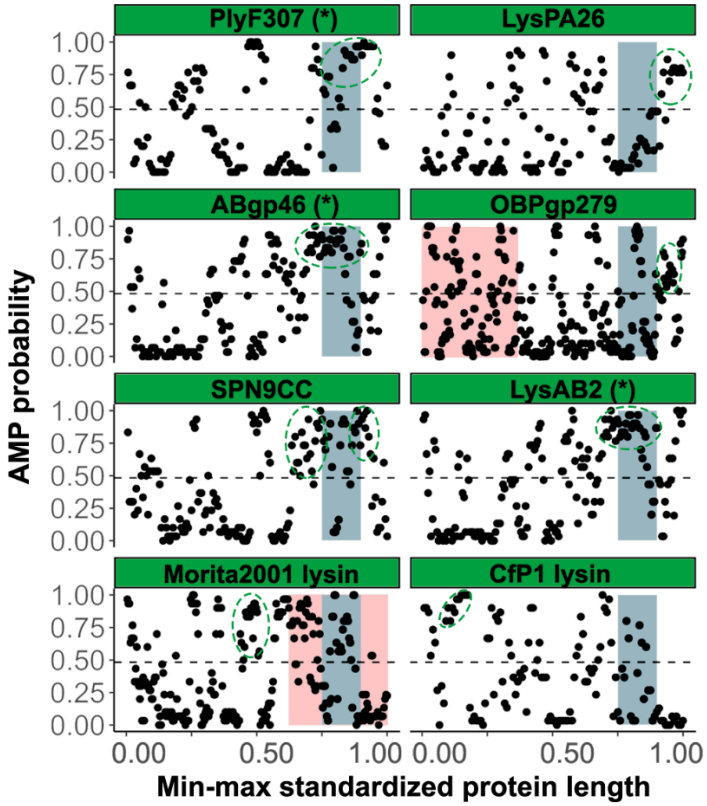


Figure 4.2.3. Local predictions of AMPs on 11-aa window peptides along the sequences of experimentally proven lysins active against Gram-negative bacteria. The blue shade indicates the [0.75, 0.9] window on which the original predictions were made for **Table 4.2.1**. The red shade indicates CWBDs. Lysins with an asterisk correspond to those in **Table 4.2.1** for which an AMP was predicted in the [0.75, 0.9] region. Plausible areas of accumulation of 11-aa predicted AMPs are indicated with a green, dashed circle.

In any case, five enzybiotic candidates were finally derived from the bioinformatic screening (**Table 4.2.2**). Three of them contained a *Phage_lysozyme* catalytic domain and the other two were *Muramidase*, both catalytic domains typically found among G− (see the previous chapter). In this thesis, we have focused on the characterization and potential antibacterial profiting of one of them, the putative endolysin of the JG004 phage, named throughout this work as Pae87 (353), on the basis of its good heterologous expression yield in a soluble form.

Table 4.2.2. Final enzybiotic candidates after the screening process.

GenBank ID	Phage	Bacterial Host	Pfam Architecture	Length (aa)	NCPR at C-t[a]	Average HM at C-t[a]
YP_009596190.1	VSW-3	*P. fluorescens*	*Phage_lysozyme*	156	0.1485	0.3764
YP_009276446.1	phi3	*P. aeruginosa*	*Phage_lysozyme*	170	0.1103	0.3756
YP_008873241.1	PPpW-3	*P. plecoglossicida*	*Phage_lysozyme*	168	0.1383	0.3990
YP_007002453.1	JG004	*P. aeruginosa*	*Muramidase*	186	0.1465	0.2816
QEM40954.1	PAP-JP	*P. aeruginosa*	*Muramidase*	187	0.1421	0.3113

[a]C-t means C-terminal region.

4. GENERAL CHARACTERISTICS OF PHAGE JG004 AND ITS ENDOLYSIN, PAE87

Myoviridae phage JG004 was originally isolated from environmental samples (sewage water) by selecting lytic activity against *P. aeruginosa* (353). No genetic elements suggesting lysogenic ability have been described for it within the 93,017 bp of its genome, and it has been used as a proof-of-concept for the formulation of nanodroplet emulsions for respiratory therapy against *P. aeruginosa* (354). The gene *PJG4_087*, encoding putative endolysin Pae87, is found in a typical lytic cassette context within the phage genome: surrounded by other genes possibly dedicated to bacterial host lysis (**Figure 4.2.4**). Specifically, downstream of the lysin gene there is a gene putatively coding for a holin (with small size, and two transmembrane regions predicted) and there may be one or more spanin genes. Interestingly, the putative holin of the JG004 phage is in an opposite location to the typical 5'-holin-lysin-3' cassette.

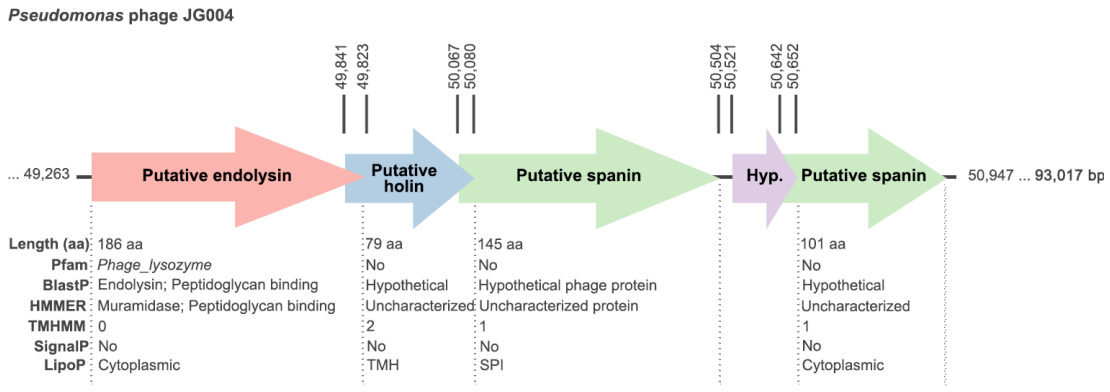


Figure 4.2.4. *Pseudomonas* phage JG004 lytic cassette. All putatively claimed functions have been proposed based on the bioinformatic analysis results depicted below each open reading frame (ORF). The rightmost bold number is the phage genome size in base pairs. Vertical numbers mark the coordinates of the corresponding ORFs.

Pae87 EAD, which is predicted to occupy the whole length of the protein (the Hidden Markov Model (HMM) alignment produced by PF covers from coordinate 19 to the very end of the sequence), belongs to the recently described *Muramidase* catalytic family (326). Pae87 shows 52.2% sequence identity with the first described member of the *Muramidase* family, *Salmonella* phage lysin Gp110, for which a muramidase activity was experimentally proven (**Figure 4.2.5**).

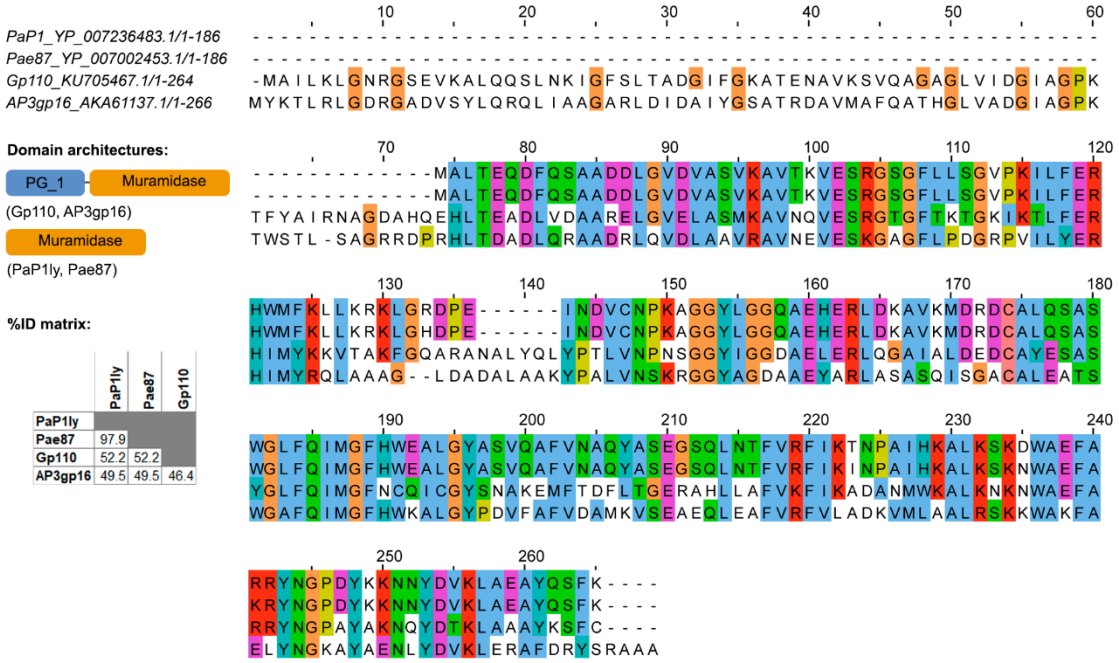


Figure 4.2.5. Multiple sequence alignment of Pae87 and other lysins with EADs from the same family reported in the literature. Conserved aa are coloured, domain architectures comprised by the analysed lysins, and respective percent identity matrices are also shown.

To date and up to our knowledge, a single experimental 3D structure example of the *Muramidase* family has been contributed: that of the *Burkholderia* phage lysin AP3gp15 (173). The fundamental difference between Pae87 and AP3gp15 is the presence, in the latter, of a *PG_binding_1* CWBD at the N-terminal end. Besides, a member of the family, the *P. aeruginosa* PaP1 phage lysin, which has been reported to degrade purified *P. aeruginosa* cell walls and inhibit *S. aureus* (but not *P. aeruginosa*) growth at very high concentrations (up to 10 mg/ml droplets on agar plates), is very similar to Pae87 (97.9% identity) (355).

5. PAE87 EXPRESSION, PURIFICATION, AND PRELIMINARY ANTIBACTERIAL ACTIVITY TESTING

Pae87 was overexpressed in an *E. coli* BL21(DE3) host, fused to an N-terminal 6 × His tag, and subsequently purified by immobilized nickel ion chromatography followed by a chromatographic buffer exchange step to remove imidazole and NaCl (**Figure 4.2.6**). Production yield was about 20 to 30 mg/l.

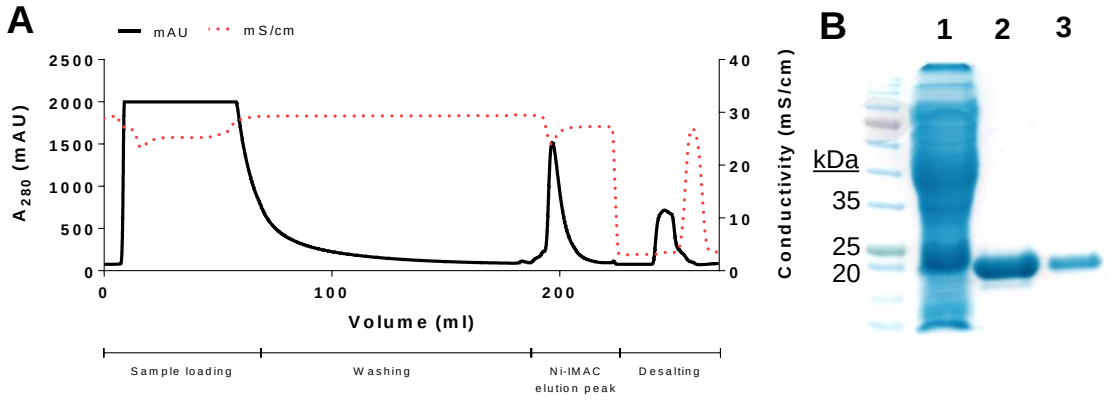


Figure 4.2.6. Purification of Pae87. A: Chromatogram of the two-step chromatographic protocol for Pae87 purification. **B**: SDS-PAGE of a crude Pae87 expression extract (lane 1) and samples from the affinity chromatography elution peak (lane 2) and the desalting peak (lane 3).

Muralytic activity assay on RBB-labelled purified *P. aeruginosa* PAO1 peptidoglycan demonstrated that Pae87 has a catalytic peptidoglycan-degrading activity since it was able to solubilize the substrate polymer in a much greater degree than the negative control, anti-pneumococcal enzyme Cpl-711 (**Figure 4.2.7A**). Also, Pae87 was observed to have a viable cell counts decrease effect (**Figure 4.2.7B**) when in contact with *P. aeruginosa* PAO1 cells, down to a maximum of ≈ 2 log units in the assayed conditions (pH 6.0, 150 mM NaCl, ≈ 10^8 CFU/ml). This apparent viability decrease, however, did not correlate with a pronounced turbidity reduction as it would be expected for a lytic enzyme (**Figure 4.2.7C**). Such bacteriolytic effect is easily observed, in general, among lysins that attack G+ bacteria (78). It is, nonetheless, normally impeded by the OM in the G− setup, and proper lysis is usually only observed in G− with wild-type lysins when the OM is permeabilized through a pre-treatment with an organic acid or adding a permeabilizing agent, such as EDTA (356). When Pae87 was used to treat PAO1 cells permeabilized with 0.5 mM EDTA, however, there were no remarkable changes in the turbidity decrease profile (**Figure 4.2.7C**). EDTA did produce, nonetheless, a synergistic bactericidal effect when applied in combination with Pae87 (**Figure 4.2.7D**), likely due to the OM disrupting effect of EDTA, which would improve the interaction of the protein with

the bacterial cell wall. While the killing activity was ≈ 2 log units in the absence of EDTA, it was around 5 log units when 0.5 mM EDTA was added.

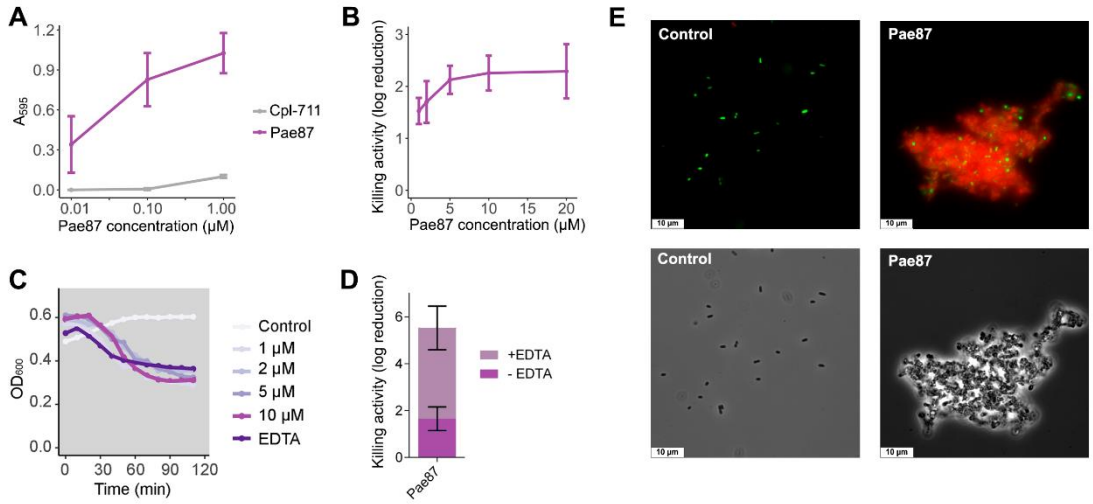


Figure 4.2.7. Pae87 activity on *P. aeruginosa* PAO1. A: Muralytic activity of different Pae87 concentrations against RBB-labelled purified *P. aeruginosa* PAO1 sacculi. **B**: Viability decrease of *P. aeruginosa* PAO1 cell suspensions treated with different concentrations of Pae87 for 2 h at 37 °C with respect to an untreated control. **C**: Turbidity decrease of *P. aeruginosa* PAO1 cell suspensions treated with different concentrations of Pae87 and with 10 μM of Pae87 + 0.5 mM EDTA. Representative results are shown. **D**: Comparative bactericidal effect of 10 μM of Pae87 on a PAO1 suspension in the presence of 0.5 mM EDTA. **E**: Fluorescence microscopy images of the PAO1 cell suspensions either untreated or treated with 10 μM of Pae87 and stained with SYTO9/PI. White bars at the lower-left corner indicate the scale (10 μm).

Fluorescence microscopy observations showed a clear cell aggregation effect induced by the Pae87 treatment as compared with the control (**Figure 4.2.7E**). This aggregation was accompanied by red fluorescence due to PI, indicative of a compromised cell surface and, possibly, as hinted by the viable cell counts, bacterial death. Also, the cells in the aggregates presented abnormal morphologies (such as rounder shapes or opacity loss when observed by phase-contrast microscopy), which altogether indicate that they are more cell debris masses than clusters of relatively healthy cells. The ability of the enzymes to keep the cells and cell debris bound together would explain the fact that a viability reduction was observed without a proper correlation with generalized lysis since cell disintegration did not occur. Actually, cell aggregation has been previously observed when testing bacterial surface-active proteins (165). This may be due to a general protein surface propensity to interact with the bacterial envelope, or, perhaps, to the existence within the protein of several affinity contact points with the cell wall. In this way, when the antibacterial proteins coat the surface of a bacterium, such a protein layer may recruit adjacent bacteria. Then, the surface interaction of Pae87 would potentially cause

membrane disruption and/or peptidoglycan degradation due to its catalytic activities, resulting in the cellular damage evidenced by **Figure 4.2.7**. The aggregation could also be the reason for the saturation of the dose-response curve in **Figure 4.2.7B**, since lysin molecules entrapped within the aggregates may not be able to attack other bacterial cells in suspension. Nonetheless, the aggregative effect, if conserved *in vivo*, could also be a therapeutic asset. Aggregation would probably prevent the dissemination of proinflammatory lysis by-products, and cellular aggregates have been shown, on occasions, to be better recognized and cleared (phagocyted) by the immune system cells (235, 357).

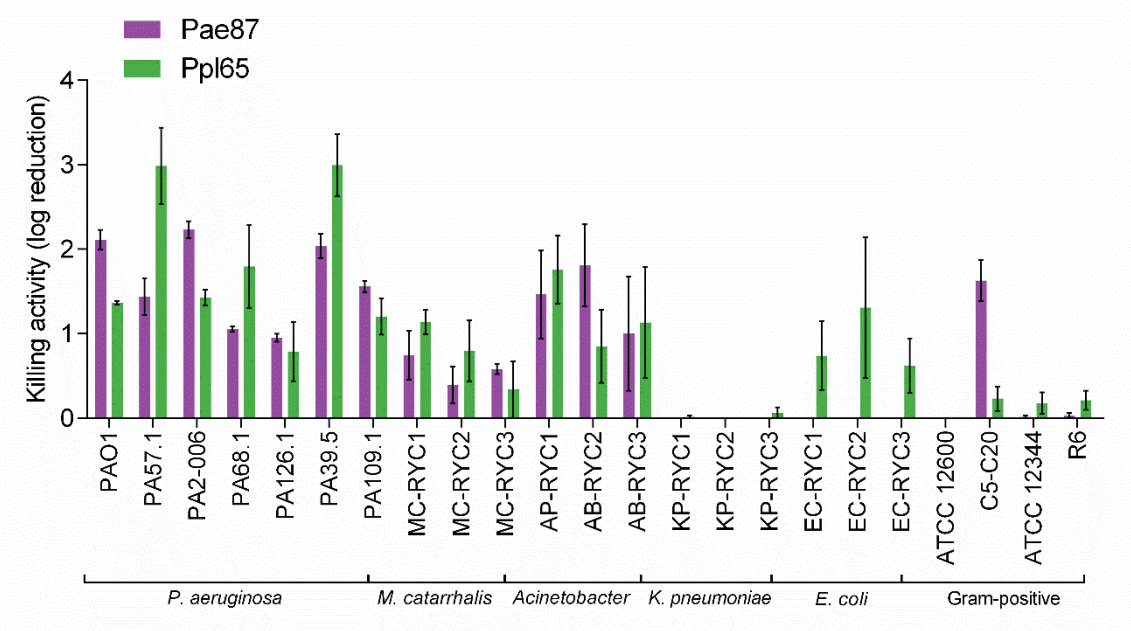


Figure 4.2.8. Bactericidal activity spectrum of Pae87 and Ppl65. The decrease in the viable cell counts of bacterial suspensions treated with 10 μM enzyme with respect to an untreated control is shown.

Finally, the bactericidal activity range of Pa87 was assessed (**Figure 4.2.8**). For comparison, the activity range of Ppl65 (endolysin of the PPpW-3 phage, **Table 4.2.2**) is also provided for comparison (255). Pae87 presented a semi-broad range activity spectrum against a collection of relevant respiratory pathogens. It displayed generalized activity against all of the *P. aeruginosa* strains tested (1-3 log kill), mild activity (≤ 1 log kill) against *M. catarrhalis*, greater activity (1-2 log kill) against *Acinetobacter*, and no activity against *K. pneumoniae* or the G+ strains in general, although it was found that the SMG strain tested was susceptible to Pae87. In comparison, Ppl65 had a similar spectrum, but somewhat broader, since some activity was found against *E. coli* strains. The SMG strain, however, was not attacked by Ppl65. The similarity in the tendencies across

Pae87 and Ppl65 activity ranges, on the other hand, points out to an analogy in their mechanisms of action.

6. STRUCTURAL-FUNCTIONAL STUDIES OF PAE87

The next aim was to elucidate the mechanism by which Pae87 interacts with susceptible bacterial cells. To this end, Mateo Seoane-Blanco in the lab of Mark J. van Raaij (CNB-CSIC) crystallized the protein, obtained X-ray diffraction data at 1.22 Å resolution, and constructed a high-quality, three-dimensional structure model of Pae87.

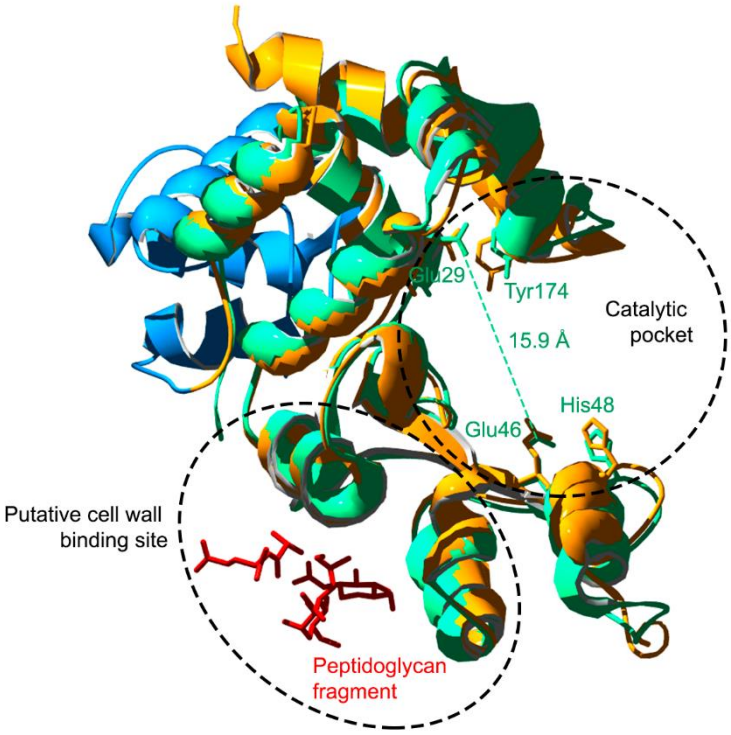


Figure 4.2.9. Structural model of Pae87 compared with AP3gp15. The 3D model of Pae87 (green) is shown superimposed on AP3gp15 structure (yellow, EAD; blue, CWBD). Relevant features are highlighted, such as the putative catalytic and substrate binding sites and conserved residues at the catalytic pocket.

The 3D structure of Pae87 was quite similar to that of the only experimental structure precedent of *Muramidase* PF family, AP3gp15 (PDB 5NM7), except for Pae87 lacking the CWBD that is present in AP3gp15 (**Figure 4.2.9**; RMSD~1.1 Å over 187 residues, 50% identity, Z-factor = 26.3, according to a structural pairwise combination at the Dali server). Pae87 was crystallized in the presence of purified PAO1 peptidoglycan previously solubilized by the action of the same enzyme. By co-crystalizing the enzyme and

substrate, the resulting 3D model showed that a fragment of peptidoglycan (NAG-MurNAc-L-Ala-D-Glu) appeared not bound into the presumed catalytic pocket, as it would be expected, but rather to a small cleft on the opposite side to the catalytic centre (**Figure 4.2.9**). A subsequent interaction analysis performed with QtPISA (358) revealed the specific residues involved in the interaction with the peptidoglycan fragment (**Table 4.2.3**, M. Seoane, personal communication).

Table 4.2.3. Total number of sidechain interactions of the respective Pae87 residues with the peptidoglycan fragment ligand.

	NAG	MurNAc	L-Ala	D-Glu
Gln79	2			
His82	1 (1)			
Glu83	1			
Asp86	1 (1)	1 (1)		
Arg93	2	5 (3)		
Gln121		2 (1)		2 (2)
Asn125				2 (2)
Tyr128	1 (1)			

Numbers between brackets indicate interactions that occur directly (not through a water molecule).

An MSA analysis of the peptidoglycan binding site using all the *Muramidase* entries of $\mathbb{S}^{LYS}$ (removing two low-coverage examples, YP_009639957.1 and YP_337984) revealed that the residues relevant for the peptidoglycan binding were much more conserved in those lysins bearing a single *Muramidase* EAD domain than in those which had an additional CWBD (**Figure 4.2.10A**, AKU43570.1 and YP_009595839.1 were reclassified as non-CWBD bearing lysins since they contained an unidentified N-terminal end that could probably be a still undescribed CWBD). Specifically, in the case of Tyr128 and Arg93, together with Gln121 and Asn125, both the aa frequency[8] and the relative BLOSUM62 score (which is a similarity metric) were close to the maximum value, 1.0, for the non-CWBD bearing entries (**Figure 4.2.10BC**). For the bimodular lysins of the *Muramidase* family, the scores were below 0.4 and 0.2 for Tyr128 and Arg93, respectively.

[8] The relative frequency at which the alluded residue appears at the said position within the MSA.

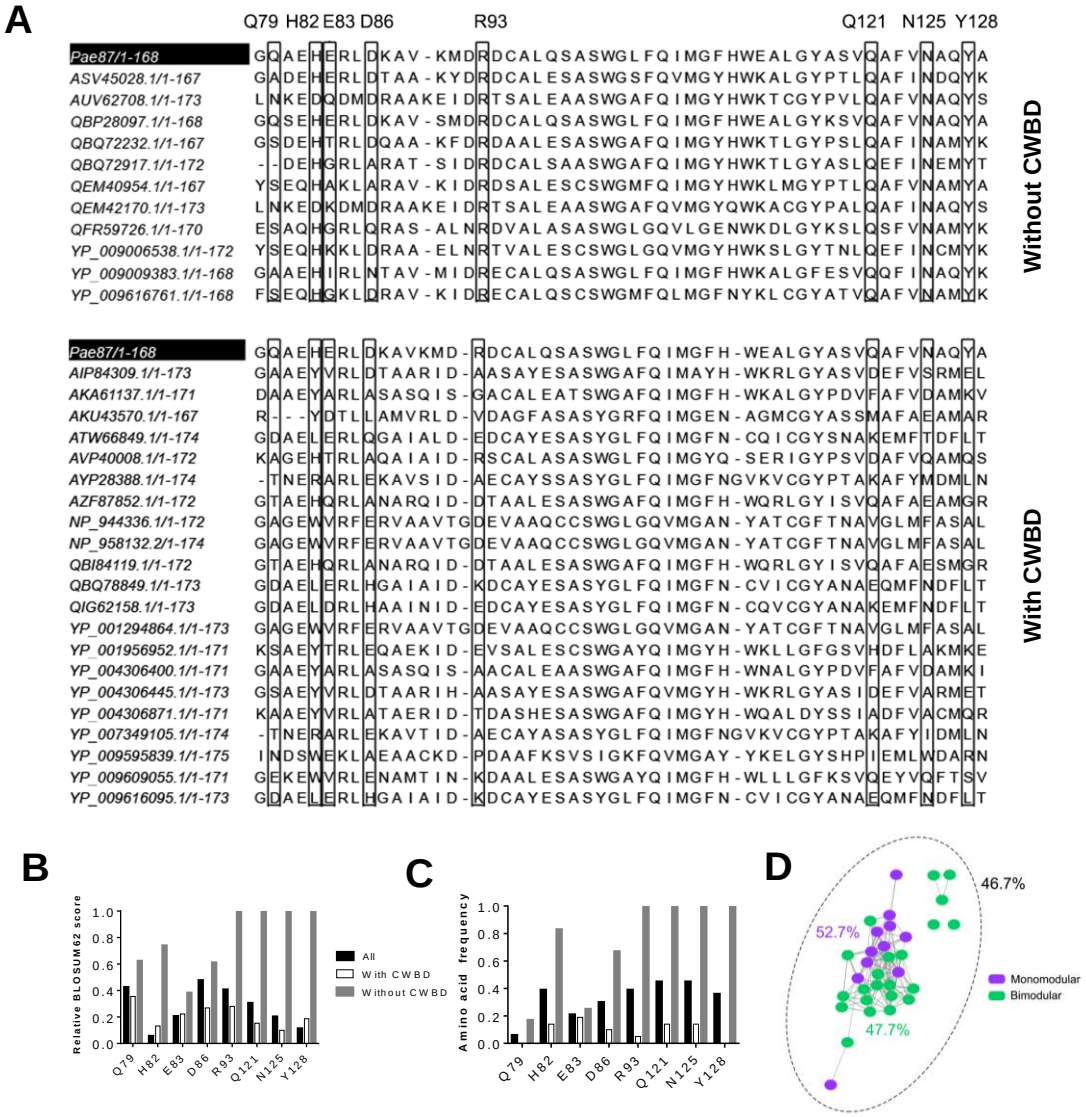


Figure 4.2.10. Multiple sequence alignments of the peptidoglycan binding site of *Muramidase* family lysins in $\mathbb{S}^{LYS}$. A: MSA of the peptidoglycan binding site of *Muramidase* family lysins grouped by the presence or absence of an N-terminal CWBD. Residue coordinates indicated at the top are those of Pae87. **B, C**: Residue conservation metrics for the residues putatively involved in contacts with peptidoglycan at the peptidoglycan binding site according to the Pae87 3D model (**B**: relative BLOSUM62 score, **C**: relative frequency of the Pae87 residue across the MSAs in each respective position). **D**: SSN of the 33 sequences of the MSAs distinguished by the presence or absence of a CWBD. Percentages are average identities for each group.

The SSN in **Figure 4.2.10D** shows that no sequence identity bias was introduced by the classification since the average identity percentage both within each subgroup and in the whole collection of sequences was around 46-52% and there was no apparent clustering. From these results, the existence of a cell wall binding site within the very catalytic domain can be proposed, whose evolutionary purpose could be to compensate for the lack of a proper, independent CWBD, given that it only appears in the examples without a CWBD. Although most enzymes contain residues within their catalytic clefts whose purpose is to interact with the substrate in such a way that it is correctly positioned

for catalysis, there is, to our knowledge, just a single example in literature of a lysin containing a cell wall binding site within the previously described catalytic domain outside the cleft. It is the case of *Bacillus* lysin PlyG, which contains a region located in the NAM-amidase catalytic domain that specifically recognizes *Bacillus anthracis* spores (359). However, it is hardly a comparable case, since for PlyG the secondary substrate binding site seems to have evolved to recognize a chemically distinct form of the target bacteria that cannot be recognized by the canonical CWBD. In contrast, for Pae87 and its close relatives, it could be argued that a region with affinity to peptidoglycan has arisen to take over the function of a lacking CWBD.

Previous structure-function studies performed with AP3gp15 suggested that the catalytic pocket was rather different from that of other common lysozymes, such as hen egg-white lysozyme (HEWL) or goose egg-white lysozyme (GEWL) (173). A single glutamic acid has been pointed out to be involved in catalysis for AP3gp15. A corresponding Glu residue is also conserved in the Pae87 structure (Glu29, **Figure 4.2.9**) and in all of the 32 *Muramidase* sequence examples in $\mathbb{S}^{LYS}$. In the same structural region, there is a second Glu residue (Glu46, according to Pae87 notation) which is perfectly conserved across all of the *Muramidase* examples considered (**Figure 4.2.11**). This conservation strongly suggests that it plays a relevant role, possibly as a second catalytic residue. The cooperation of two acidic residues at the catalytic site has been extensively reported as the main known mechanism of catalysis for glycosidase enzymes (360). Other conserved residues facing the catalytic cleft were His48 and Tyr174, which, therefore, could also play a significant role in the catalytic pocket integrity and function (**Figure 4.2.11**). In fact, it has been described that Tyr (and Trp) residues are common within glycosidase active sites since their combination of hydrophobic character and an ability to form hydrogen bonds favour sugar binding (360). Aromatic side chains have also been shown to be pivotal in sugar binding through CH−π interactions (between the electropositive patches of sugar C−H bonds and the π electron cloud of the aromatic side-chain) (361). As for the His48 residue, it also bears catalytic potential, since it has been described that His may be involved in catalysis, either as a sole catalytic residue or in cooperation with carboxylic aa (362, 363).

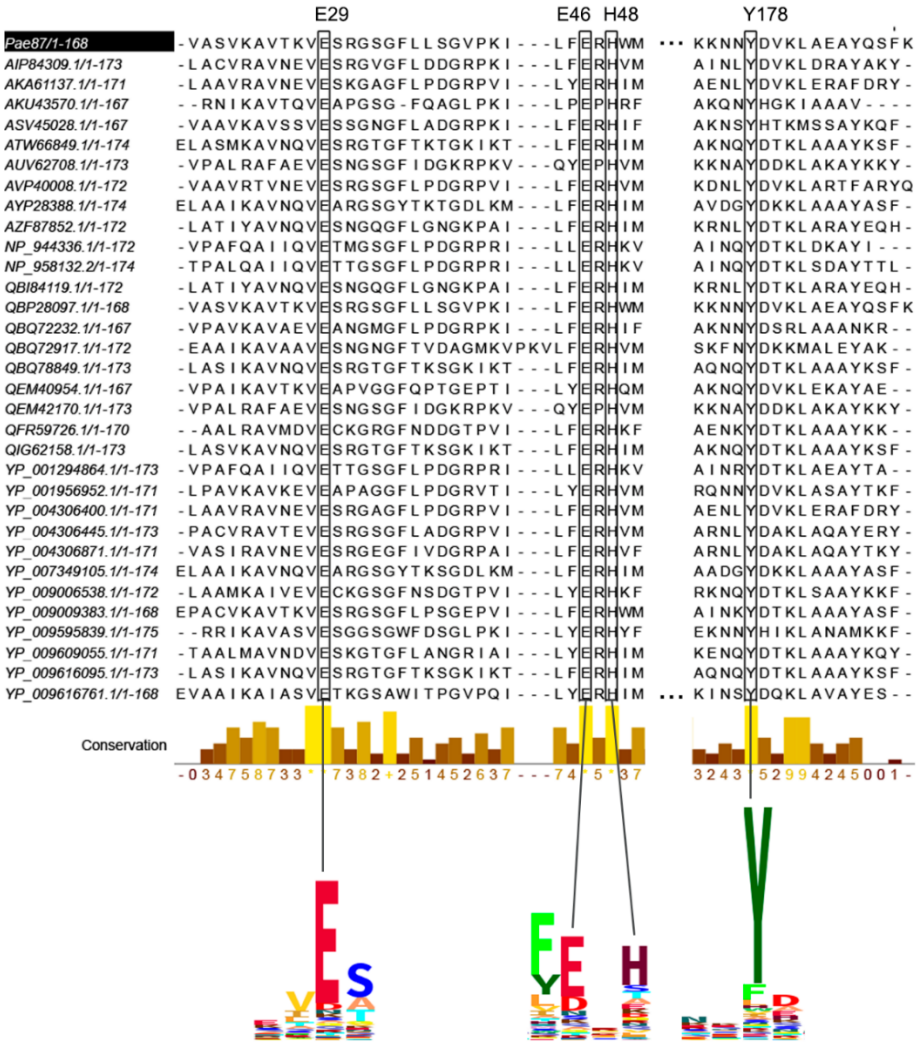


Figure 4.2.11. Multiple sequence alignment of *Muramidase* family domains in $\mathbb{S}^{LYS}$ showing the conservation of residues facing the catalytic pocket. Selected segments of the PF HMM logo of the family are displayed at the bottom. Relevant positions are connected to their respective columns in the MSA.

His48 may play its role by providing electrostatic stabilization of transition states or pKa value modulation within the active site. All of the aforementioned residues also presented high conservation levels within the PF HMM of the family, judging by the HMM logos (**Figure 4.2.11**), strengthening their presumptive functional role. Previous research on the *Muramidase* family catalytic center did not focus on many of the hereby reviewed conserved residues due to their approach, which relied on the structural comparison with known lysozymes. It was claimed, however, that the studied lysin, AP3gp15, should have an “unusual catalytic mechanism” due to the lack of suitable equivalent residues at the key positions of the presumed catalytic center (173).

Since the canonical catalytic mechanism for glycoside hydrolysis relies on a couple of carboxylic acids, a mutational analysis was conducted to confirm the catalytic implication of both Glu29 and Glu46. Pae87 single mutants E29A, E46A, and the double mutant, E29A/E46A, were constructed by overlap extension PCR and their muralytic and bactericidal activities were tested (**Figure 4.2.12**). Both residues, Glu29 and Glu46, were deemed relevant for the catalytic degradation of PAO1 peptidoglycan since all the mutants displayed a remarkable decrease in their cell wall solubilization ability when compared to wild type Pae87 (**Figure 4.2.12AB**). At the maximum non-saturating concentration (0.1 μM), Pae87 retained about 75% of its maximum activity detected, while the mutants displayed no peptidoglycan solubilization activity.

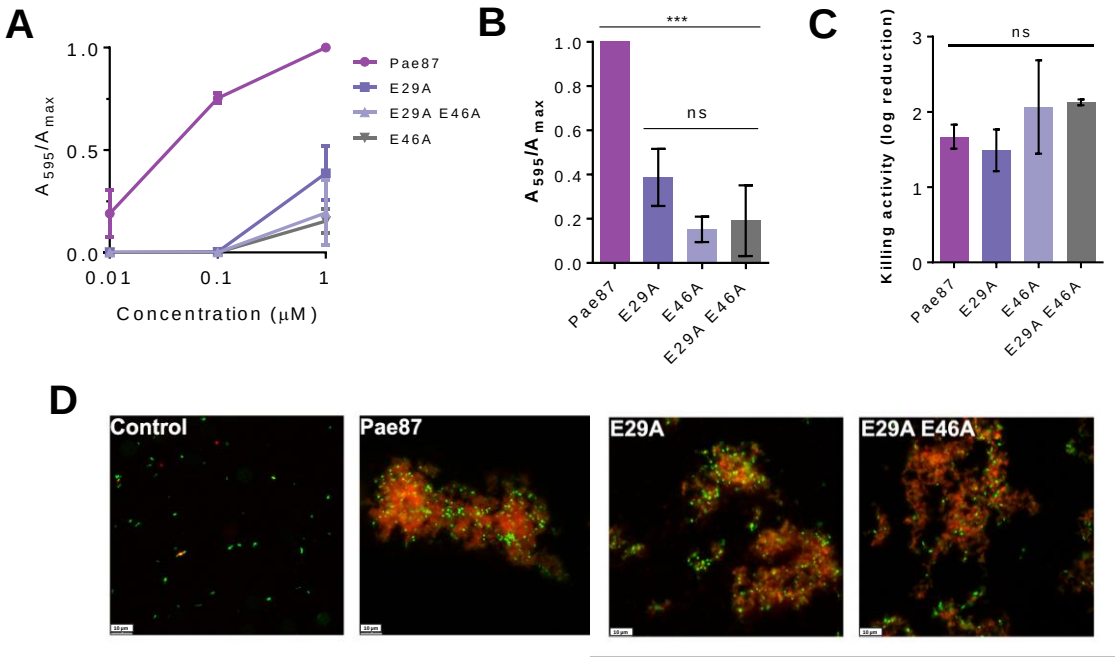


Figure 4.2.12. Activities of wild-type Pae87 and of mutants in which conserved catalytic site glutamic acid residues were mutated. **A**: Muralytic activity of Pae87 and its mutants on RBB labelled purified PAO1 cell walls at different concentrations, as explained in MATERIALS AND METHODS (**Section 6.1**). **B**: Comparison of muralytic activities at the maximum concentration tested (1 μM). **C**: Bactericidal activity of 10 μM of Pae87 and its mutants against bacterial suspensions of PAO1 (2 h, 37 °C). **D**: Fluorescence microscopy of some of the treated suspensions in **C**, dyed with SYTO9 and PI. White bars at the lower-left corner indicate the scale (10 μm). One-way ANOVA was used in **B** and **C** followed by Tukey post-test to perform an all-against-all multiple comparison (ns = non-significant difference; *** = $p \leq 0.01$).

The three mutants also retained some residual activity at the maximum concentration tested (less than 50% than the wild type enzyme), suggesting that there may be other residues involved in catalysis (perhaps His48 and/or Tyr174). It is also worth noting that, if we consider both Glu-29 and Glu46 to be involved in catalysis, as our results suggest, it would mean that the Pae87 catalytic cleft is unusually large since the average distance

between the side chain oxygens of the acidic residues in our 3D model is 15.9 ± 1.6 Å (**Figure 4.2.9**). For context, the typical O···O intercarboxylic distance has been recorded to be, in average, of 8.5 ± 2.0 Å in inverting β-glycosidases, while it is shorter (4.8 ± 0.3 Å or 6.4 ± 0.6 Å) in β-glycosidases that use the retaining mechanism (364). A putative conclusion from these observations is that the hydrolysis mechanism of Pae87 may operate with net inversion of the anomeric configuration. The size of the Pae87 catalytic cleft remains, however, remarkably large, which was also noted for the previously reported comparisons between AP3gp15 and other canonical lysozymes such as HEWL, GEWL, or T4, and it was suggested to imply an extended substrate-binding surface for *Muramidase* family enzymes when compared with other more typical glycosidases (173). It cannot be ruled out, however, that there may exist some hinge flexibility that could allow the two lobes, each with a Glu residue, to approach each other and perform the catalytic function with a more typical spatial configuration.

Regarding the catalytic activity of Pae87, the disaccharide found in the soluble product bound into the crystallized protein (NAG-MurNAc) already points out a lysozyme activity, as described in the literature for this family. The degradation products analysis by RP-HPLC coupled to mass spectrometry firstly showed that both Pae87 and the positive control (the lysozyme cellosyl) mobilized an array of soluble compounds, in contrast with the untreated blank (**Figure 4.2.13A**). When comparing the mass spectrometry chromatograms of cellosyl and Pae87, a coincidental pattern for the main degradation peaks was found (**Figure 4.2.13B**). This observation supports the catalytic nature of Pae87 as a muramidase. Moreover, the CID spectrum for one of the main peaks of Pae87-degraded peptidoglycan (the one eluted at 4.3 min) presents peaks compatible with the loss of a non-reduced NAG (−203.078 mass units) from a NAG-MurNAc-Ala-Glu-mDAP-Ala or a NAG-MurNAc-Ala-Glu-mDAP fragment (**Figure 4.2.13C**). Conversely, there is no evidence coherently compatible with losing a reduced NAG (−223.106). This serves as proof for a muramidase activity, which leaves a MurNAc end susceptible to be reduced. NAG is susceptible to reduction when the glycan strand is degraded by a glucosaminidase activity (326, 365).

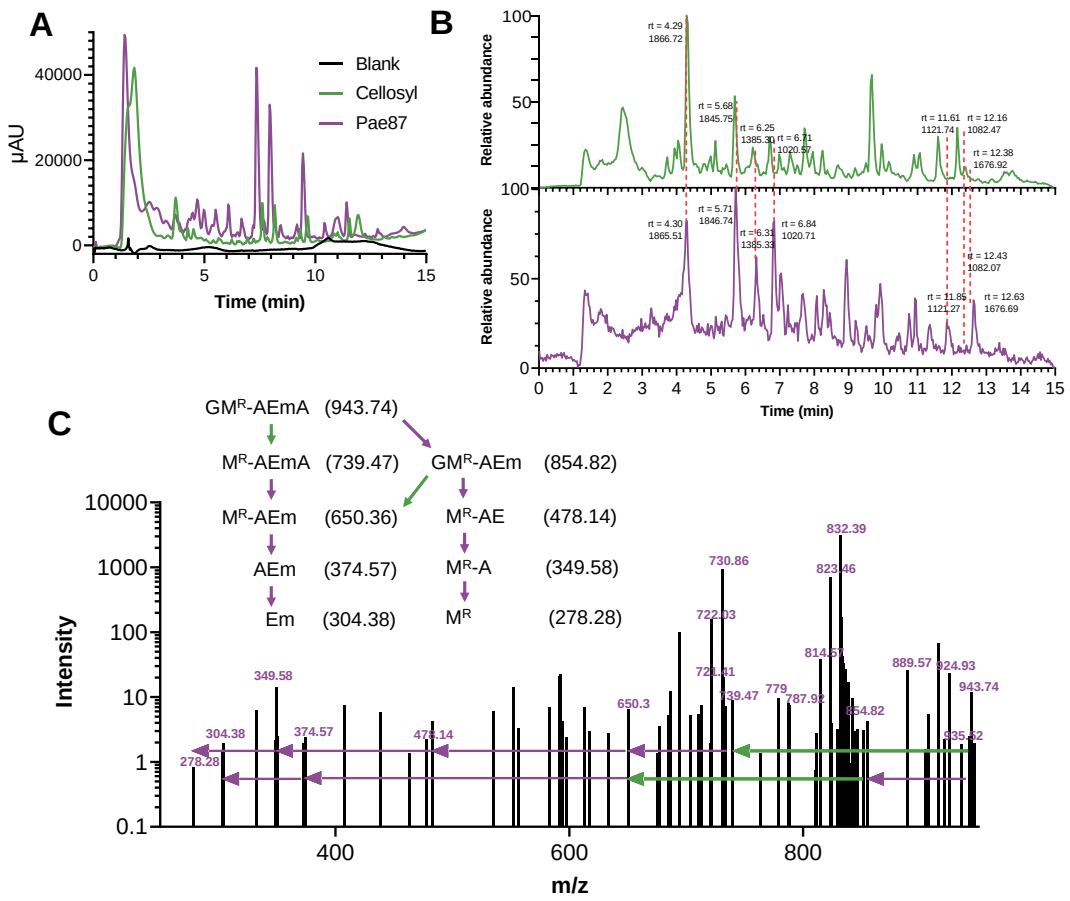


Figure 4.2.13. Analysis of the degradation products of Pae87 activity on *P. aeruginosa* PAO1 peptidoglycan. A: UV (204 nm) chromatograms of PAO1 peptidoglycan solubilized with either cellosyl or Pae87 and of the soluble fraction of an untreated sample. **B**: Liquid chromatography-mass spectrometry chromatograms of the degradation products of cellosyl (green) and Pae87 (purple) activities on PAO1 peptidoglycan. Retention time (rt) and a representative monoisotopic mass value are shown for selected peaks. **C**: CID spectrum of the rt = 4.30 min peak of the Pae87 peptidoglycan degradation. Selected m/z values are displayed. The dissociation of a GM^R-AEmA fragment is presented. G = nonreduced NAG; M^R = reduced MurNAc; A = alanine; E = glutamic acid; m = *meso*-diaminopimelic acid.

On the other hand, there were no significant differences in the observed bactericidal activity against PAO1 between Pae87 and its mutants or in the fluorescence microscopy images (Figure 4.2.12CD). This may mean that the main effective mechanism in the Pae87 antibacterial activity would be a noncatalytic one. However, the mutation of the negatively charged, exposed Glu residues into Ala would increase the surface positive net charge, and such change may improve the affinity of the protein for the bacterial surface, perhaps making up for the loss of the catalytic activity up to some point. This observation does not necessarily downplay the conclusion that most or some of the Pae87 activity is due to noncatalytic interactions, but rather provides a further explanation to the nearly identical bactericidal effect observed for the mutants with respect to the parental enzyme.

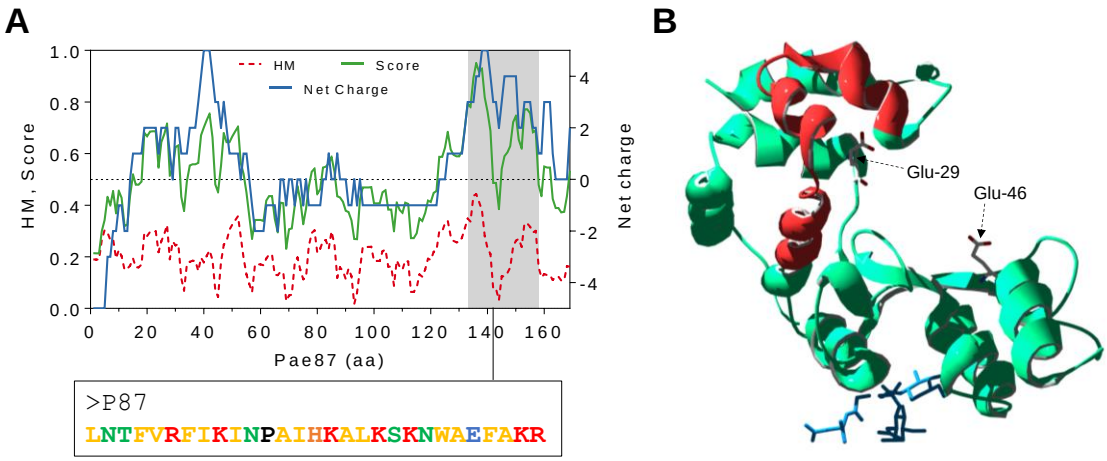


Figure 4.2.14. Definition and localization of P87 peptide within Pae87. A: Physicochemical profile of Pae87, depicting HM and net charge in each of the 11-aa windows of the protein, as well as the score used to measure both of the properties. P87 sequence is also shown code-coloured regarding the properties of each residue (yellow, nonpolar; green, polar without charge; red, positively charged; black, proline; blue, negatively charged). **B**: 3D model of Pae87 with peptide P87 highlighted in red. Catalytic Glu residues and peptidoglycan ligand (in blue) are displayed as spatial references.

The leading hypothesis in this work is that the noncatalytic activity of Pae87 may reside in the previously detected C-terminal AMP. A closer examination of the physicochemical profile of Pae87 allowed us to postulate a specific 29-aa peptide, hereinafter named P87, which would make up the AMP-like region (**Figure 4.2.14**). The definition of P87 was based on the concurrent maximization of net charge and HM that takes place in such coordinates of Pae87. For measuring said properties simultaneously, a score variable was calculated consisting in the half-sum of the min-max standardized net charge and HM for each 11-aa fragment of Pae87 (**Equation 1**):

$$\frac{1}{2}\left(\frac{HM_i - \min(HM)}{\max(HM) - \min(HM)}\right) + \frac{1}{2}\left(\frac{NC_i - \min(NC)}{\max(NC) - \min(NC)}\right) \quad \textbf{Equation 1}$$

Where HM_i and NC_i are specific values of HM or net charge, respectively, and HM and NC the respective sets of values for each variable. In this way, the score measured the magnitude of both parameters with equal weight and on a 0-1 scale. The P87 sequence contains seven positively charged Arg or Lys residues (and a negatively charged Glu), interspersed with mostly nonpolar residues. This structure, together with the fact that it forms three alpha-helices within the Pae87 structure (**Figure 4.2.14**), suggests that it may form amphipathic helices with membrane-interaction potential. Nonpolar residues of P87 are mostly buried within Pae87, but not in all cases. For example, Ile144 or even Phe141 are especially exposed. Lys and Arg residues, on the other hand, are

located at the outer surface of the protein, therefore available for electrostatic interaction with negatively charged elements of the bacterial surface (namely, the phosphate groups of the lipopolysaccharide).

Peptide P87 was synthesized and provided by GenScript for further analysis of its antimicrobial properties. Circular dichroism spectra in the presence of increasing concentrations of TFE provided evidence on the ability of P87 to form amphipathic helices in the presence of biological membranes (**Figure 4.2.15A**).

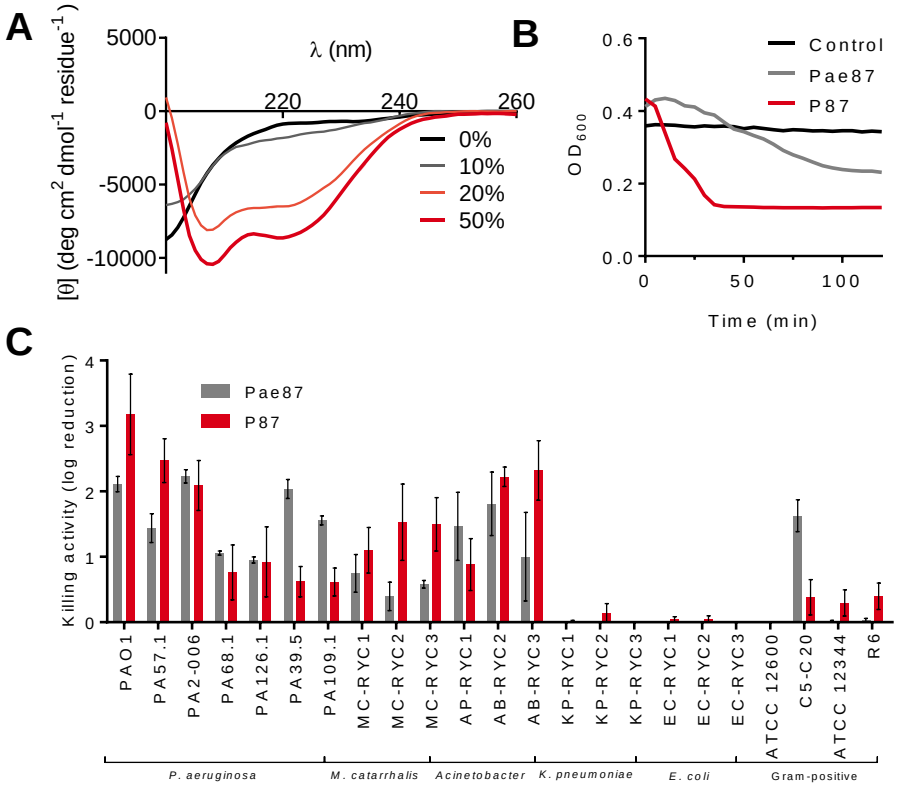


Figure 4.2.15. Antimicrobial activity of P87 peptide. A: Far UV **c**ircular dichroism spectra of 20 μM P87 in 20 mM NaPiB, pH 6.0, with 100 mM NaCl and different v/v concentrations of TFE (as indicated in the legend). **B**: Turbidity decrease assay of a 10 μM P87 or Pae87 treatment to PAO1 cell suspensions. **C**: Bactericidal activity range of 10 μM P87. Pae87 bactericidal range from **Figure 4.2.8** is also shown for comparison.

P87 presented an unordered conformation in aqueous solution, but, beginning at 20% TFE, it started to shift towards an α-helix conformation, as evidenced by the typical peaks at, roughly, 222 nm and 208 nm. It is thus presumed that the acute lytic effect that P87 causes when in contact with a PAO1 bacterial suspension is due to the formation of amphipathic helices induced by the presence of the cell membranes (**Figure 4.2.15B**). Such helices could then insert into the membranes in such a way that leakage of the cell content occurs, leading to osmotic shock and lysis. It differs, in this regard, with the poor

lytic outcome of Pae87 treatment (**Figure 4.2.15B**). This difference may be due to the much smaller size of P87, which could enable it to properly insert into the membranes rather than interact superficially, as it is assumed that Pae87 does. Peptide P87 was also able to kill PAO1 and other bacteria in a range and magnitude similar to Pae87 (**Figure 4.2.15C**). The antimicrobial capacity of P87 is thus considered proven, and therefore a role in the Pae87-bacterial surface interaction can be proposed with supporting evidence.

7. Determination of the bactericidal mechanism of Pae87 and P87

Since the OM has traditionally been labeled as the main obstacle when attacking G− bacteria from without, a fluorescent probe was used to detect OM permeabilization induced by Pae87 and its derivatives (*i.e.*, the noncatalytic mutants and the antimicrobial peptide P87). NPN is a hydrophobic molecule that fluoresces when reaches the phospholipid layer of the inner membrane upon OM permeabilization. Typically, NPN uptake kinetics are recorded with short incubation times (with a maximum of 10 min and a minimum of 3 min) (366, 367). However, we extended the incubation period up to 2 h due to the increasing fluorescence values of Pae87 kinetics over such time (**Figure 4.2.16A**). At the average kinetic estimations shown in **Figure 4.2.16A**, two different tendencies can be observed: the 'canonical' OM permeabilization peak induced by P87 in the first minutes after peptide addition, and the slow but steady increase in fluorescence for Pae87 and its mutants. For the sake of statistical comparison, the average values of NPN fluorescence reached 5 min after reagent addition were considered (**Figure 4.2.16B**). At such a short period, the NPN fluorescence of Pae87 and its noncatalytic mutants is still low, although, for Pae87, significantly different than the control. The fluorescence induced by P87 was above that obtained for EDTA positive control. Fluorescence microscopy images at the end of incubation visually confirmed the damage to the OM for all the compounds tested (**Figure 4.2.16C**). The first conclusion drawn is that both the peptide and the proteins can permeabilize the OM, although with remarkably different kinetics. Also, catalytic activity is not required for OM permeabilization, as it would be expected. A fluorescent version of Pae87 labeled with Alexa488 (see Materials and Methods **Section 5.4** for details) was used in combination with PI, a DNA-intercalating fluorescent probe. In this way, the temporal localization of Pae87 and its effect on a suspension of PAO1 cells were traced (**Figure 4.2.17**).

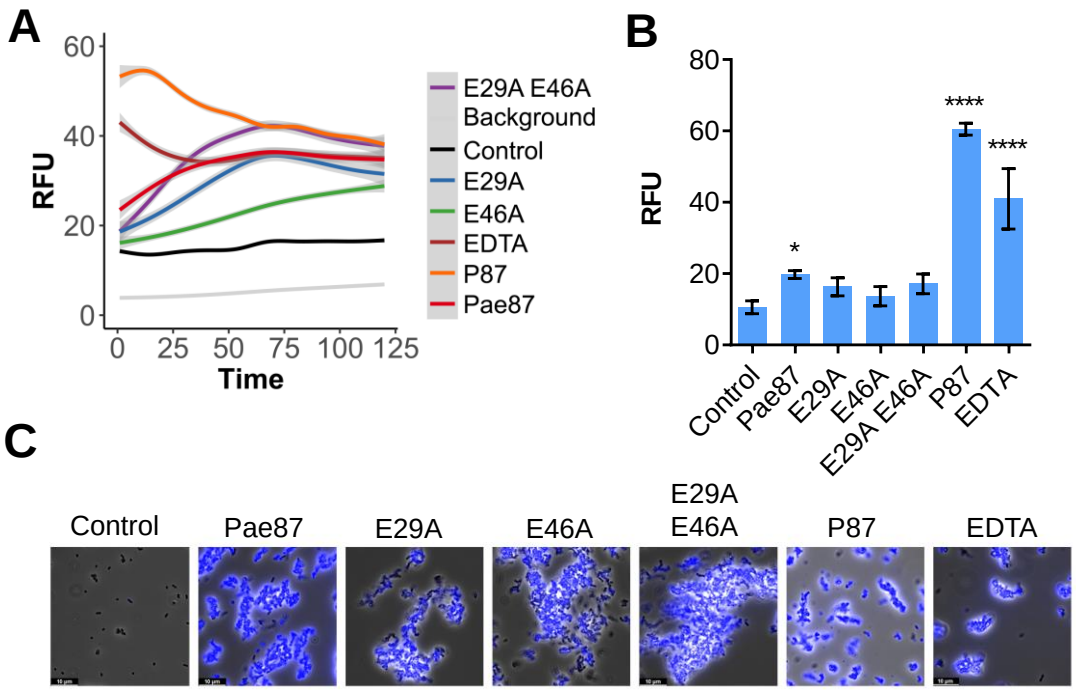


Figure 4.2.16. OM permeabilization assays with fluorescent probe NPN. A: GAM estimation of the average tendencies of the NPN fluorescence kinetics (excitation wavelength was 350 nm, emission was recorded at 420 nm). Estimation was based on three independent replicates, mean estimation ± 95% C.I. (grey shade) is shown for each experimental condition. **B**: Comparison of the average NPN signal (minus background fluorescence) after 5 min incubation. Mean ± sd of three independent replicates is shown; a one-way ANOVA test with Dunnett post-test was applied to statistically compare each condition with the control (untreated cells in the presence of the probe). *, $p \leq 0.05$; ****, $p \leq 0.0001$. **C**: NPN fluorescence microscopy observation of each experimental condition after 2 h incubation. Superimpositions of phase-contrast images with blue fluorescence signal observed with an A filter cube (excitation bandpass 340-380) are shown. RFU = fluorescence units relative to the maximum value achieved during the assay.

At t_0, Pae87 already began binding to the bacterial surface, as green fluorescence rims were observed around the *P. aeruginosa* cells (**Figure 4.2.17A**). Over time, the Pae87 molecules bound to the cell walls promoted an increasing aggregation among close bacteria. According to the visual estimations of the areas of such aggregates, their maximum size was reached approximately after 30-60 min under the assay conditions (**Figure 4.2.17B**). Also, at t_0, a discrete intracellular spot of red fluorescence (PI) was observed. Afterwards, PI fluorescence underwent a gradual diffusion concomitant to the aggregation. At 60 min of incubation, red fluorescence appeared as a halo around the aggregates, with a distorted bacterial morphology. Altogether, the conclusion is that the enzyme promotes aggregation as well as damage to the cell surface, with DNA leakage ensued, and with no dispersion nor disintegration of the cell debris. Therefore, these observations, together with the NPN fluorescence results, support the mechanism of action of Pae87 proposed in **Section 5** of this chapter with the occasion of the results presented in **Figure 4.2.7**.

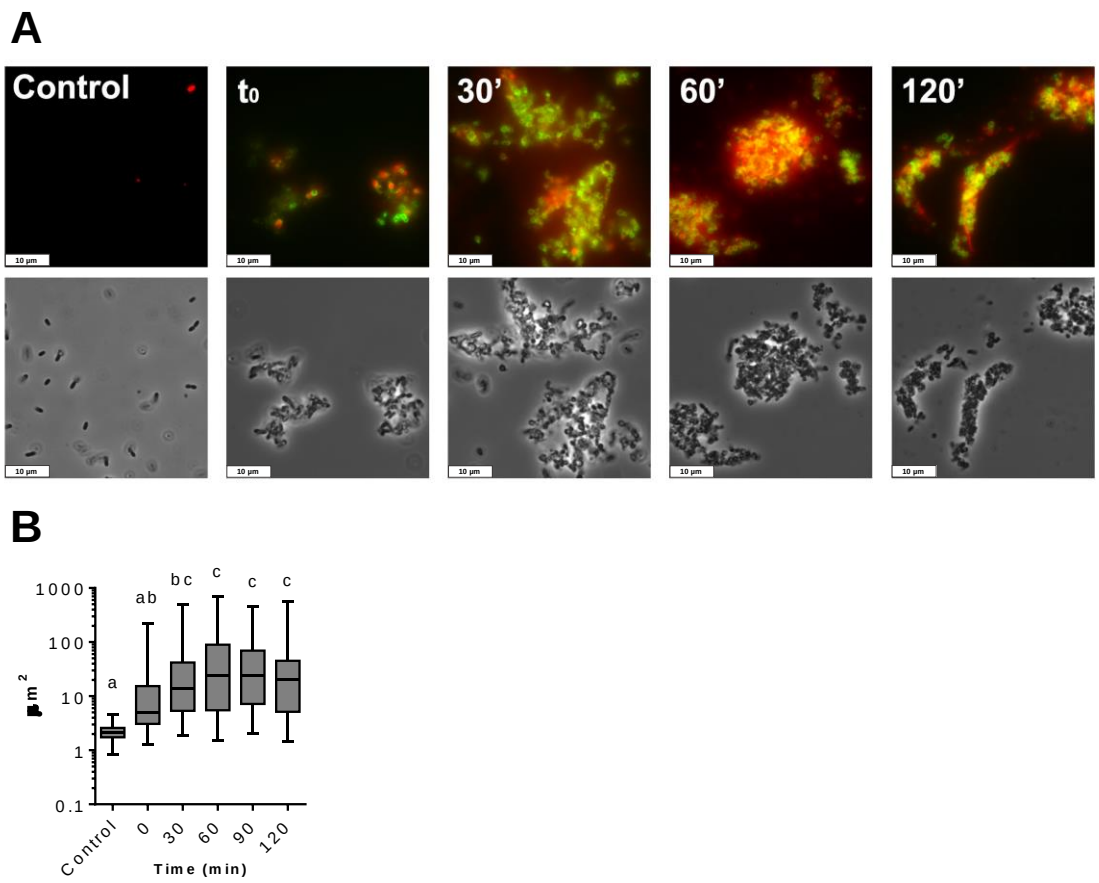


Figure 4.2.17. Microscopic observation of a PAO1 suspension treated with Pae87 over time. A: Fluorescence and phase-contrast images representative of the observations made along 2 h of incubation (37 °C, 20 mM NaPiB, pH 6.0, 150 mM NaCl) of a PAO1 suspension (≈10^8 CFU/ml) treated with 10 μM of Alexa488-labelled Pae87 and stained with PI. White bars indicate 10 μm. **B:** Time-wise distributions of cells or aggregates areas estimated using LAS X microscopy image analysis software in, at least, ten frames per time point. A one-way ANOVA followed by Tukey's post-test was applied for multiple comparisons. Distributions marked with different letters are significantly different from each other, while, between those indicated with the same letter, no significant differences were found.

8. Influence of different reaction conditions on Pae87 and P87 bactericidal activity

Several incubation conditions were examined to check their importance in Pae87 and P87 killing activity (**Figure 4.2.18**). The peptide:bacteria stoichiometry was critical for the killing efficacy of P87 peptide and, to a lesser extent, for Pae87 (**Figure 4.2.18A**).

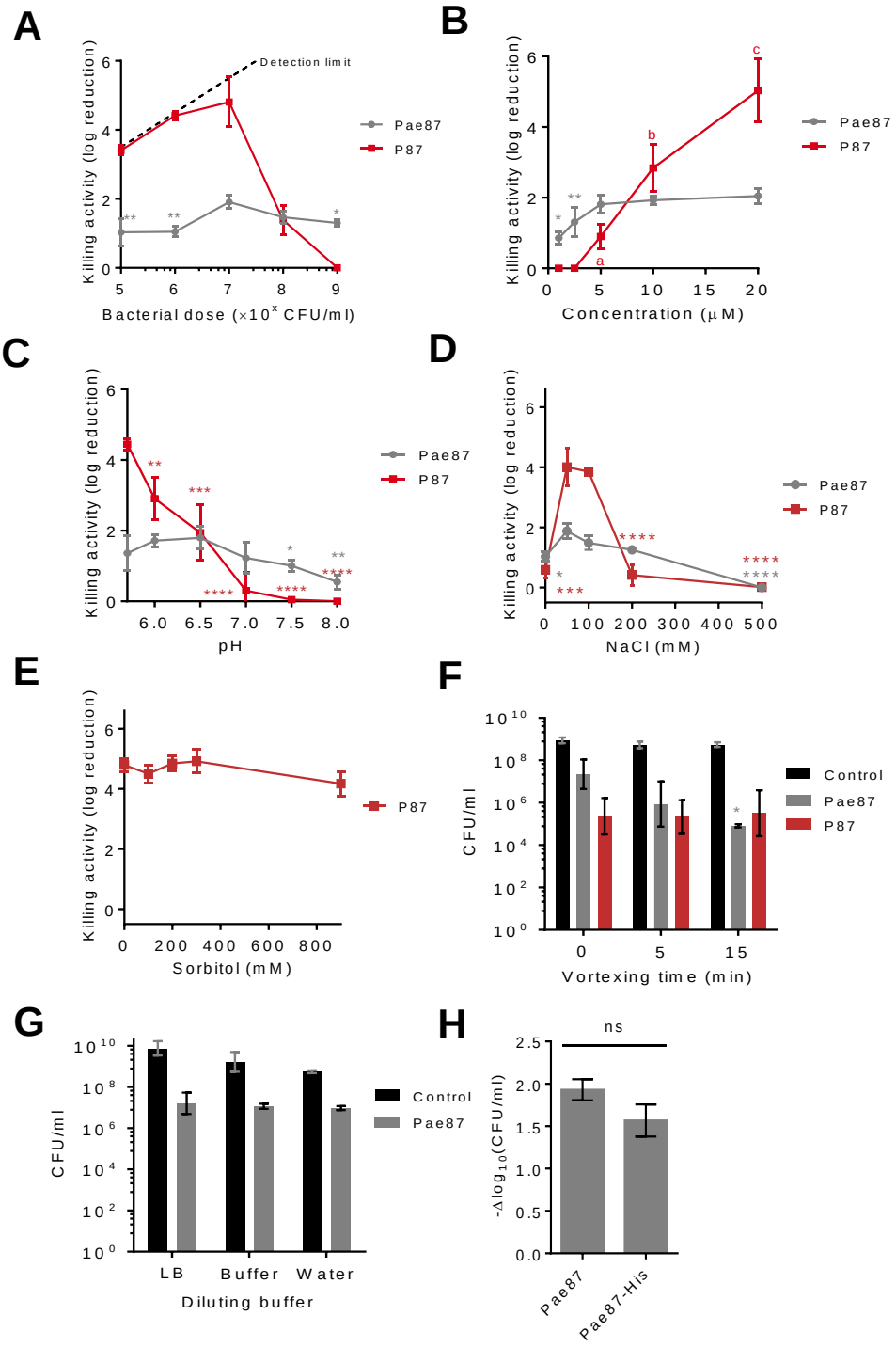


Figure 4.2.18. Parameters of Pae87 and P87 bactericidal activity on *P. aeruginosa* PAO1. **A**: Enzyme:bacteria or peptide:bacteria stoichiometry. **B:** Dose-response curves. **C**: Variation of killing activity with pH. **D**: Variation of killing activity with ionic strength. **E**: Variation of killing activity with the concentration of a non-ionic osmolyte (sorbitol), maintaining a fixed concentration of 50 mM NaCl. **F**: Viable counts after a 0, 5, or 15 min vortexing. **G**: Effect of the nature of the diluent for plating. **H**: Role of His-tag in the observed Pae87 activity (Pae87-His = Pae87 lacking the His-tag). One-way ANOVA with Dunnett post-test was applied to compare with the control or the highest value condition (**A**: 10^7 CFU/ml; **B**: 20 μM; **C**: 6.5; **D**: 50 or 150 mM; **E**: 0 mM sorbitol; **F**: the corresponding viable count without vortex treatment, **H**: Pae87 treatment). Unless otherwise stated, the incubation conditions were 20 mM NaPiB, pH 6.0, 150 mM NaCl, ≈10^8 CFU/ml PAO1, 10 μM of the bactericidal compound, 37 °C, 2 h. Asterisks indicate *p*-values of significant comparisons (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$), non-significant comparisons are not indicated. When marked by letters, an all-against-all multiple comparison was applied with ANOVA plus Tukey post-test. Different letters indicate significantly different results.

In both cases, the optimal ratio was achieved at around 10^7 CFU/ml for 10 μM Pae87 or P87, which roughly corresponds to some 10^{10} molecules per cell. At 10^7 CFU/ml and below, the maximum detectable killing was observed for P87, while higher bacterial doses reduced the bactericidal effect, surely due to a sub-optimal number of antimicrobial molecules per cell. P87 maximum activity at 10 μM, against 10^7 CFU/ml PAO1, was above 4-log kill, while with increasing bacterial concentrations it decays to nil activity at 10^9 CFU/ml. This is in agreement with a cooperative mechanism of action where a threshold number of bound peptides is required for the bactericidal activity. As for the dose-response curves (**Figure 4.2.18B**), the one of Pae87 quickly saturated, as it was shown before in **Figure 4.2.7**. This was speculated to be due to the entrapment of Pae87 molecules within the aggregates. P87 presents a more canonical curve, in line with the effect of the peptide:bacteria ratio, with an increasing activity with higher concentrations (up to 5-log kill at the maximum concentration tested). pH does not seem a major determinant of Pae87 activity (**Figure 4.2.18C**), although the best killing numbers were obtained at mildly acidic pH, as it was for P87 too. This could be explained due to the higher positive charge that both protein and peptide may have at acidic pHs, improving interaction with the negatively charged bacterial surface. Given that the activity increase was observed at pH 6.0 and below, protonation of histidine residues is the most plausible explanation, at least in the case of P87. Although P87 was almost inactive at near-physiological pH (≈ 7.5), the fact that it was highly effective at acidic pH is relevant for infection treatment, since it has been many times suggested that the pH at the infection site is acidified by a combination of bacterial metabolic activity and immune system responses (368, 369). This acidification is especially relevant in certain conditions, such as cystic fibrosis, whose patients are already infection-prone at the respiratory tract, with *P. aeruginosa* being one of the main causative agents of cystic fibrosis exacerbation (370). Ionic strength did not have a remarkable impact at near-physiological concentrations (50-200 mM NaCl) for Pae87, but P87 only displayed a relevant killing activity between 50 and 150 mM (**Figure 4.2.18D**). It is possible that a slight salt concentration might be necessary for P87 proper solubility (in fact, the Aggrescan server (371) detects an N-terminal hotspot for aggregation), but that higher concentrations might shield charged residues. This latter hypothesis is supported by the results in **Figure 4.2.18E**, in which the activity of 10 μM P87 was tested in the presence of different concentrations of a non-ionic osmolyte (sorbitol). In such results, the killing activity of the peptide was not affected by the increasing concentrations of the non-ionic solute. The 2 h-treated PAO1 suspensions were subjected to different vortexing treatments before plating, to check whether the observed aggregative effect was underestimating the viable cell counts (**Figure 4.2.18F**). The vortex treatment would disintegrate the aggregates and thus release viable cells

that, grouped within the aggregates, might have produced a single colony. The results in **Figure 4.2.18F**, however, suggest that this is not the case, *i.e.*, the observed decrease in viable cell counts cannot be solely attributed to the mere aggregation, and thus a bactericidal mechanism must be occurring. This is because no significant differences were found in the viable counts after vortexing with respect to the non-vortexed samples, except with the samples treated with Pae87 and vortexed for 15 min. In this latter case, the cell count in fact decreased over the control without vortexing. This is in agreement with our previous proposal that aggregates actually comprise damaged cells and, thus, the mechanical shaking would increase the apparent killing by definitely harming these already damaged bacteria. Of note, when observing the vortexed samples under the microscope, while the control cells appeared clearly unharmed, a mostly blank background was observed for the Pae87-treated ones (data not shown), further supporting the aforementioned interpretation of this experiment. It has been previously shown that a high turgor pressure (*i.e.*, using a low tonicity reaction buffer or using a high tonicity reaction buffer and then applying an osmotic shock by diluting in a low tonicity buffer) can be crucial to enable the efficient killing activity of an inherently active anti-G− lysin (372). This was claimed to be due to a sub-lethal peptidoglycan degradation effect. We tested this hypothesis by diluting the PAO1 cells treated in the standard buffer with 150 mM NaCl in different media: LB, the same high tonicity buffer (20 mM NaPiB, pH 6.0, 150 mM NaCl) and just distilled water (which was the usual diluent in the other assays). Our results show that turgor pressure has not a clear impact on the antibacterial activity itself: nor the low tonicity reaction buffer enhances activity (**Figure 4.2.18D**) nor does a higher tonicity of the diluent (LB or NaPiB plus 150 mM NaCl versus water) increase the bacterial count (**Figure 4.2.18.G**). The previously cited work, in which the importance of turgor pressure is discussed, presents an inherently active lysin ‘from without’ that does exert a measurable bacteriolytic effect, unlike Pae87, which implies that it does not bear such an aggregative effect as Pae87 (although no explicit evidence on this matter is provided in the article). The aggregative effect of Pae87 may enhance its contact with the bacteria and thus improve the lethal effect upon them, probably by increasing the contact time and adding a multivalent effect. In this way, the turgor pressure-independence of Pae87 may be explained. The activity of Pae87 on the bacterial surface and membranes can be due to or enhanced by the presence of the N-terminal His-tag, which can acquire a relevant positive charge when protonated at lower pHs. To test the involvement of the His-tag in the bactericidal activity, a non-His-tagged version of Pae87 was tested against *P. aeruginosa* PAO1 at equimolar concentrations (**Figure 4.2.18H**). No significant differences between the bactericidal activity exerted by Pae87 and its analogue protein

lacking the His-tag were found. Therefore, the current evidence does not support a specific role for the His-tag in the observed antibacterial activity of Pae87.

IV. RESULTS

CHAPTER 3 – Application of Pae87 and its derivatives as antimicrobial agents

1. INTRODUCTION

In the previous chapter, an enzybiotic candidate (Pae87) was mined out of a putative phage lysins collection based on its G− surface interaction potential. Evidence was provided on its antimicrobial ability and mode-of-action, concluding that at least part of its effect was due to a nonenzymatic action most probably located at the C-terminal peptide termed P87. Hereafter, we intended to explore the applicability of Pae87 and its derivatives, including P87-based AMPs, as molecules designed for antibacterial purposes. This was accomplished by a double strategy: 1) devising ways to improve Pae87 and P87 activities; and 2) testing such original and derived products in a variety of experimental platforms aimed to provide preliminary results on their validity in practical settings. Approaches for the former aim were increasing HM and net charge of peptide P87 by introducing five mutations, while for trying to improve Pae87, it was fused to an additional cation-binding domain. In this way, improved antimicrobial molecules P88 and Pae87-F were constructed. As for the second aim, such molecules were tested for synergy with standard-of-care antibiotics and against *in vitro* grown *P. aeruginosa* biofilms.

2. MUTATIONAL OPTIMIZATION OF P87: DESIGN OF PEPTIDE P88

A most relevant starting point for this work is the assumption that a higher net charge and HM within a proteinaceous structure implies a better chance for it to interact with the G− surface. On this basis, several modifications of peptide P87 were proposed to increase both its net charge and the HM values to improve its bactericidal activity (**Figure 4.3.1A**). In this way, first-hand evidence on the relevance of HM and net charge for anti-G− activity would also be provided. Similar mutation-based tuning has been previously reported in the literature (373, 374). The five mutations in P87 proposed to design the enhanced antimicrobial peptide P88 theoretically increase the net charge by ≈ 2 units while also increasing amphipathicity (from 0.63 to 0.81). Given JPred secondary structure prediction, such mutations were not expected to significantly hamper the peptide helix propensity (**Figure 4.3.1B**).

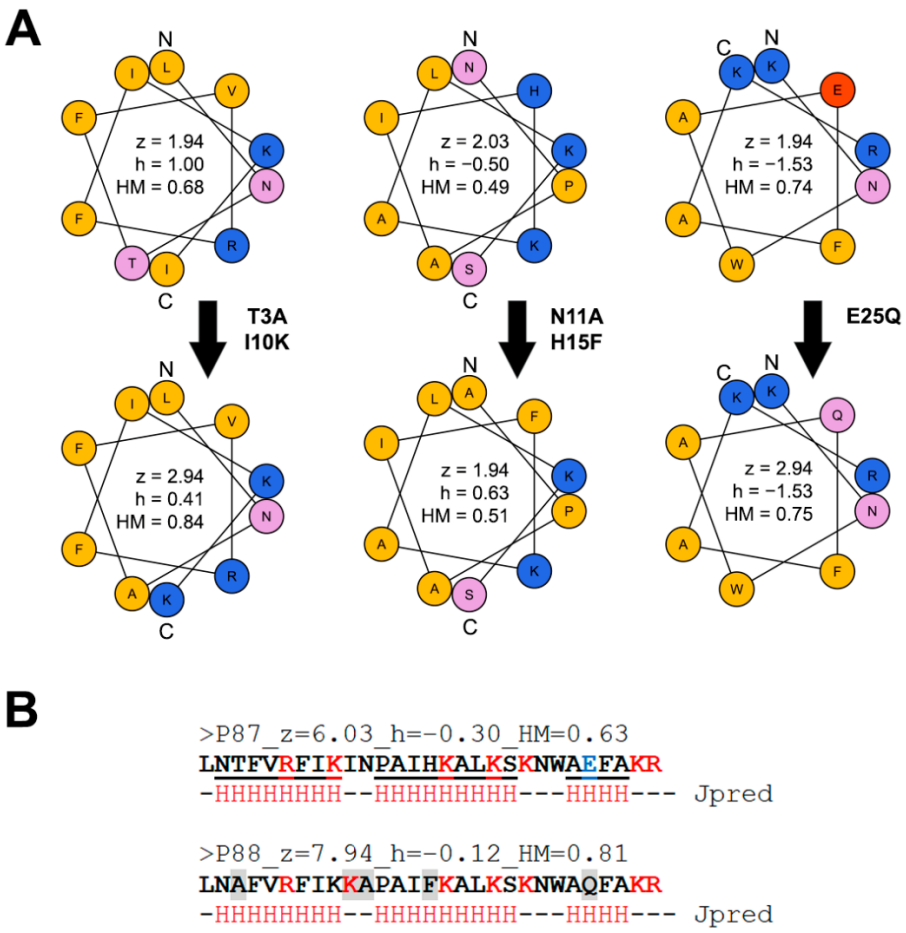


Figure 4.3.1. P88 peptide design. A: Wheel representation of the three possible α-helices of P87 and the derivative P88. The five P88 mutations with respect to P87 are displayed, as well as the net charge (z), hydrophobicity (h), and HM calculated for each of the helices. **B**: P87 and P88 sequences, together with the full sequence net charge, hydrophobicity and HM values, and the JPred secondary structure prediction (a residue marked with an "H" indicates that it is involved in an α-helix). Residues shaded in grey are those mutated in P88.

P88 retained the ability to form α-helices as demonstrated by the circular dichroism far-UV spectra in the presence of different concentrations of TFE (**Figure 4.3.2A**). Assays with membrane permeabilization probes also showed that equimolar concentrations of P88 permeabilized the OM (**Figure 4.3.2B**) and both the outer and the inner membrane (**Figure 4.3.2CD**) more effectively than P87. Altogether, these results prove that P88 has better disruptive interaction with the biologic membranes, and thus the designed mutations accomplished their predicted aim.

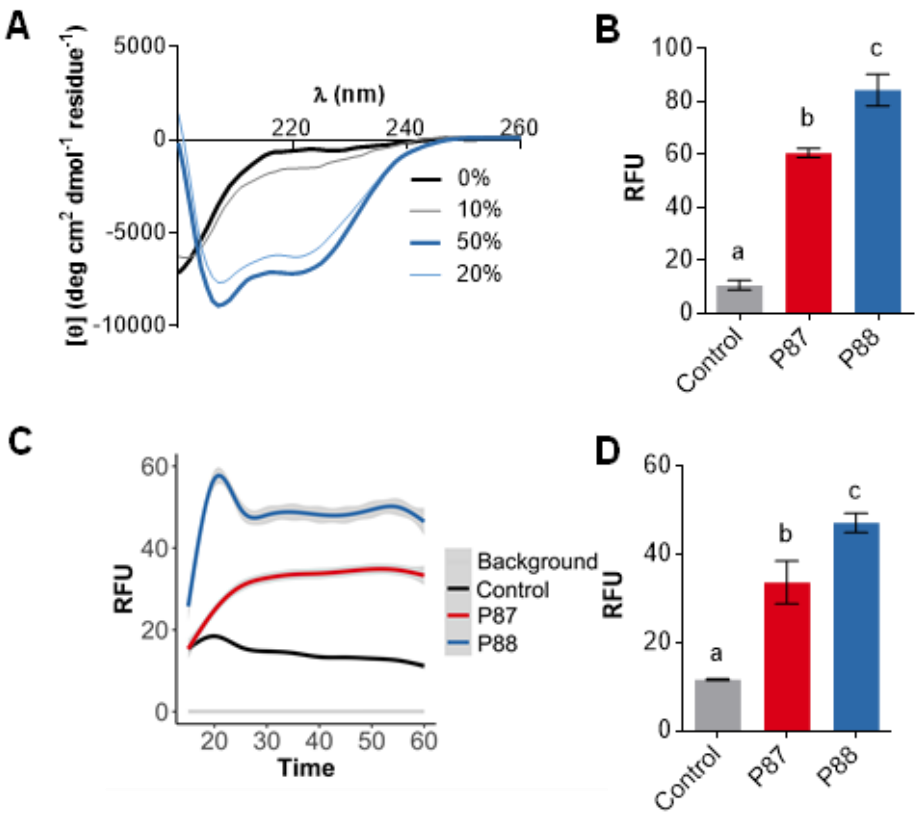


Figure 4.3.2. Ability of P88 to form amphipathic helices and permeabilize biological membranes. A: Far-UV circular dichroism spectra of P88 in the presence of different concentrations of TFE. **B**: Average NPN fluorescence signal after 5 min incubation. **C**: GAM estimation of the average tendencies of the SYTOX fluorescence kinetics. Estimation was based on three independent replicates, mean estimation ± 95% C.I. (grey shade) is shown for each experimental condition. **D**: Average SYTOX fluorescence after 60 min incubation. In **B** and **D**, means ± sd of three independent replicates are shown and a one-way ANOVA test with Tukey post-test was applied for multiple comparisons. Different letters indicate significant differences.

To further check whether the enhanced properties of P88 correlate with greater antibacterial ability, the modified peptide was tested against *P. aeruginosa* PAO1 suspensions in comparison with equimolar amounts of P87 (**Figure 4.3.3**). Like P87, P88 also exerted a rapid lytic effect when added to a PAO1 suspension (**Figure 4.3.3A**). It is worth noting that P88 lytic kinetics seemed faster than that of P87. While for P87 a relevant dependency on the peptide:bacteria ratio was observed, yielding that maximum activity at 10 µM P87 was achieved with a bacterial dose of 10^7 CFU/ml or lower, the activity of 10 µM P88 was not hampered with as much cell density as 10^8 CFU/ml (**Figure 4.3.3B**). This indicates that killing efficiency was increased 10 times with P88 mutations in comparison with P87.

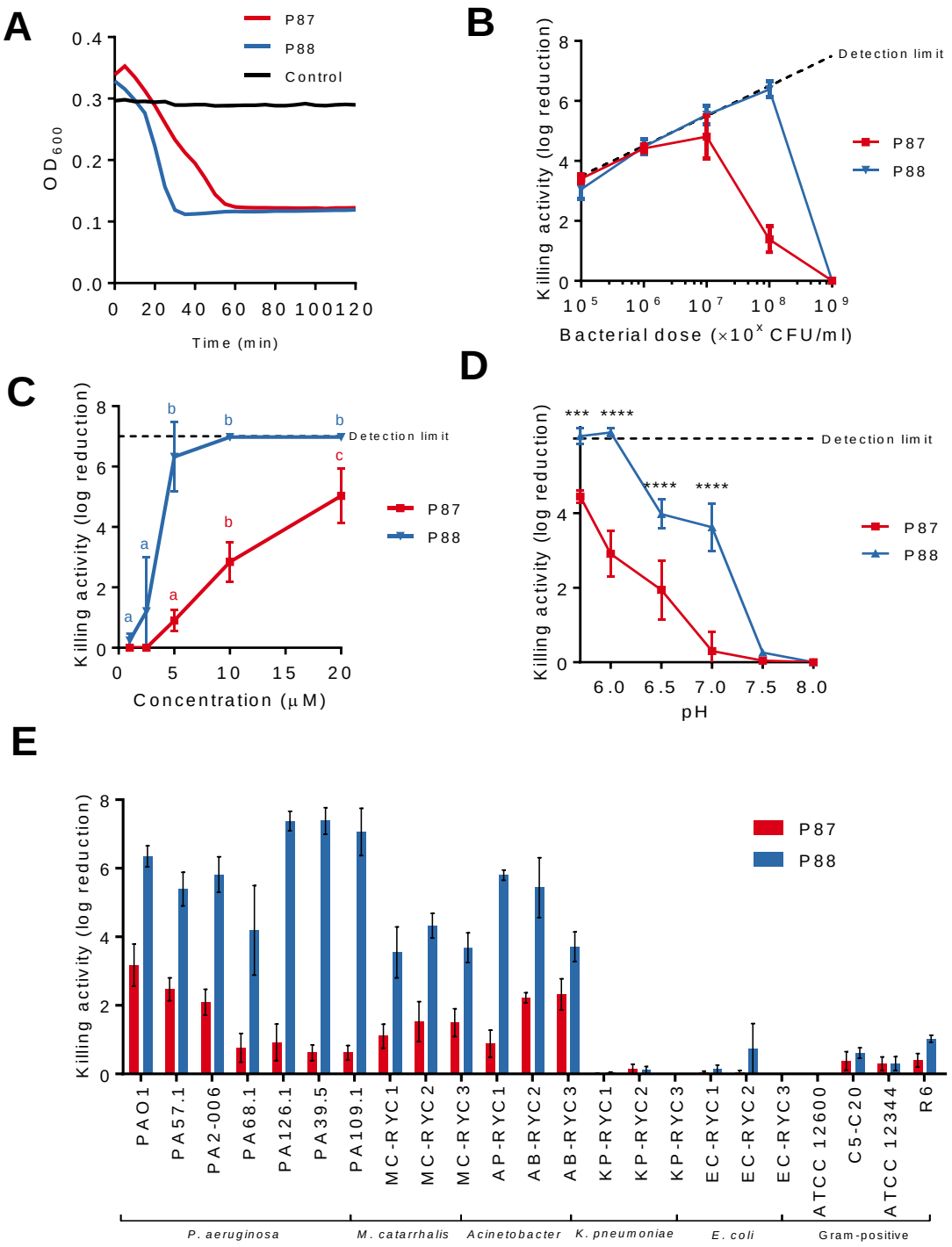


Figure 4.3.3. Parameters of P87 and P88 bactericidal activity on different bacteria. A: Bacteriolytic effect. Representative results are shown. **B**: Peptide:bacteria stoichiometry. **C**: Dose-response curves. One-way ANOVA plus Tukey *post hoc* test was applied to compare concentrations within each curve. Different letters indicate significantly different results. **D**: Variation of killing activity with pH. Two-way ANOVA followed by Dunnett post-test was applied to test significant differences with respect to pH 5.7 (*** $p \leq 0.001$, **** $p \leq 0.0001$) **E**: Bactericidal range. Unless otherwise noted, standard assay conditions were: 20 mM NaPiB, pH 6.0, 150 mM NaCl, ≈10^8 CFU/ml PAO1, 10 μM of the bactericidal compound, 37 °C, and 2 h incubation. P87 data are those generated and displayed in **Chapter 2** of this thesis.

Likewise, the dose-response curve of P87 up to 20 μM against 10^8 CFU/ml PAO1 was linear, while P88 rapidly achieved maximum activity at 5 μM (**Figure 4.3.3C**). As for the variation of activity with pH, P88 replicated the trend observed for P87 in the previous chapter, *i.e.*, its activity is maximized at acidic pHs (**Figure 4.3.3D**). Nonetheless, P88 maintained a remarkable killing activity (around 4 log units) at pHs 6.5−7.0, which could be explained by its higher net charge at the same pHs. The bactericidal spectrum of P88 was the same as P87, pointing to a common mechanism of action, but, again, P88 displayed much higher killing values for all of the susceptible bacteria tested (**Figure 4.3.3E**). Finally, P88 activity was tested against a collection of clinical *P. aeruginosa* strains to check the impact of bacterial genetic variability on its antimicrobial potential (**Figure 4.3.4**). Such experiments were performed by Antonio Doménech-Sánchez at the laboratory of Sebastián Albertí (Universidad de las Islas Baleares, UIB).

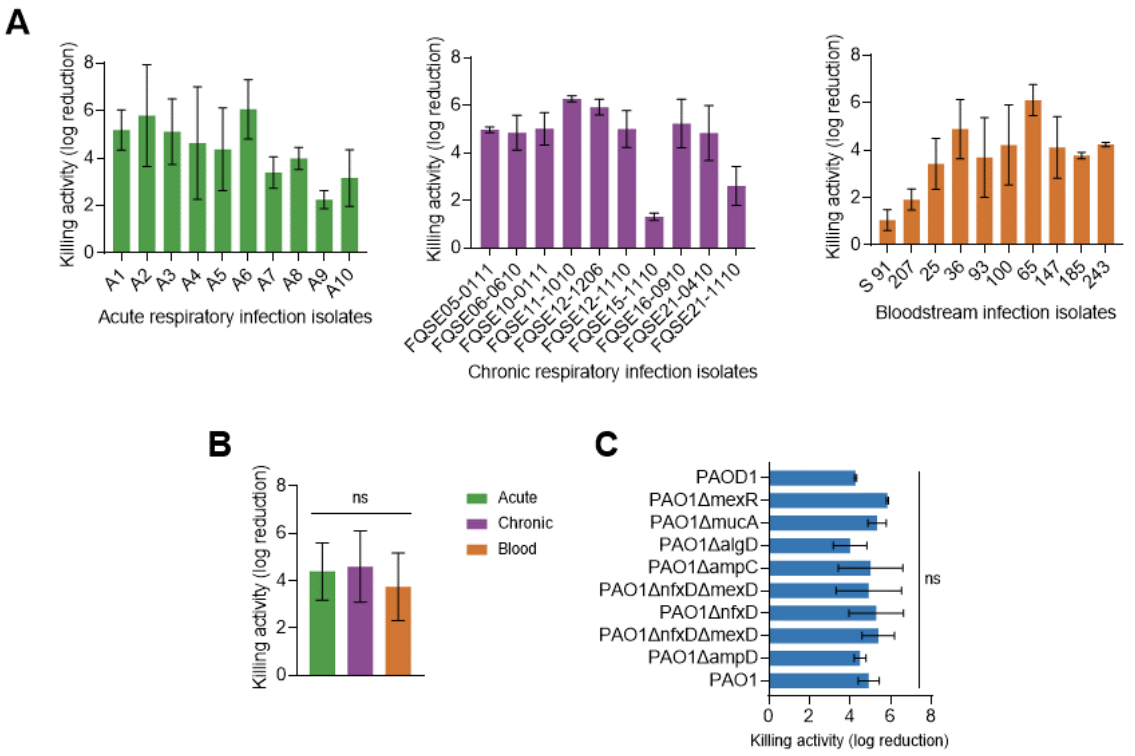


Figure 4.3.4. Impact of *P. aeruginosa* genetic variability on P88 activity. Experiments were carried out by A. Doménech-Sánchez at the laboratory of S. Albertí (UIB). **A**: Bactericidal activity of 10 μM P88 on 10^5 CFU/ml *P. aeruginosa* clinical strains suspensions (in 20 mM NaPiB, pH 6.0, 150 mM NaCl) incubated for 2 h at 37 °C. **B**: Mean activity of P88 against the clinical isolates from different origins (275, 278). One-way ANOVA followed by Tukey post-test was applied to analyse statistical differences. **C**: Activity of 10 μM P88 on 10^5 CFU/ml *P. aeruginosa* PAO1 and a collection of isogenic mutants (276, 277). One-way ANOVA and Dunnett post-test were used to detect statistical differences with respect to the activity against PAO1.

As shown in **Figure 4.3.4A**, all 30 *P. aeruginosa* clinical isolates tested demonstrated to be susceptible to P88 activity. A low variability was observed in the magnitude of the killing effect (ranging between 4-6 logs) except for four more recalcitrant strains. While it is not uncommon to find that isolates from different niches present diverse antibiotic resistance profiles, our results showed no significant differences in susceptibility between the strains from different origins (**Figure 4.3.4B**). In addition, a panel of different isogenic mutants of *P. aeruginosa* strain PAO1 bearing different phenotypes (non-encapsulated, highly encapsulated, etc.) was checked. Again, no specific P88 resistance determinants were found (there were no statistically significant differences in P88 susceptibility among the mutant strains, **Figure 4.3.4C**). All these observations demonstrate that the enhanced ability of P88 to interact with the bacterial membranes with respect to P87 makes it a better antimicrobial.

3. APPLICABILITY OF P88 AS AN ANTIMICROBIAL: CYTOTOXICITY AND SYNERGY

One of the drawbacks usually adduced for the therapeutic application of AMPs is their possible cytotoxic effect against eukaryotic cells. Indeed, our experiments rendered relatively low IC_{50} values as calculated after examining the viability of epithelial A549 lung cell cultures ncubated with different P88 peptide concentrations for 48 h (**Figure 4.3.5**). While the parental enzyme, Pae87, displayed no remarkable cytotoxic effect at the concentrations tested (**Figure 4.3.5A**), both P87 and P88 were somewhat toxic for the A549 cells (**Figure 4.3.5BC**). After fitting the dose-response curves (for details, see **Figure 4.3.5** legend), the calculated IC_{50} was 49 μM for P87 and 34.4 μM for P88. Although a cytotoxic effect may hamper the therapeutic potential of the peptides, the positive note is that, while the results of the previous section show that P88 is ten times more efficient against bacteria than P87, the fold change in toxicity is just 1.42. The known toxic peptide melittin, present in honeybee venom and included as a positive control in the study, had an IC_{50} of 1.9 μg/ml (**Figure 4.3.3D**). This value (≈ 0.7 μM) is much lower than those estimated for the Pae87-derived peptides. On the other hand, the observed P88 MICs for PAO1 and the clinical strain 39.5 were, respectively, 20 μM and 50 μM (**Table 4.3.1**). The ratio between MIC and IC_{50} is much more reasonable in the case of P88 than in that of the toxic melittin, which is described to have a MIC against *P. aeruginosa* strains ranging between 64 and 128 μg/ml (more than 30 times the IC_{50}) (375, 376). Of note, such cytotoxicity should not be an issue whatsoever if the peptides are purposed for such applications as disinfection or food preservation.

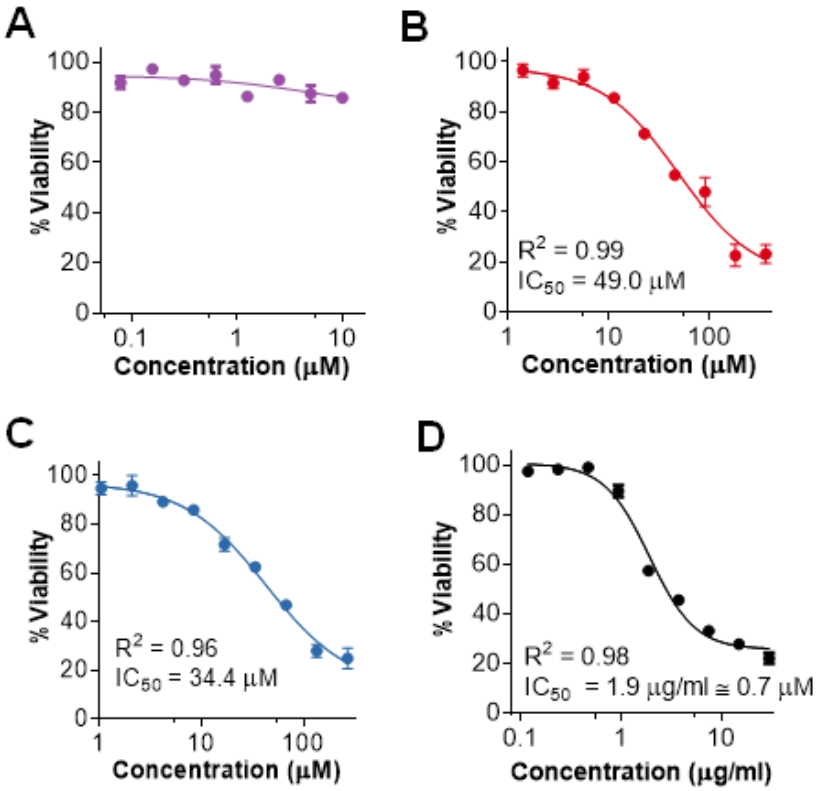


Figure 4.3.5. Cytotoxicity of different Pae87-based peptidic compounds on A549 cell cultures. The residual MTT measured viability with respect to an untreated control after 48 h treatment of A549 cells with different concentrations is shown (**A**: Pae87; **B**: P87; **C**: P88; **D**: melittin as a cell toxicity positive control). Actual data were non-linear fitted to a negative sigmoidal curve (with generalized equation $Viability = 100/(1 + 10^{HillSlope \times (LogIC50 - Concentration)})$) using GraphPad Prism 6. In this way, IC_{50} was calculated when the fitted sigmoid presented clear top and bottom sections (*i.e.*, **B**, **C**, and **D**). R^2 is also presented as a goodness of fit metric.

In the spirit of reducing the effective concentration of P88 and its side cytotoxic effect, experiments on the synergistic activity of the peptide together with several antibiotics were conducted. Both the MICs of P88 and several antibiotics were calculated for the standard strain PAO1 and the clinical isolates 39.5 and 126.1 (**Table 4.3.1**).

Table 4.3.1. MICs of P88 and a range of antibiotics against *P. aeruginosa* strains.

	PAO1	39.5	126.1	
P88	20	50	50	µM
LVX	0.25	0.13	0.25	
ERY	512	512	256	
KAN	16	256	256	µg/ml
CHL	512	512	128	
TET	64	1024	32	
GEN	0.20	>16.00	2.00	
AZI	128	512	256	

Using these MIC values, both a checkerboard assay and an end-point viability determination were performed with different combinations of P88 and the antibiotics, using initially PAO1 and strain 39.5 as models (**Figure 4.3.6**). The checkerboard assay demonstrated synergy for ERY, CHL, and TET in 39.5 strain (**Figure 4.3.6ABC**), but not for LVX, KAN, or GEN (data not shown). The common criterium to prove synergy of a minimum FICI value of 0.5 or below is hereby met since minimum FICI values were 0.46, 0.44, 0.48, and 0.5 respectively for ERY, CHL, TET, and AZI.

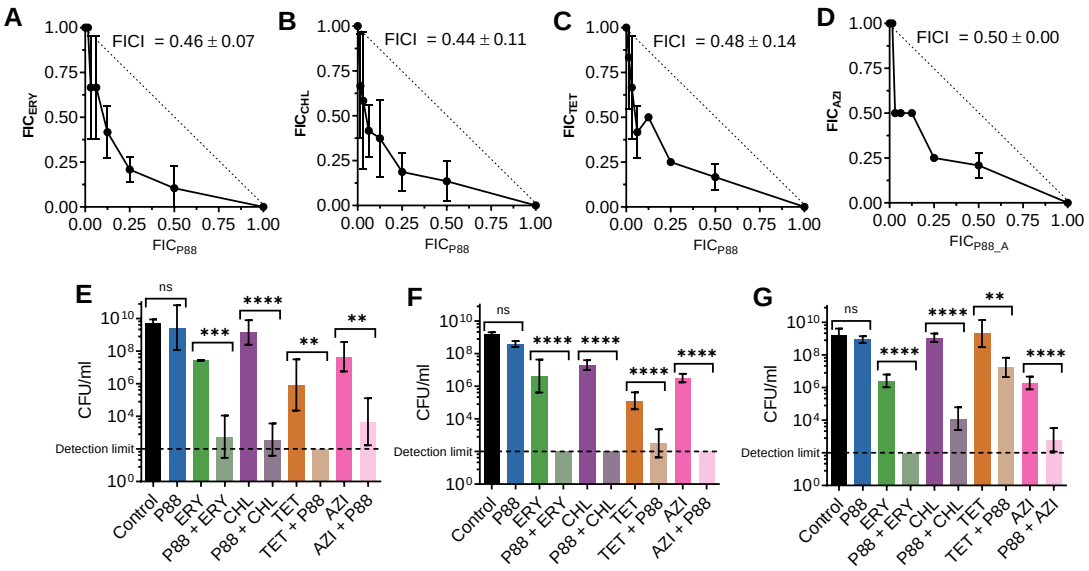


Figure 4.3.6. Synergy of peptide P88 with several antibiotics. A, B, C: Checkerboard assay isobolograms of different P88-antibiotic combinations against *P. aeruginosa* 39.5 (**A**: ERY; **B**: CHL; **C**: TET; **D**: AZI). Mean ± sd of the minimum FICI values are displayed. **D, E:** Viable counts of *P. aeruginosa* 39.5 (**E**), PAO1 (**F**) or (**G**) 126.1 treated overnight (≈16-20 h) at 37 °C in Mueller-Hinton medium (initial inocula ≈ 10^5 CFU/ml) with 0.25 × MIC of an antibiotic, P88 or a combination of 0.25 × MIC of each, plus the untreated control.

Further experiments confirmed synergy both in strain 39.5 and in PAO1 (**Figure 4.3.6DE**). In this case, overnight cultures of each strain in the presence of 0.25 × MIC of P88 did not display a significant decrease in the viable cell counts. A concentration of 0.25 × MIC of each antibiotic caused a 1-3 log units decrease in the case of ERY and CHL and ≈ 4 log units for TET. However, the combination of a mildly active antibiotic concentration with a peptide concentration without effect in viability yielded viable counts at least 2 logs below the ones of the cultures treated only with the antibiotic. In fact, in some cases, the viable counts of the peptide-antibiotic combination were below the detection limit of the assay. Therefore, the viability criterium for synergy (an increase in killing effect ≥ 2 logs higher in the combination than in the sum of the effects of each compound alone) was also met. As for the possible synergy mechanism, it is quite straightforward to say that AMPs (and particularly those which interact with biological membranes) can synergize with antibiotics by enhancing the bacterial cell permeability and, thus, the access of antibiotics to their targets (377). However, this does not provide a clear explanation of the actual reason why P88 demonstrates synergy with some antibiotics with intracellular targets and not others, even more so if we consider the lack of correlation of any of the antibiotic properties displayed in **Table 4.3.2** with the presence or absence of synergy. In reality, the mechanisms of synergy are not well understood and may even rely on AMP antibacterial activities other than cell permeabilization, or their particular mode of action (378).

Table 4.3.2. Properties of the antibiotics used in this work, as taken from Drugbank (379).

Antibiotic	Molecular Weight (g/mol)	LogP[a]	Mechanism	Synergy with P88
LVX	361.4	2.1	Interferes with DNA replication	NO
ERY	733.9	2.6	Interferes with protein synthesis	YES
KAN	484.5	-6.3	Interferes with protein synthesis	NO
CHL	323.1	1.14	Interferes with protein synthesis	YES
TET	444.4	-1.3	Interferes with protein synthesis	YES
GEN	482.5	-4.1	Interferes with protein synthesis	NO

[a] LogP = partition coefficient

The presence of a biological fluid (lysed horse blood), had a slightly detrimental effect on P88 antimicrobial activity, increasing strain 39.5 MIC up to ≈128 μM with 50% lysed horse blood (**Figure 4.3.7A**). The AZI MIC, however, dramatically decreased with the blood concentration (down to 8 μg/ml at 50% lysed horse blood), as already reported in the literature (380). This is explained by an increased cell permeability in complex media, and such a phenomenon is also susceptible to hamper the synergistic cooperation of

antibiotic and peptide. Our results in 25% lysed horse blood indeed showed a slight decrease of the synergistic activity, but synergy can still be observed (**Figure 4.3.7B**).

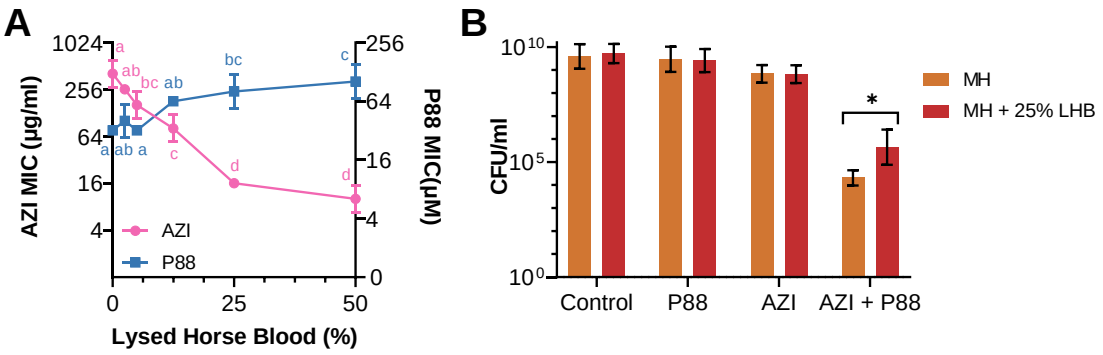


Figure 4.3.7. Impact of blood on the antimicrobial activity of P88 and AZI and their synergy. A: MIC determined in Mueller-Hinton broth supplemented with different concentrations of lysed horse blood. A two-way ANOVA followed by Tukey post-test was applied to find statistically significant differences between concentrations. Different letters indicate significantly different results. **B**: Viable counts of *P. aeruginosa* 39.5 treated overnight (≈16-20 h) at 37 °C in Mueller-Hinton medium (initial inocula ≈ 10^5 CFU/ml) with 0.25 × MIC of an antibiotic, P88 or a combination of 0.25 × MIC of each, plus the untreated control, with or without 25% lysed horse blood.

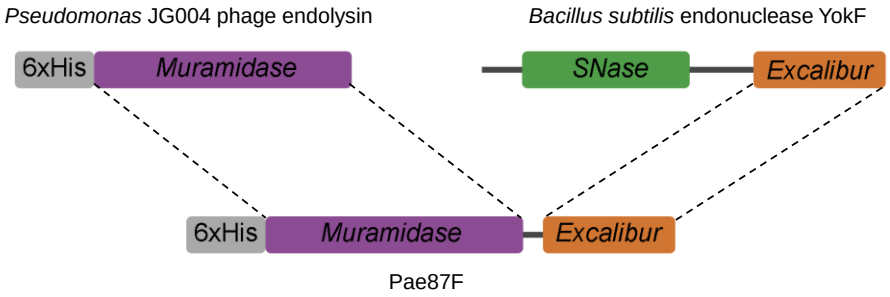
In conclusion, synergy provides a possible strategy to therapeutically use P88 with concentrations well below the IC_{50} (the effective concentrations in the presence of sub-inhibitory amounts of antibiotics are 5.0 µM and 12.5 µM for PAO1 and 39.5/126.1, respectively) and allows to postulate such P88 as a molecule able to resensitize bacteria to antibiotics. Nonetheless further *in vivo* proofs shall be needed to ascertain the therapeutic value of P88 and its combined administration with antibiotics.

4. IMPROVING PAE87 ACTIVITY BY FUSING IT TO A CATION-BINDING DOMAIN: PAE87F

Although the full antimicrobial potential of Pae87 is difficult to assess without proper *in vivo* studies, it is clear that its bactericidal effect numbers are mid-to-low, especially in comparison with other engineered anti-G− lysins (such as artilysins). To increase the bactericidal activity of Pae87, we proposed the yet undescribed strategy of fusing the enzybiotic to a cation binding domain. The rationale for this approach lays in the fact that the G− OM is electrostatically stabilized by positive cations (Mg^{2+} and Ca^{2+}) that bind to the negatively charged lipid A moiety of the lipopolysaccharide (381). This is why cation chelating agents, such as EDTA or organic acids have a disturbing effect on the OM. With this in mind, we conducted a literature search for bacterial cation-binding domains. An *Excalibur* (PF) domain from *Bacillus subtilis* endonuclease YokF was selected to be

fused to Pae87, based on previous evidence of Ca-binding abilities for this domain family and a positive net charge (even if the proposed residues responsible for cations binding are acidic ones).

A



B

>Excalibur_YokF_z=+2
TFKNCTELRKKYPNGVPSSHPAYQSKMDRDHDNYACER

>Randomized_F
DRVSKYPKGPFDYAKNQESNCHMSKNATTECRRLPDHY

Figure 4.3.8. Design of Pae87F. A: Chimeric lysin Pae87F design comprising an EAD from Pae87 and cation-binding domain from exonuclease YokF. **B**: Sequence of *Excalibur* domain from YokF and a "randomized" version of the same domain, containing the same residues in random positions. Predicted net charge (z) at pH 7.0 is displayed, and basic (red) and acidic (blue) residues are indicated.

With this approach, we intended to enhance Pae87 tropism towards the bacterial OM and even increase the OM disrupting abilities of the new protein, named Pae87F. The resulting protein contained an N-terminal 6×His tag to facilitate purification, contributed by the expression vector pET-28a(+), the full *Muramidase* domain of Pae87, a $(GS)_3$ flexible short linker, and the 38 aa long *Excalibur* domain from YokF (**Figure 4.3.8**). Pae87F was expressed and purified as described in MATERIALS AND METHODS. A Pae87F variant named Pae87RF (**Figure 4.3.8B**), which contains a randomized *Excalibur* sequence that maintains net charge but is designed not to keep the cation binding ability, was also constructed, produced, and purified. Next, a set of antibacterial activity tests were run against suspensions of *P. aeruginosa* PAO1 and other bacteria (**Figure 4.3.9**).

Firstly, the turbidity decrease profile of equimolar amounts of Pae87F did not differ substantially from that of Pae87 (**Figure 4.3.9A**), from which no dramatic changes in the mode-of-interaction are deduced for the chimeric protein.

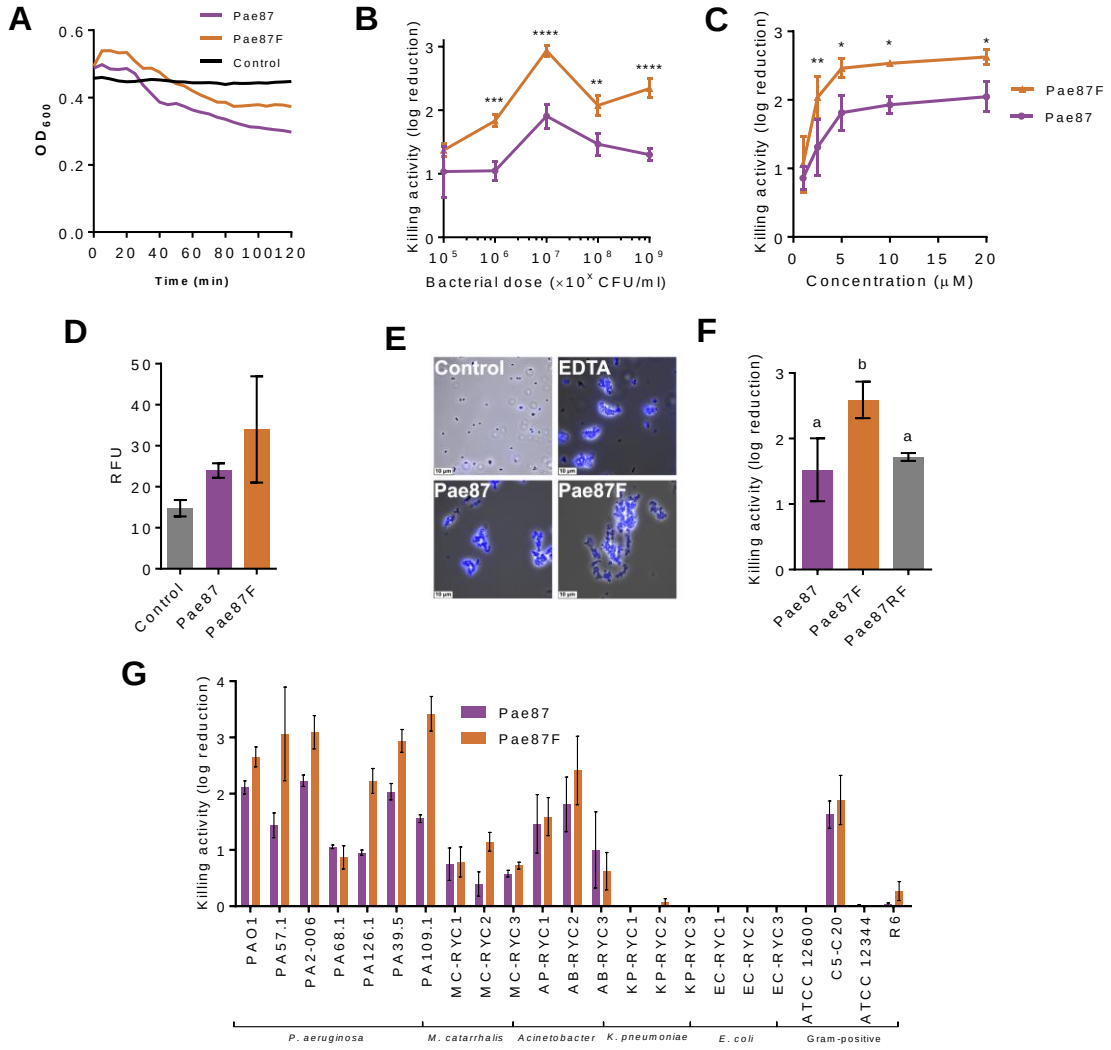


Figure 4.3.9. Comparative parameters of Pae87 and Pae87F bactericidal activity on *P. aeruginosa* PAO1 and other bacteria. A: Bacteriolytic activity curves on PAO1 strain. Representative results are shown. **B:** Protein:bacteria stoichiometry. **C**: Dose-response curves. **D**: Average NPN fluorescence signal after 5 min incubation. RFU = fluorescence units relative to the maximum value achieved during the assay. **E:** Fluorescence microscopy observation of fluorescence due to NPN after 2 h incubation, including a 2 mM EDTA positive control. Superimpositions of phase-contrast images with blue fluorescence signal observed with an A filter cube (excitation bandpass 340-380) are shown. **F**: bactericidal activity of Pae87, Pae87F, and an equimolar concentration of Pae87RF. **G:** Bactericidal range**.** Unless otherwise noted, standard assay conditions were: 20 mM NaPiB, pH 6.0, 150 mM NaCl, ≈10^8 CFU/ml PAO1, 10 μM of the bactericidal compound, 37 °C and 2 h incubation; 100 mM sorbitol was included in the NPN assay. One-way ANOVA with Bonferroni post-test was applied to compare between Pae87 and Pae87F results within each experimental condition when indicated with asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$). When marked with letters, an all-against-all multiple comparison was applied using ANOVA plus Tukey post-test. Different letters indicate significantly different results. Pae87 data are those generated and displayed in **Chapter 2** of this thesis.

However, **Figure 4.3.9BC** showed a consistent 1 log unit increase in bacterial killing in the case of Pae87F compared to Pae87, which is maintained across the different protein:bacteria ratios and protein concentrations tested. The quantification of fluorescence due to NPN incorporating into the phospholipidic membrane upon addition of Pae87 or

Pae87F (**Figure 4.3.9D**) was not significantly different, although comparisons, in this case, should be made with care because of the aggregative effect that is also present in Pae87F (**Figure 4.3.9E**). The relatively lower activity displayed by Pae87RF indicates that the observed increase in bactericidal activity should be due to such specific activity. Such bactericidal enhancement was observed in many but not all of the susceptible bacteria tested in the bactericidal range (**Figure 4.3.9G**). Although the present results do suggest that the improvement must be due to the cation binding, it is not possible yet to claim a specific mechanism for such enhancement. It is considered, however, that the fusion to a cation-binding domain could be a valid strategy to improve enzybiotic activity. The improved effect of Pae87F over Pae87 has been hereby provided as the first example of such a strategy. However, the optimization of Pae87 activity by enhancing the interaction with the bacterial surface has not yielded too notable results, probably because Pae87 already showed signs of a remarkable ability to interact with the cell wall.

5. TESTING OF PAE87 AND ITS DERIVATIVES IN BIOFILMS

The proteinaceous antimicrobials developed in this thesis were also tested in an *in vitro* biofilm model, given the currently acknowledged relevance of biofilms in clinical infectious settings. To that end, our available *P. aeruginosa* strains were tested for biofilm formation in polystyrene plates. A fixed inoculum of each strain (≈ 10^6 CFU/ml) was added into each well in TSB medium and biofilms were formed overnight (≈ 16-20 h) at 37 °C without shaking. **Figure 4.3.10A** shows a rather high variability among the CV-quantified biofilms in the different strains, which is not uncommon. Two clinical strains (126.1 and 57.1) were selected to proceed on the basis of their good-biofilm forming capacity and susceptibility to the compounds based on Pae87. In both cases, a rather mild disaggregation effect (as evidenced by the CV measurements) correlated with some log units of killing (**Figure 4.3.10BC**). Not too surprisingly, P88 exerted the highest killing activity, reaching up to three log units of killing, however, somewhat below its previously shown activity against planktonic cultures. Given the structure of biofilms, it is to some point expected a hampering to the effectiveness of the different antimicrobials. However, the fact that a relevant bactericidal effect is detected against biofilms of two different clinical *P. aeruginosa* strains is a good prospect towards their application as therapeutics or disinfectants.

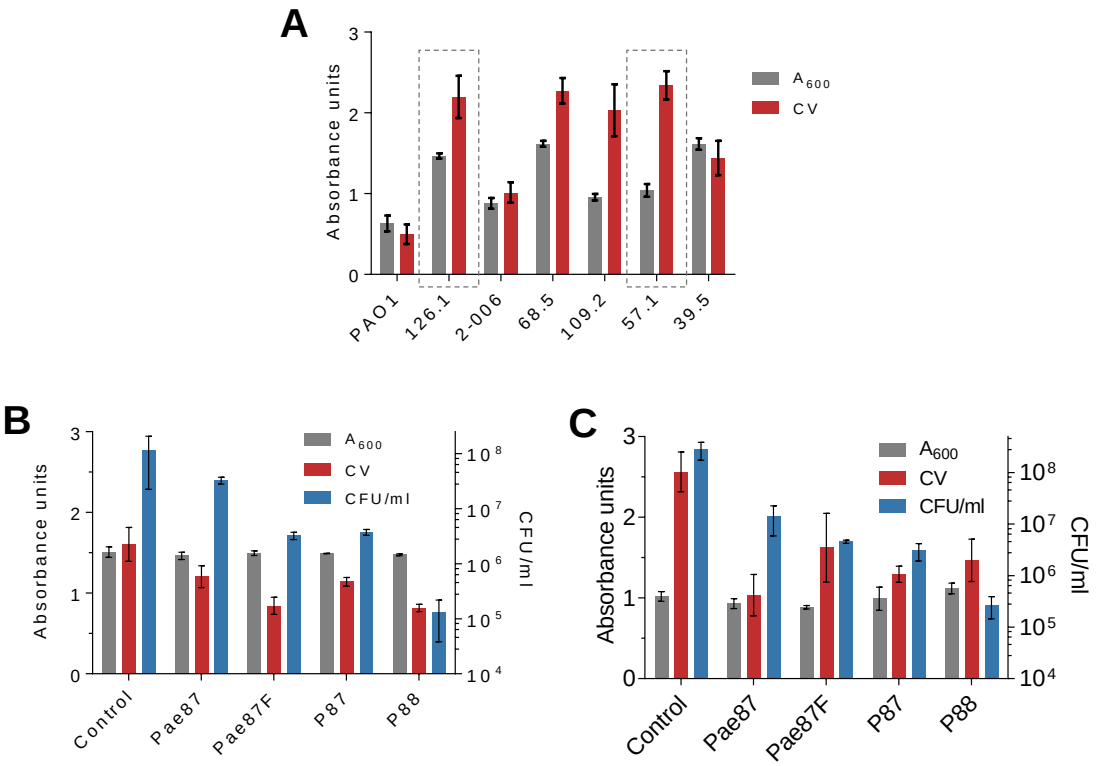


Figure 4.3.10. Activity test of Pae87 and its derivatives against *P. aeruginosa* biofilms. A: Evaluation of the biofilm formation capacity of our available clinical *P. aeruginosa* strains. **B, C**: Biofilm disaggregation assays of *P. aeruginosa* with 10 μM of different compounds applied for 2 h at 37 °C.

IV. RESULTS

CHAPTER 4 – Development of a pneumococcus-biomimetic material for encapsulation and controlled release of choline-binding proteins

The results in this chapter are intended to be published in the following manuscript in preparation:

Roberto Vázquez*, Francisco J. Caro-León*, Alberto Nakal, Susana Ruiz, Carmen Doñoro, Luis García, Blanca Vázquez-Lasa, Julio San Román, Jesús Sanz, Pedro García*, María Rosa Aguilar*. (2021). DEAE-chitosan nanoparticles as a pneumococcus-biomimetic material for the development of antipneumococcal therapeutics.

1. INTRODUCTION

The introduction of enzybiotics as novel therapeutic agents may face some challenges because of their nature. Common antimicrobial chemotherapy has relied on small molecules (antibiotics), which, in general, are less susceptible to provoke immune reactions or to be degraded by specific host mechanisms. Enzybiotics, as proteins, are suspected to induce an immune reaction in which neutralizing antibiotics may be produced (although it has been shown that neutralization is not absolute), and they are also susceptible to host proteases. Because of these particularities, the current interest of enzybiotics research has derived towards the formulation of lysin-based preparations in which certain properties relevant for clinical applications are enhanced. This novel generation of enzybiotics has been labelled as a 'Third Generation', following the first one of wild-type phage lysins and the second one of lysins engineered to enhance biochemical or antimicrobial features. Among the many approaches to develop third-generation lysins, one of them is the immobilization of the enzybiotics in polymeric scaffolds, either in the form of layers of wound dressing bandages or as NPs for systemic or local administration. The advantages of such immobilization range from the possibility to administrate the active compound into niches difficult to reach (such as the deep lung), the administration in specific areas, such as the skin, by formulating emulsions or wound dressings, or the sustained release of the active compound. This latter property of enzyme-immobilized systems may pose a great gain in terms of active compound stability in *in vivo* settings. Bearing this in mind, in this chapter, the design and testing of a novel chitosan-based material with the ability to bind CBPs is presented. Such chitosan, derivatized with DEAE radicals (ChiDE) was used to prepare NPs into which a previously developed enzybiotic (Cpl-711) was encapsulated for controlled release.

2. RATIONALE FOR THE DEVELOPMENT OF A PNEUMOCOCCUS-BIOMIMETIC CHITOSAN

As explained in the INTRODUCTION (**Section 6**), *S. pneumoniae*, the main bacterial pathogen causing CAP in the world, contains choline residues grafted onto the teichoic acids of its G+ cell wall (**Figure 4.4.1A**). Such choline residues are of the utmost importance for the physiology of the bacterium since they serve as ligands for several functionally relevant surface proteins (CBPs), such as murein hydrolases responsible for peptidoglycan remodeling, autolysis, or fratricide (151). Besides, most phages that infect pneumococcus have evolutionarily imitated such host enzymes by acquiring choline-binding murein hydrolases as lysins for the release of their progeny. Moreover, CBPs also recognize molecules chemically analogous to choline, such as the DEAE radicals

(**Figure 4.4.1B**) that are widely used, for example, in ion-exchange chromatography. This is why a chitosan-based material was designed in which some of the deacetylated C-2 sites of the polymer chains were derivatized with DEAE radicals, aiming to imitate the multivalent disposition of choline in the pneumococcal surface (**Figure 4.4.1C**).

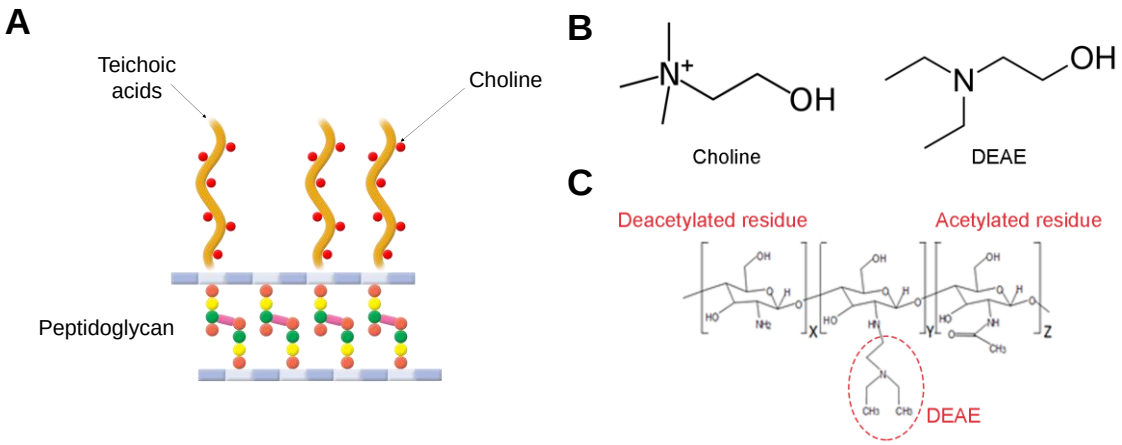


Figure 4.4.1. Rationale for the consideration of ChiDE as a mimic of the pneumococcal cell wall. A: Scheme of the pneumococcal cell wall, remarking the choline residues present in its teichoic acids. **B**: Chemical structures of choline and DEAE. **C:** Chemical representation of ChiDE.

Such derivatization was carried out by Javier Caro-León at the laboratory of María Rosa Aguilar (ICTP-CSIC) by reacting chitosan with DEAE-Cl at a slightly basic pH. The grafting of DEAE onto the polysaccharide was confirmed by ^{1}H NMR, rendering a 46 mol% degree of substitution (J. Caro-León, personal communication). Using this ChiDE, NPs (ChiDENPs) were obtained by ionotropic gelation, based on the interaction between the cationic amino groups of the ChiDE and the anionic phosphate groups of tripolyphosphate molecules (382). The mean hydrodynamic diameter of the ChiDENPs, as recorded by DLS, was 168 ± 5.4 nm, with a PDI of 0.137 ± 0.010 (J. Caro-León, personal communication). Using the same procedure, NPs of unmodified chitosan (ChiNPs) were also obtained for comparison, displaying a similar size (152 ± 4.8) but greater PDI (0.217 ± 0.013).

3. ChiDENPs prevent pneumococcal autolysis and promote chain formation

Given that the aim was to investigate the biomimetic properties of ChiDENPs with respect to the *S. pneumoniae* cell wall, firstly, the interactions between the NPs with pneumococcal cells were analyzed (**Figure 4.4.2**). *S. pneumoniae* R6 cultures grown in

the presence of up to 50 µg/ml ChiNPs showed more pronounced autolysis than the control cells and the ones treated with ChiDENPs (**Figure 4.4.2AB**). A clearer effect, however, was observed when the treatment was applied to resting cell suspensions (**Figure 4.4.2CD**). In this latter case, turbidity decrease due to inherent pneumococcal autolysins during 1 h incubation time was diminished in a dose-dependent manner for ChiDENPs, while ChiNPs induced a more marked turbidity decrease, rather than lysis protection. This timewise decrease of OD_{550} with ChiNPs was accompanied by biomass sedimentation at the bottom of the test tube and correlated with lower counts of viable cells (**Figure 4.4.2E**). On the other hand, no changes were found in cell counts of ChiDENPs-treated resting pneumococci with respect to the control (**Figure 4.4.2F**). Taken these observations together, the conclusion is that ChiDENPs were able to interfere with the normal CBP physiology.

Under this hypothesis, ChiNPs directly interact with the bacteria, as hinted by the aggregation and aberrant morphologies —loss of spheroidal shape, presence of cell debris— observed in the micrographs of **Figure 4.4.2IK**. But ChiDENPs would rather sequester CBPs, protecting cells from CBPs-mediated autolysis and promoting their chaining (**Figure 4.4.2JL**), in contrast with the typical diplococcal morphology of the control cells (**Figure 4.4.2G**). The hijacking of CBPs has long been associated with a chained phenotype (383). The simplest way of getting such phenotype is, indeed, adding an excess amount of choline. Assuming a 46 mol% degree of substitution, 50 µg/ml ChiDENPs would approximately contain 126 µM DEAE residues. As shown in **Figure 4.4.2H**, 200 µM free choline is not able to exert a comparable chaining effect. In fact, the lowest concentration threshold for choline to provoke a relevant pneumococcal chained phenotype is in the millimolar order (234) (specifically, some experiments point out to 25 mM, J. Sanz, personal communication). Therefore, an exponentially enhanced CBP-binding effect of ≈ 200-fold was observed due to the multivalent disposition on the ChiDENPs, which is better recognized by CBPs, as previously observed for other multivalent choline-bearing NPs (232). Such increased efficiency has been proposed to be beneficial for the application in clinical settings, in which the chained phenotype can improve the ability of the immune system to detect and clear pneumococci (235, 357).

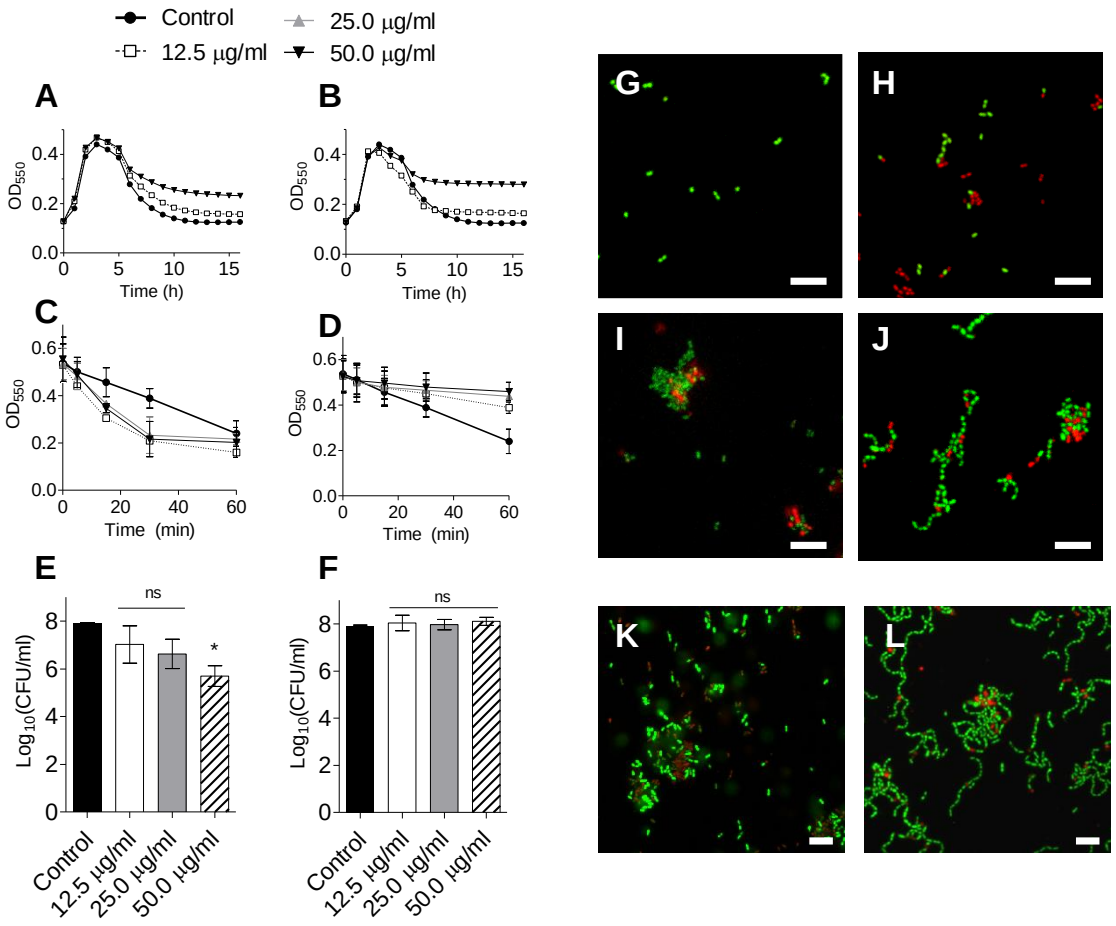


Figure 4.4.2. Effect of ChiNPs and ChiDENPs on *S. pneumoniae* R6 cultures. A, B: Pneumococcal growth in C+ Y medium in the presence of different concentrations of ChiNPs (**A**) or ChiDENPs (**B**). **C, D**: Turbidity change over time of pneumococcal resting cell suspensions in the presence of different concentrations of ChiNPs (**C**) or ChiDENPs (**D**). Values in panels **A** and **B** are representative of three independent replicates. Results in panels **C** and **D** are means ± sd of three independent replicates. **E, F**: Viable cell counts of pneumococcal resting cell suspensions after 1 h incubation with ChiNPs (**E**) or ChiDENPs (**F**). Mean values ± sd of three independent replicates are shown and a one-tailed ANOVA followed by a Tukey post-test was performed to assess significant differences with respect to the control. The asterisk means *p*-value ≤ 0.05. **G, H, I, J**: Projections from CLSM imaging of pneumococcal cultures at early stationary growth phase (≈ 4 h incubation) either untreated (**G**), or grown in the presence of 200 µM free choline (**H**), 50 µg/ml ChiNPs (**I**) or 50 µg/ml ChiDENPs (**J**). **K, L**: Fluorescence microscopy images of the late stationary growth phase of pneumococcal cultures (≈6 h incubation) grown either in the presence of 50 µg/ml ChiNPs (**K**) or 50 µg/ml ChiDENPs (**L**). White bars in microscopy images represent 5 µm, green fluorescent cells are alive cells, while red fluorescent cells are membrane-compromised or dead cells.

4. ChiDENPs specifically bind proteins through DEAE affinity mediated by CBDs

To explicitly test whether ChiDENPs were able to bind CBPs through affinity interaction, an enzybiotic CBP (namely, the chimeric lysin Cpl-711, **Figure 4.4.3A**) was incubated in NPs suspensions, as explained in MATERIALS AND METHODS. Such suspensions were subsequently filtered through 0.1 µm pore size filters and both the whole and the filtered fractions were tested for bactericidal activity on *S. pneumoniae* R6 resting cell suspensions at a final enzyme maximum concentration of 2 µg/ml and maximum ChiNPs

or ChiDENPs concentration of 50 µg/ml (**Figure 4.4.3B**). Free Cpl-711 that had been incubated in the absence of NPs caused a 2-log cell viability decrease with respect to the water-treated control. ChiNPs decreased viability with respect to the control by around 1 log, but only in the whole fraction and not in the filtrate, indicating that most of the NPs are indeed being retained in the filter. On the other hand, ChiNPs incubated together with Cpl-711 exerted a viability decrease more similar to that of free Cpl-711, but only in the whole fraction, with a significant difference with the viable cells treated with the filtrate. This indicates that the enzybiotic was in the whole fraction (*i.e.*, bound to the ChiNPs), but it still was able to release upon contact with the bacteria and exert its bactericidal activity. This points out to non-specific adsorption of the enzyme that is easily overcome by the presence of the natural affinity ligand of the CBP (*i.e.*, the bacterial cell wall). On the other hand, ChiDENPs, as seen before, had no impact on cell viability with respect to the control, either alone or those incubated with Cpl-711. The fact that there was no activity in the filtrate of ChiDENPs incubated with Cpl-711 suggests that the enzyme was indeed associated or bound to the ChiDENPs. But, since there was also no activity in the whole fraction, in contrast with the ChiNPs, this association must be different in nature, *i.e.*, stronger and most probably due to an affinity interaction. This also clearly supports the ability of ChiDENPs to neutralize the physiological activity of natural CBPs in the presence of *S. pneumoniae*.

When observing Alexa488-labeled Cpl-711 interaction with the different NPs under the microscope (**Figure 4.4.3CDEF**), large fluorescent aggregates were visible in the case of ChiNPs (**Figure 4.4.3C**), while with ChiDENPs, fluorescence was located in discrete, nanometric particles (**Figure 4.4.3E**). Immediately after treatment with 2% choline, ChiNPs incubated with Alexa488-Cpl-711 showed little morphological difference with respect to their appearance before treatment (**Figure 4.4.3D**), while ChiDENPs lost the previously observed locally-focused fluorescence (**Figure 4.4.3F**), suggesting the elution of the protein from the ChiDE.

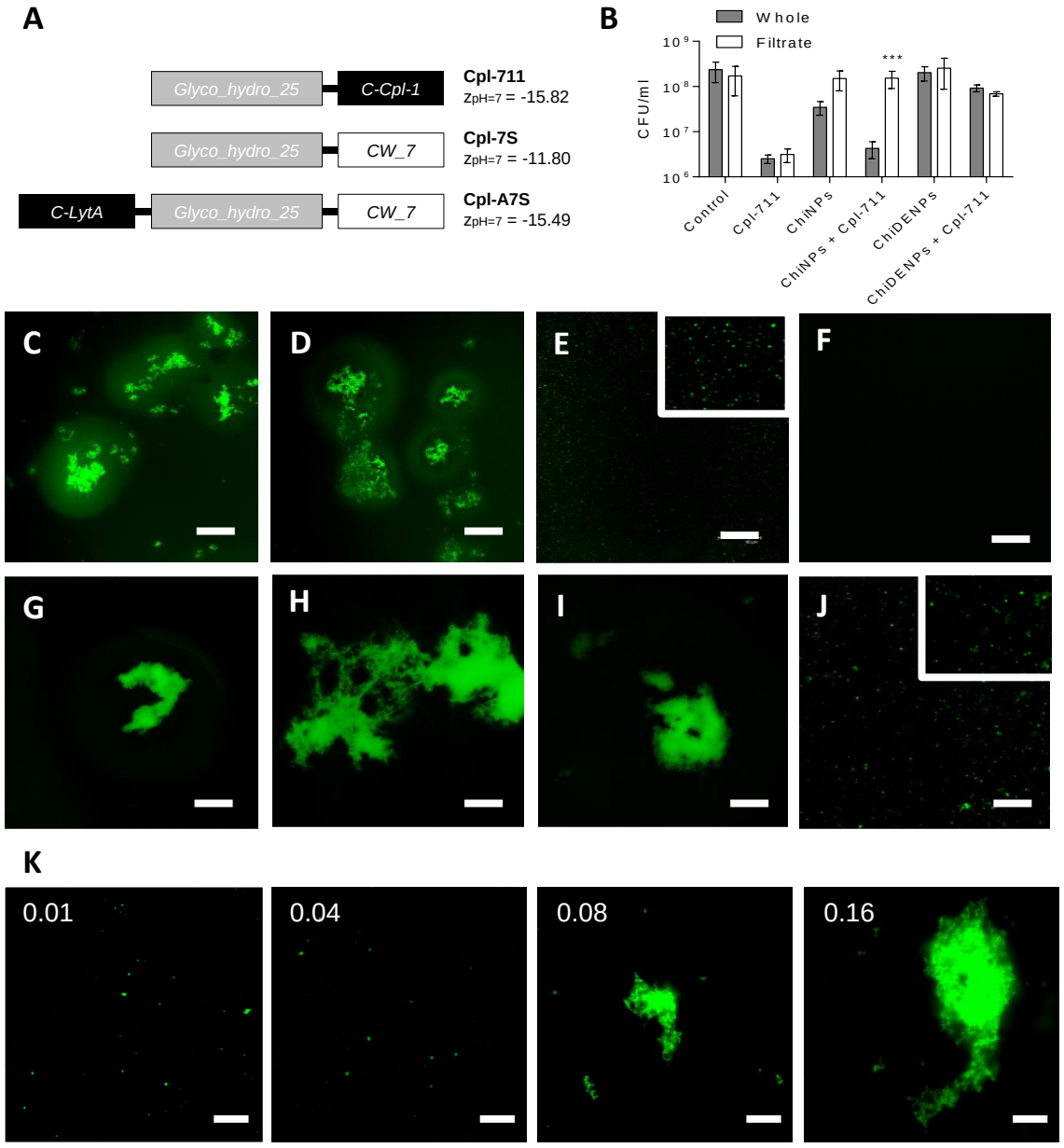


Figure 4.4.3. Surface interaction of CBPs and non-CBPs with ChiNPs and ChiDENPs. A: Structural schemes of the proteins used throughout this work. Black boxes are CBDs, grey boxes are catalytic domains and white boxes are non-CBD CWBDs. **B**: Cpl-711 surface adsorption assay where cell viable counts are shown from *S. pneumoniae* R6 suspensions treated with water (control), 2 µg/ml Cpl-711 or 50 µg/ml ChiNPs/ChiDENPs pre-incubated or not with Cpl-711 at a 0.04 protein/polymer mass ratio. Viable cells were assayed both from the whole treated pneumococcal suspension fractions or from the filtrated fractions through a 0.1 µm pore diameter filter. Mean values ± sd are shown for three independent replicates, and statistical significance is indicated (***: p-value ≤ 0.001) according to two-way ANOVA followed by Tukey post-test to compare whole fractions *vs.* filtrated ones. **C,D,E,F**: CLSM images from ChiNPs (**C,D**) or ChiDENPs (**E,F**) preincubated for 30 min at room temperature with Alexa488-labeled Cpl-711 at a mass ratio of 0.04 and then treated with 2% choline (**D,F**). **G,H,I,J:** CLSM images from ChiNPs (**G,I**) or ChiDENPs (**H,J**) preincubated for 30 min at room temperature with Alexa488-labeled Cpl-7S (**G,H**) or Cpl-A7S (**I,J**). **K**: Fluorescence microscopy images of different mass ratios of Alexa488 labeled Cpl-711/ChiDENPs. Inserts in E and J are 2x zoomed images. White bars in all images represent 40 µm.

To confirm that the nature of the specific interaction of Cpl-711 with ChiDENPs was due to the DEAE binding by Cpl-711 CBD, the same experiment was carried out with the

non-CBP enzybiotic Cpl-7S. Cpl-7S has a similar structure to that of Cpl-711, exclusively differing in their C-terminal CWBDs (**Figure 4.4.3A**). Similar to what was observed for Cpl-711, ChiNPs incubated together with Alexa488-Cpl-7S produced large fluorescent aggregates (**Figure 4.4.3G**). Both enzymes also have a relatively high negative net charge, which would explain a nonspecific, electrostatic interaction with ChiNPs, given its positive ζ-potential as has been previously observed with other proteins in water (384-386). But, furthermore, aggregates were also produced when incubating Cpl-7S with ChiDENPs (**Figure 4.4.3H**). A fusion protein based on Cpl-7S but bearing an N-terminal tag consisting of a CBD derived from pneumococcal autolysin LytA (C-LytA) was also designed to determine if the presence of a CBD would impact the mode of interaction of Cpl-7S with the NPs. This fusion protein, termed Cpl-A7S, was indeed very similar to Cpl-711 in terms of net charge but was substantially larger (56.87 kDa versus 38.58 kDa). Alexa488-Cpl-A7S showed an identical propensity to aggregate in the presence of ChiNPs than its parental enzyme, Cpl-7S (**Figure 4.4.3I**), but not when put together with ChiDENPs (**Figure 4.4.3J**). In this latter case, as with ChiDENPs in the presence of Cpl-711, a dispersion of fluorescent particles was observed rather than aggregates. Taking all these results together, it is clear that the specific interaction between the DEAE moieties of ChiDENPs and the CBD of a protein maintains the stability of the nanoparticulate suspension better than the non-specific interaction that keeps together the proteins with the NPs.

Finally, the ability to retain a stable nanoparticulate state in the presence of different concentrations of CBPs was investigated by observing ChiDENPs incubated with different concentrations of Alexa488-Cpl-711. Up to a polymer/protein mass ratio of 0.04, the suspensions appeared nanoparticulate, but above such ratio, aggregates started to be apparent (**Figure 4.4.3K**). It could be argued that, while the DEAE residues provide a net positive charge to the ChiDENPs that keep them apart by electrostatic repulsion, binding of the CBPs causes hiding of these groups from the solvent, resulting instead in non-polar particle-particle interactions that produce the same aggregative effect as observed with the non-functionalized ChiNPs.

5. Encapsulation of Cpl-711 within ChiDENPs and its sustained release

In a step forward to achieve therapeutically relevant in vivo concentrations and release of therapeutic proteins, the ChiDENPs were loaded with the enzybiotic Cpl-711, thus, exploring the possibility of this system (ChiDENPs-711) to release a substantial amount of the protein in a controlled manner. In this case, such protein was included together with the chitosan or ChiDE solution during the ionic gelation process in the

fabrication of NPs. In this way, we expected to get a nanoparticulate system with a higher active compound concentration, provided that Cpl-711 could be included within the NPs, and bound to inner DEAE residues that otherwise would be non-accessible. A range of protein/polymer ratios well above the aggregation threshold described in the previous section (0.1-0.3) was studied to determine the optimum ratio for the production of ChiDENPs-711. At a ratio of 0.25, an encapsulation efficiency[9] of ≈70% and a loading capacity[10] of ≈ 20% were achieved while maintaining reproducible and good values of hydrodynamic diameter (J. Caro-León, personal communication). Therefore, successive experiments were conducted with ChiDENPs-711 produced at a protein/polymer mass ratio of 0.25. The incorporation of Cpl-711 into the NPs was confirmed by a range of techniques, including thermogravimetric analysis, attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR), or X-ray diffraction by J. Caro-León at the ICTP-CSIC.

The next step was to check the ability of such a system to release its cargo and allow its subsequent activity against *S. pneumoniae*. After incubation of ChiDENPs-711 at 37 ºC and pH 7.4, an initial burst release of the enzybiotic (2-3 first hours, accounting for almost 60% of release) was observed (**Figure 4.4.4A**), followed by a constant release. After 12 h, the cumulative release of Cpl-711 from ChiDENPs was ≈ 82% of the total encapsulated mass. The protein released from the NPs retained the same molecular mass as free Cpl-711, as proved by the SDS-PAGE analysis (**Figure 4.4.4B**). Furthermore, such released enzyme demonstrated to retain activity against *S. pneumoniae* (**Figure 4.4.4CD**). However, activity assays showed a particular effect: both turbidity decrease (**Figure 4.4.4D**) and bactericidal activity (**Figure 4.4.4C**) were greater at the highest dilutions tested of the first release samples (*i.e.*, dilutions −3 and −4 from the release sample taken after 1 h incubation and dilutions −2 to −4 from the 2 h sample). Only in the 3 h sample, there was an activity decrease upon dilution. Additionally, the Cpl-711 released from samples produced turbidity decrease kinetics with a pronounced sigmoidal shape, in contrast with the exponential decrease of free Cpl-711. One possible explanation for this observation is the presence of degradation products or smaller NPs bearing DEAE groups in the released samples.

[9] The proportion of initially added enzyme that was incorporated into the NPs.
[10] The fraction of the net weight of the freeze-dried NPs that can be attributed to the incorporated enzyme.

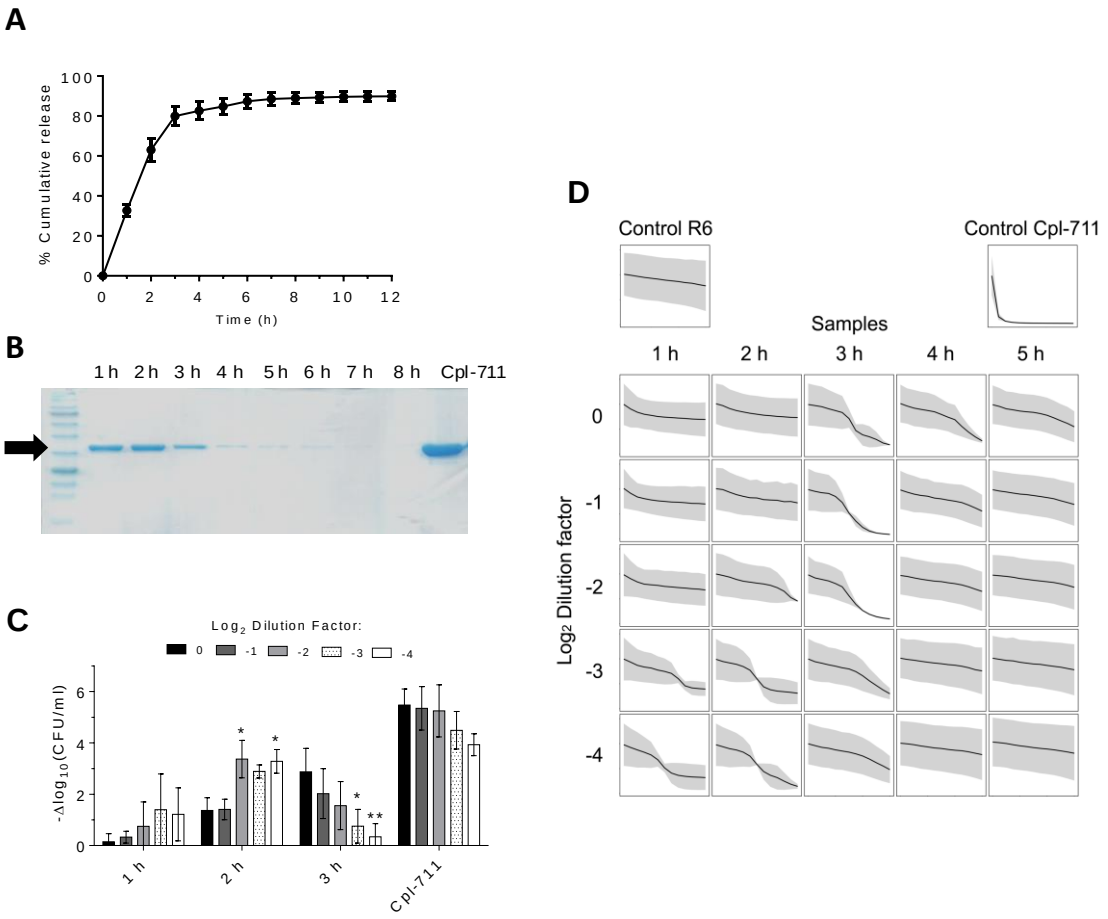


Figure 4.4.4. In vitro release profile of Cpl-711 from ChiDENPs-711. **A**: Cumulative mass release profile of Cpl-711 from ChiDENPs as estimated from three SDS-PAGE replicates (means ± sd are shown). **B**: SDS-PAGE of the samples from the release assay. The arrow shows approximately 39 kDa (corresponding to Cpl-711 molecular weight) and the final lane shows a sample from a Cpl-711 solution at the maximum theoretical concentration (1.5 µg/ml) loaded into ChiDENPs-711. **C**: *S. pneumoniae* R6 killing activity of 2-fold serial dilutions of samples from the *in vitro* release assay (those corresponding to 1 h, 2 h, and 3 h incubation) or a 1.5 µg/mL of Cpl-711 solution. A two-way ANOVA was applied with Tukey post-test to compare the effect of each dilution to that of the undiluted sample for each time (* = $p \leq 0.05$; ** = $p \leq 0.01$) **D**: Turbidity decrease assay of 2-fold serial dilutions of samples from the *in vitro* release assay (1 to 5 h), together with a 1.5 µg/ml Cpl-711 positive control and a NaPiB-treated R6 negative control. Light grey shades represent ± sd of each data point.

This kind of product would be available as ligands for the CBDs of the released Cpl-711. If this would be the case, when put together with pneumococcal cells, these Cpl-711 molecules with their CBDs occupied with DEAE moieties would require a relatively low DEAE to pneumococcal choline ratio to displace the binding equilibrium from the DEAE to the bacterial surface. Therefore, when decreasing the release products concentration by dilution while maintaining the same bacterial concentration, an increase in the bactericidal activity is observed (**Figure 4.4.4C**).

6. BIOCOMPATIBILITY OF CHIDENPS-711

To propose the use of ChiDENPs-711 for in vivo treatment, its biocompatibility was tested by exposing A549 cell cultures to serial dilutions of the release products during different periods (**Figure 4.4.5**).

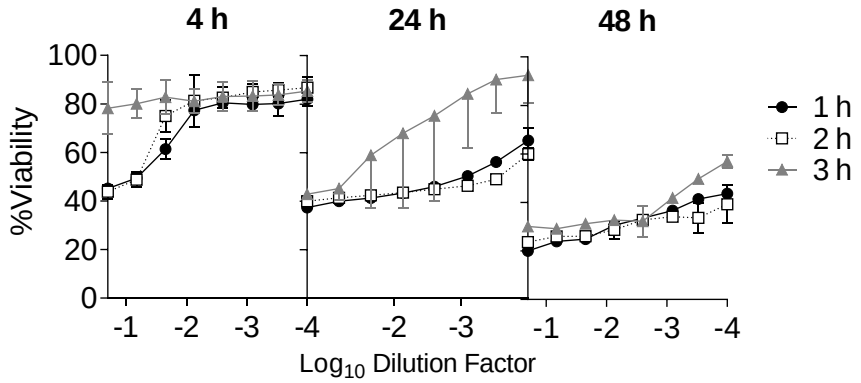


Figure 4.4.5. Cytotoxicity of ChiDENPs-711 release products on A549 cell cultures. Residual % viability is shown with respect to the untreated control after either 4 h, 24 h, or 48 h of exposure to 10-fold serial dilutions of the release product samples corresponding to 1 h, 2 h, or 3 h incubation.

After 4 h incubation, the viability of the cells treated with the 3 h-release products had not significantly decreased (it maintained around 85%), while those of the cells treated with the more concentrated samples (1 h and 2 h) decreased to ≈ 45%. Nonetheless, upon dilution, even with the latter samples the viability rapidly improved. With longer incubation times, however, a certain toxic effect was indeed observed. After 24 h incubation viability ranged between 40% and 85%, and, after 48 h, it was only between 20% and 50%, depending on the dilution factor or the sample. This altogether points out a toxic effect of the released products of ChiDENPs-711 that is dependent on the local concentration and the time of exposure to said products.

V. DISCUSSION

As it has been made clear by the data provided in the INTRODUCTION of this thesis, one of the main health-related problems to be urgently tackled in the years to come are infections caused by antibiotic-resistant bacteria (19). Among many other issues that should be combined in multipronged strategies involving many public actors, the development of novel antimicrobial agents is a priority (387). Such a new generation of antimicrobials would fill the void left by the dried antibiotic pipeline, providing alternative and complementary tools to treat bacterial infections. The current state of affairs suggests that the better positioned technological alternative to antibiotics to be implemented in the short term is that posed by phages and their derived products (388). While phage therapy itself is currently making important advances into the clinical practice by opening new paths in the way we understand antimicrobial therapeutics both from a biological and a regulatory point of view (54, 60), the so-called enzybiotics may better fit into the current paradigm of pharma industry business ecology (54). This is one of the reasons why the research in repurposing phage lysins as enzybiotics has tremendously thrived in recent years. But not only the possible economic and regulatory framework for enzybiotics seems better suited to an imminent entry to the market, almost assured by the phage lysins currently undergoing phase III clinical trials[11]. Also, enzybiotics present some critical advantages over antibiotics and even whole virions that make them, in our opinion, especially suited for a post-antibiotic era, as shown in **Table 1.1**. An obvious example is their experimentally proven lower chance to induce the selection of enzybiotic-resistant bacteria since modification of the cell wall and/or the targeted essential elements of the bacterial surface comes with a great cost for the bacterial fitness (81). In addition, the modular architecture of phage lysins is being profited for protein engineering strategies that rely on the long-time proposed approach of 'domain shuffling' to create chimeric proteins with novel emergent properties (389). The current state-of-the-art of domain shuffling has moved towards the high-throughput, type IIs restriction enzyme-based genetic engineering, which allows creating wide chimeric protein libraries to be screened looking for some pre-determined traits (154). So, taken all these advances together, the real question today in lysins research is: where do we take enzybiotic-based therapy in the long term? How can enzybiotics contribute to the revolution of antimicrobial therapy to come? As already hinted in the INTRODUCTION, J. P. Pirnay creatively paints a not-so-science-fiction future in which integrated technologies could automatically provide a personalized phage-based drug directly delivered to treat the specific infection of a specific patient (60). It could be argued that not only full virion therapy can fit into this new

[11] https://www.contrafect.com/pipeline/exebacase

personalized framework, but also lysins constructed by modular engineering, provided that we can count on a strong knowledge base on the structure-function relationship of lysins.

Many benefits can be derived from engineering-based knowledge production. For example, the aforementioned high-throughput strategies are already providing rapidly and easily derived 'rules' for lysin engineering (155). Rules that the classical production of chimeras would only produce after exhausting efforts. But such eminently applied (or top-down) approaches still need to be properly contextualized by a more basic research, which would provide the appropriate foundation for a productive interpretation and implementation of the technological advances. In this thesis, we have adopted an integrated approach that encompassed the elucidation of a data-sustained paradigm onto which to build the whole process of designing and preclinically developing an antimicrobial. Thus, we have intended both to take advantage of the novel, holistic knowledge-derivation techniques (such as the data mining presented in **Chapters 1** and **2**) and also of a more traditional, knowledge-driven design of phage lysin antimicrobials (as developed in **Chapters 2** and **3**).

In **Chapter 1**, the construction and analysis of an exhaustive database comprising 2,182 lysin sequences have produced a series of relevant conclusions towards both understanding phage-mediated bacterial lysis and designing phage-derived antimicrobials. To begin with, important architectural differences between lysins from phages that infect either G+ or G− bacteria have been brought out, thus probably correlated to the fundamental difference between the G− and the G+ envelope. The primary difference between such groups of lysins is that while G+ ones are typically multimodular, their G− counterparts usually have just a single EAD. Indeed, most carbohydrate enzymes known to date pose a CWBD that enables them to approach their substrate, since polysaccharidic substrates are insoluble and thus non-diffusible (390). Therefore, if the enzyme is not able to encounter its target substrate in solution, it must be able to approach it. It does so, then, by possessing a tropism towards the polymeric substrate, which lies on the CWBD. It can be speculated that since G− bacteria contain just a thin peptidoglycan layer, the passive pumping of massive amounts of lysins may be enough to degrade this single layer without needing an additional affinity towards the insoluble substrate. However, our results in **Chapter 2** serve to hypothesize the presence of a substrate-binding function in at least one family of EADs (the *Muramidase* PF family) that make up the lysins of many G− phage lysins. Moreover, the proposed essential residues for the affinity towards peptidoglycan are only conserved in those G− phage lysins that do not contain a separate CWBD. These observations altogether point out an evolutionary compensation for the lack of a distinct CWBD. Therefore, the typically maintained hypothesis of G− phage

lysins lacking or not needing a cell wall binding function should be perhaps revised once sufficient evidence is gathered on i) whether these regions function as true CWBDs and ii) how widespread this trait is. If we accept that monomodular lysins may contain internal cell wall binding subdomains, it is tempting to speculate that this may be explained by the existence of two differential evolutionary routes: one that has evolved multimodular lysins by fusing domains through horizontal gene transfer (148); and another in which functional, substrate-binding monomodular enzymes are evolved by fixation of certain beneficial mutations within the EAD.

The multimodular architecture of G+ phage lysins has even given rise to lysin architectures in which not just two but three domains are involved. In many cases, trimodular lysins bear two EADs and a CWBD. A possible explanation for multicatalytic lysins is an increased lytic efficiency over monocatalytic ones since activities attacking different sites of the peptidoglycan are known to act synergistically in peptidoglycan degradation, as our results have previously demonstrated (86). Since the G+ peptidoglycan layer is much thicker than the one of G− bacteria, it is conceivable that such synergistic cooperation may pose a significant benefit towards efficient lysis in G+ but not in G−. Such synergy could also imply a decreased chance for the appearance of resistant peptidoglycan mutants (391). It has also been shown that the synergistic concurrence of both activities is sometimes needed for full activity. Thus, some phages may have evolved a regulatory mechanism to avoid the lysis of other potential host cells relying on the proteolysis of bicatalytic lysins by host-cell proteases. Then, both EADs would be disjointed by proteolysis upon host cell lysis and the degraded lysins would no longer be active against the nearby bacterial population (392). This should be especially relevant for phages infecting G+ bacteria, which lack a protective OM hindering the lysis of other bacterial cells from without, and hence the exclusiveness of the bicatalytic architecture among phages infecting G+. In some other cases, however, it is the high affinity of the CWBD that has been proposed as the mechanism that maintains lysins tightly bound to cell debris preventing widespread lysis of the bacterial community (301), which is also an argument for the widespread presence of CWBDs among G+ and not among G–. Additionally, it has also been shown that, at least in some cases, the central domain predicted as an EAD may have an auxiliary role in substrate binding, rather than being an actual EAD (393).

These architectural differences uncovered by our analysis of lysin sequences just remark the importance of bacterial surface to make sense of phage lysins structural-functional specialization. Even peptidoglycan structure (or chemotype) has been pointed out by our data as a possible determinant for the appearance of specific traits in the corresponding lysins. For example, it has been shown that while peptidases are very diverse among lysins from phages that infect bacteria with A1 subtype peptidoglycan (lacking an

interpeptide bridge between peptide subunits), it is not the case in bacteria with an interpeptide bridge in their peptidoglycan. In A3-subtype bacteria phage lysins, the only generally predominant peptidase would be the *CHAP* family. It makes sense that a more complex peptide moiety (such as that found in the A3 chemotype versus A1) would have driven the evolution of a specific peptidase. On the other hand, of course, there are examples of widely distributed domain families, such as *LysM* or *Amidase_2*, that probably target generally conserved bacterial surface traits. It is, however, risky to state anything about the ecological importance for a phage lysin to be highly specific to the bacterial host or somewhat wide-ranged, since it would probably be case-dependent.

Among the evolutionary specializations of lysins, which are already quite evident, the analyses presented in **Chapter 1** have uncovered some traits of a physicochemical nature. Such differential physicochemical properties have allowed us automatically distinguishing between lysins from G– and G+ infecting phages using a machine learning algorithm. We can conclude from our results that such difference resides in the presence of a cationic C-terminal region, akin to AMPs, which is present at least in a subpopulation of G– phage lysins well represented in our dataset. This observation does not imply that the presence of an AMP-like region is a *sine qua non* feature of G– lysins but rather suggests that it is widespread among the currently known G– phage lysin architectural variants. To illustrate this with some numbers, our results in **Chapter 2** indicated that an AMP could be predicted in at least 32% of the G– entries of the dataset. Such a result is sure biased by the selection of the specific region that was screened by our algorithm within the sequences, as also shown by the results in **Figure 4.2.3**. Consequently, chances are that this number underestimates the actual presence of AMP-like subdomains, as hinted by the results in **Chapter 1** (particularly, by the association of the cationic C-terminal end with several widespread domain families in G– phage lysins).

As for the possible biological explanation for the presence of AMP-like regions in G– lysins, several hypotheses can be proposed. For example, it could be argued that they may take over the function of a typically lacking CWBD, by granting a tropism towards the cell surface. But two reasons prevent us from arguing that this should be the case: i) a positively charged region would be surely useful for the lysin to approach the bacterial surface ‘from without’, given the negative charge that the phosphate groups confer to the LPS that makes up the outer face of the OM. But it is unclear how this would benefit a lysis ‘from within’, considering that the peptidoglycan itself bears no charge and that the features of the AMP-like region make it reportedly able to interact with membranes, but not with polysaccharides nor proteins (163, 174, 394); ii) the already mentioned possibility of the presence of a cell wall binding subdomain within the EAD would make a second cell wall binding region redundant. Another proposal, bearing in mind the aforementioned

ability of AMP-like regions to interact with membranes, would be that they could act somewhat as spanins, providing a complementary or substitutive membrane-disturbance ability necessary for the phage-mediated lysis of G– cells (71). Indeed, there have been described some G– phages that lack spanins (and have evolved alternative mechanisms to compensate) (395). Since spanins have been suggested to be essential for lysis, their role could be fulfilled by AMP-like regions in lysins. This possibility could be supported by a preferential appearance of AMP-like subdomains in lysins from 'spaninless' phages. However, although this should be closely examined in the available data, the variety of lysis cassettes found in the AMP-positive candidates derived from the screening presented in **Chapter 2**, many of which contain plausible spanin candidates, urges to tone down the generalization of this proposal. Finally, since there are also proofs in the literature of some lysins being able to transverse the inner membrane by themselves (without the concurrence of a holin), and even linking this ability to AMP-like subdomains (162, 394), AMP-like regions may also have a holin-like function. Whichever the case, and given that they appear preferentially among G–, it is clear to us that their evolutionary sense arises in the context of the G– cell wall, and, specifically, in the presence of a dual membrane structure. Therefore, while at this point the available evidence does not allow proposing nor discarding one hypothesis or another, the conclusion may be that AMP-like regions enhance the interaction with G– membranes and plausibly plays a role in the execution of efficient lysis.

Taking forward this message, in **Chapter 2** we used the prediction of AMP-like regions to obtain enzybiotic candidates foreseen to have an increased ability to interact (and perhaps disrupt) the *P. aeruginosa* OM as well as the peptidoglycan. The finally selected candidate, Pae87, was able to interact with the G– bacterial surface, as shown, for example, by its aggregative effect (165) and by its intrinsic bactericidal activity. Regarding the bactericidal activity of Pae87, certainly, the killing values reached (1-3 log reduction in CFU/ml with respect to the untreated control) cannot compete with the bactericidal efficiency displayed by *ad hoc* engineered lysins: *e.g.*, artilysins Art-175 or 1D10 achieve 4-5 log reduction against *P. aeruginosa* or *A. baumannii* with 3 to 10 times less amount of enzyme (154, 180). However, its bactericidal potency is comparable to that of some other previously reported lysins intrinsically active against G− from without. For example, lysin OBPgp270 caused a 1-log decrease in viable cells at a concentration of 1.5 µM against *P. aeruginosa* PAO1 in the absence of EDTA and 4-log killing with 0.5 mM EDTA, very much like Pae87 (179). In this regard, the bactericidal activity results obtained with our candidate are within the expected range. The fact that Pae87 was able to solubilize purified *P. aeruginosa* peptidoglycan and the degradation peaks observed in the RP-HPLC analysis confirmed that Pae87 is a peptidoglycan hydrolase. Moreover, the

observed peaks are compatible with the literature reports stating that the members of the *Muramidase* family, to which Pae87 EAD belongs, are lysozymes. The experimental data provided in this thesis, however, indicate that the membrane activity is the major determinant for Pae87 antimicrobial potential, rather than the catalytic activity, as pointed out for other intrinsically active lysins (166). In this work, a specific AMP-like C-terminal region (P87) with intrinsic membrane-permeabilizing and bactericidal activity has been identified as the most probable part of the enzyme responsible for the aforementioned effect. The three-dimensional model provided by the CNB "Structural Biology of Viral Fibers" group confirmed that P87 was located at the surface of the protein, supporting the hypothesis that it would be able to directly interact with membranes. Moreover, Pae87 had an OM permeabilizing activity of its own. Based on these results, a mechanism for Pae87 activity from without was proposed: i) the P87 region of the enzyme would coat the OM and cause the aggregation of adjacent cells; ii) then the membrane-permeabilizing action would act, probably together with the peptidoglycan hydrolysis activity, to disrupt the cell wall, iii) causing the leakage of intracellular components and cell death without provoking a full disintegration of the bacteria ('lysis'), but rather keeping the cell debris tightly bound in compact aggregates. As hinted before, this 'death without lysis' could be beneficial from the point of view of *in vivo* therapy, since it would prevent the dissemination of pro-inflammatory factors. Also, bacterial aggregates have been previously shown to be better cleared by the immune system (235, 357). However, the true potential of this kind of antimicrobial agent should be tested *in vivo* to clarify its possible benefits.

On another hand, using the Pae87 3D model and sequence analysis based on MSAs, two acidic residues within the catalytic cleft have been identified as relevant for the peptidoglycan hydrolase activity. One of such residues has been for the first time identified in the *Muramidase* family, which, in any case, given its recent discovery, must still be fully described. It can be stated, anyway, that the structure of Pae87 and, in general, the *Muramidase* family, is somewhat unusual due to the large distance between the catalytic residues. The presence of such a large catalytic cleft may indicate an unusual, alternative catalytic mechanism. Alternatively, the two catalytic lobes could approach down to a more canonical distance enabling a classical glycoside hydrolysis mechanism, provided that the hinge between such lobes would allow some flexibility. Knowledge of the precise mechanisms that operate in the activity of phage lysis is important towards a more precise, even 'surgical' engineering of enzybiotics. For example, the results achieved by increasing the surface interaction potential of Pae87 by fusing it to a calcium-binding domain (in the Pae87F chimera, **Chapter 3**) suggest that no important benefits could be gained with this engineering strategy in this specific case. But, perhaps, an enhancement

of the hydrolytic activity by tuning the structural composition of its catalytic center may have yielded a better antimicrobial activity. In any case, the results of this thesis show, for the first time, that cation-binding domains can also be considered as enzybiotic activity enhancers in the future, perhaps rendering better results when coupled to other types of EADs.

In **Chapter 3**, we intended to unfold the true potential of Pae87 and its derivatives as antimicrobial agents. To that end, P87 was rationally engineered to maximize its HM and net charge. As a result, the derivative peptide P88 was obtained. This peptide achieved a greater bactericidal efficiency (our results indicate that the killing efficiency was 10-fold increased with respect to the parental peptide). For many of the bacterial strains tested, 10 μM P88 caused a decrease of the viable cell counts below the detection limit of the experiment (*i.e.*, ≤ 10^2 CFU/ml). This altogether makes P88 the better antimicrobial agent candidate proposed in this thesis in terms of bactericidal power. Additional experiments provided by the S. Albertí group of the UIB also proved that P88 had no major resistance determinants among a somewhat wide collection of *P. aeruginosa* strains, furthermore highlighting the potential of P88 as an antimicrobial. Nonetheless, if P88 were to be used in an *in vivo* therapy, the main risk would be its cytotoxicity. A toxic effect of AMPs is often cited as one of their drawbacks for systemic treatment. Our results with A549 epithelial lung cell cultures demonstrate that, indeed, P88 has a cytotoxic effect, as does P87. However, while P88 is ten times more active on bacteria than P87, the toxicity increase of P88 over P87 is not as high. Therefore, our path towards positioning P88 as a valid antimicrobial was to test its ability to enhance bacterial susceptibility to antibiotics via synergistic cooperation between antibiotic and peptide. Several antibiotics (CHL, TET, ERY, and AZI) showed synergy when tested together with P88, achieving viable cell counts below the detection limit when the synergistic combination was applied in a range of *P. aeruginosa* strains and lowering the peptide concentration needed for growth inhibition to 0.25 × MIC (12.5 μM) in the synergy checkerboard assays. At such concentrations, P88 only displayed a mild cytotoxic effect (viability of ≈ 75% with respect to the untreated control). The fact that synergy was detected in the four strains tested also suggests that it may not be a strain-dependent effect. With these data, P88 could be tested *in vivo* in combination with antibiotics as a viable therapy. Nonetheless, as pointed out in the INTRODUCTION, other rational engineering approaches could be adopted to improve the therapeutic index of P88. For example, such strategies as diminishing the angle subtended by the positive charge at the polar side of the helices or introducing D-enantiomers have been shown to reduce cytotoxicity and increase bacterial selectivity (207, 208). Regarding the synergy, it has been many times described that AMPs may act synergistically with antibiotics, especially those which have intracellular targets, and

thus explaining the synergy as a mechanism in which the AMP improves the cell envelope permeability and thus the antibiotic gains access to its target more easily (378). The synergy displayed with macrolides (ERY, AZI) is interesting on its own, since, although *P. aeruginosa* strains usually show rather high *in vitro* MIC values for macrolides, this kind of antibiotics are successfully used in the clinic to treat *P. aeruginosa* respiratory infections. The reason behind this is that macrolides seem to impede the production of virulence factors (396). It has also been shown that the macrolides MIC determined for *P. aeruginosa* strains in media supplemented with biological fluids is much lower than in the commonly recommended medium (Mueller-Hinton broth). This is due to an enhanced OM permeability in such supplemented media, which would thus explain to a certain point the success of macrolides in anti-infective chemotherapy against *P. aeruginosa* (380). The trade-off of this phenomenon is that a further increase of OM permeability, such as that provided by P88, may be redundant *in vivo*. Although our results show that synergy is still present when a blood lysate is added to the testing medium, the synergy of P88 with macrolides should be tested *in vivo* to ascertain its true therapeutic potential.

Finally, **Chapter 4** dealt with the possibility of further engineering enzybiotics together with other elements, such as polymeric scaffolds, to improve the *in vivo* performance of a lysin-based therapeutic preparation. In this case, we have followed a rather popular design strategy that has been termed ‘biomimicry’ (397). To that end, a chitosan-based polymer bearing DEAE moieties was produced (by the M. R. Aguilar group at the ICTP) to emulate the multivalent disposition of choline on the pneumococcal surface. The final aim was to obtain a nanoparticulate material able to bind CBPs, such as pneumococcal surface does. The results obtained with our ChiDENPs preparations, in comparison with the ChiNPs, clearly show that the derivatization with DEAE has an enormous impact on how NPs behave in the presence of elements from the pneumococcal system (*i.e.*, the proposed aim was achieved). To begin with, the multivalent disposition of DEAE was shown to efficiently bind housekeeping CBPs and provoke pneumococcal chaining, which could be profitable for therapy (235, 357). As a side-effect of the integrated ChiDE-CBP system, the protein-coated ChiDENPs were more stable, less prone to aggregation than NPs bearing proteins bound just by unspecific interactions (in the case of ChiNPs, electrostatic interaction is proposed to be the main interactive effect given the opposite charges of the tested proteins and the NPs surface). The ChiDENPs loaded with Cpl-711 have been shown to sustainedly release the cargo, at 37 °C, in a 3 h period. This time span has the potential to increase the systemic *in vivo* half-life of lysins, which has been reported to be in the range of 20-60 min (149). In addition, the results of this thesis prove that the released enzyme still retains activity against *S. pneumoniae*. These released products, however, display a peculiar behavior: the first supernatants collected

during the release experiment have a higher bactericidal effect when they are the most diluted. This phenomenon can only be explained by the presence of an inhibitory compound in the released fractions. Since the system contains DEAE, which is a CBPs activity competitive inhibitor (as shown by the pneumococcal chaining effect of the ChiDENPs), it is plausible to assume that the presence of either smaller ChiDENPs deficiently separated by the centrifugation step or either low molecular weight fragments of ChiDE produced by the degradation of the ChiDENPs may be present in such supernatants. In an *in vivo* setting, where the dilution rate of the released products may be higher than in an *in vitro* experiment within a confined vessel, and there is a high concentration of pneumococcal cells due to infection, the ratio of DEAE vs pneumococcal choline should be low enough to allow the action of the enzybiotic. Whichever the case, this should not be an issue when using other non-CBPs tagged with an additional CBD to be used in the ChiDENPs system. The cytotoxicity results support a model in which the enzyme is released because of or accompanied by the degradation of the polymer. Since the first released supernatants (the ones collected in hours 1 and 2) provoke a higher cytotoxic effect they must have a higher concentration of the cytotoxic product. Such toxic product cannot be the enzyme, since enzybiotics have been repeatedly shown to be non-toxic *in vivo* (due to the bacterial and phage populations that cohabit in our organism, we are constantly exposed to these enzymes). Therefore, ChiDENPs-derived products should be present in greater amounts in the first samples, *i.e.*, they are being collected at a similar pace as the released protein. Additional experiments should cast some light on the actual release mechanism, but current literature points to host lysozymes as the effectors of chitosan biodegradation *in vivo* (398, 399). Therefore, it is not unreasonable to propose that Cpl-711 may be contributing to the ChiDENPs degradation since it contains a lysozyme activity. Moreover, a previous attempt to load Cpl-1 (a parental enzyme of the chimeric Cpl-711) into ChiNPs reported a burst release time of 12-14 h (210). The release behavior observed in our system might be perhaps explained by a faster autodegradation of the ChiDENPs-711 due to the presence of DEAE moieties that could potentially improve the interaction between Cpl-711 lysozyme and the polymer. In this sense, ChiDENPs-711 degradation and, hence, Cpl-711 release could be tuned by modifying the substitution degree of DEAE moieties in the polysaccharide backbone. These pending questions deserve future investigation. On the other hand, certain toxicity is to be expected in the case of polycationic polymers, according to the literature. For example, for dextran-DEAE, similar concentration- and time-dependent toxicity has been observed (400). Toxicity has also been shown to be dependent on the molecular weight of the polymer (it increases with the molecular weight) (400, 401); therefore, future developments might include chitosan modifications such as using lower MW polymer,

modifying it to facilitate the breaking of the chains and/or reducing the DEAE substitution degree. With the current formulation, the therapeutic utility would heavily depend on a hypothetical therapeutic window in which the active cargo is released upon reaching the infection site while degraded ChiDE is quickly cleared and excreted in such a way that the antimicrobial benefit surpasses the toxic hazard. Current literature suggests that ChiNPs are absorbed and excreted rapidly, mostly in the first hours after pulmonary administration (402, 403), but, again, *in vivo* confirmation shall be provided.

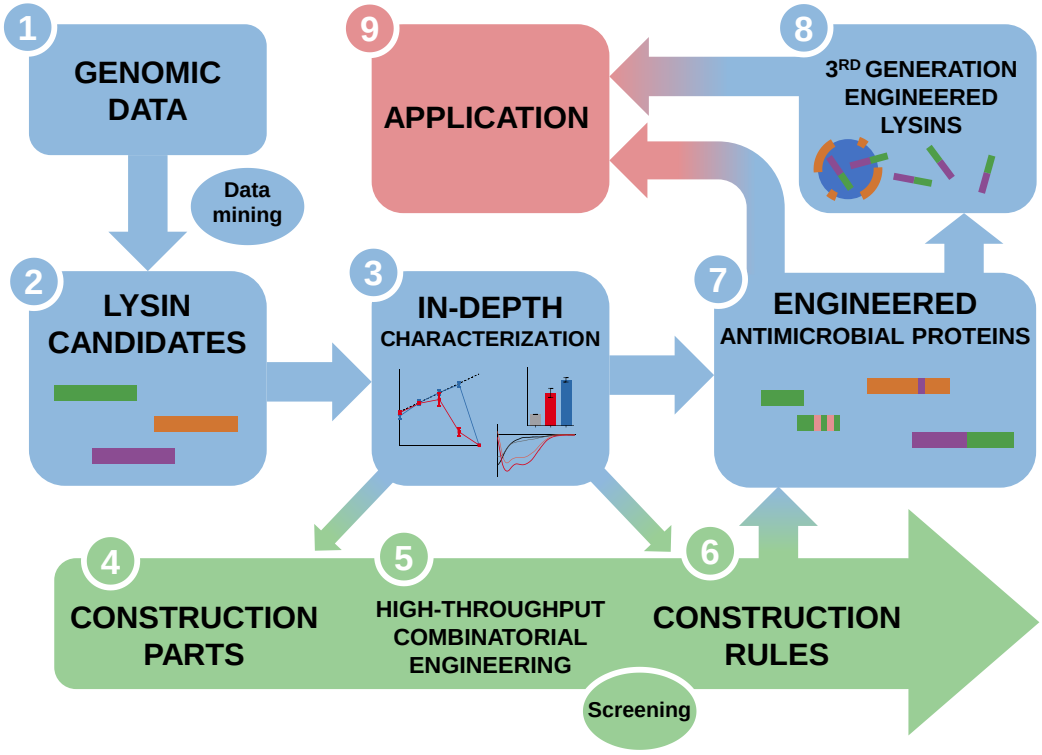
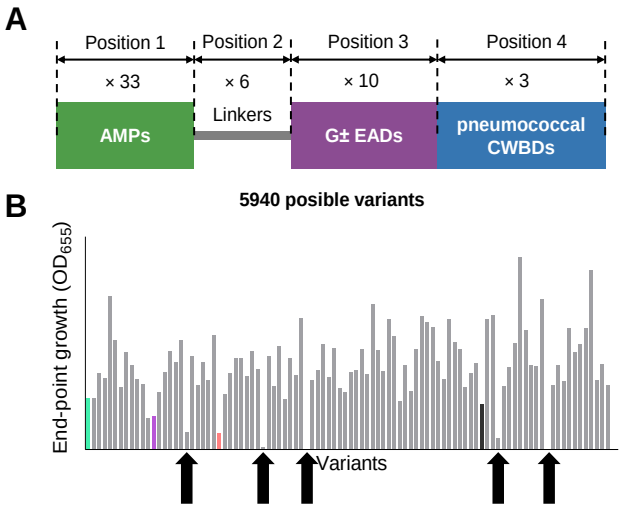


Figure 5.1. Workflow for the development of lysin-based antimicrobials under the current state-of-the-art.

To wrap up, the different strategies for the design and development of protein antimicrobials presented in this thesis aim to contribute to a novel workflow for the development of lysin-based antimicrobials (**Figure 5.1**). Such workflow shall benefit from the newest developments lately provided by the lysin research community. For example, as demonstrated here, the massive amounts of genomic information currently available online can be mined based on literature to obtain enzybiotic candidates foreseen to possess some predetermined features. Such candidates, very much as conducted throughout this work, can be then in-depth characterized and such information can be used either to design and develop engineered proteinaceous antimicrobials or to define knowledge-driven

strategies for future engineering efforts. On the other hand, the state-of-the-art high-throughput combinatorial engineering platforms, such as VersaTile (154) shall benefit from the previously characterized elements (modules) to use them as construction parts of multimodular lysin variants. Such combinatorial libraries can be experimentally screened to detect architectural variants of interest from which sets of 'construction rules' can be derived. Of course, the screening method selected determines the construction rules that are discovered and thus such rules must be semantically tagged and characterized to provide the desired results when applied *ab initio*.



C

Variant	Position 1	Position 2	Position 3	Position 4
1.1αB2	Peptide 1	Linker 1	EAD 1	CWBD 1
1.1αD11	Peptide 2	Linker 2	EAD 2	CWBD 2
1.1βF4	Peptide 3	Linker 3	EAD 3	CWBD 2
1.1γB7	Peptide 2	Linker 2	EAD 3	CWBD 2
1.1γC9	Peptide 2	Linker 2	EAD 2	CWBD 2
1.1γD5	Peptide 4	Linker 3	EAD 2	CWBD 2

Figure 5.2. Summary of a typical high-throughput combinatory cloning experiment. A: Design of the library. For each position, a certain number of different modules are added to the reaction mixture. **B:** Screening experiment (growth inhibition measurement of an *S. pneumoniae* culture treated with the protein variants expressed in a 96-well plate). The arrows indicate those wells whose protein variant was able to inhibit pneumococcal growth. **C**: Domain structure of some positive hits derived from the screening experiments. Experimental details can be found in (154).

During the course of this thesis, I did a short stay to learn about the high-throughput modular genetic engineering technology developed by Yves Briers' group at the University of Ghent (the VersaTile platform). Due to the recent circumstances, the results we obtained there could not be fully developed and included in this thesis. But, just as an example of how the derivation of 'construction rules' from the high throughput semi-random combination of modules works, some results are presented in **Figure 5.2**. In these experiments, EADs and CWBDs from the pneumococcal system were combined with AMPs and domains derived from G– phage lysins, with a pre-defined order within the constructed variants (**Figure 5.2A**). The initial purpose was to obtain viable structural variants with a dual activity both against *S. pneumoniae* and other G– pathogens that sometimes cause mixed infections of the respiratory tract (such as *M. catarrhalis* or *H.*

influenzae) (404). The full library of variants was tested against *S. pneumoniae* and those who rendered a growth inhibition activity against this bacterium were sequenced (**Figure 5.2B**). In this way, a set of anti-pneumococcal lysins bearing specific AMPs and linkers at the N-terminal region was obtained (**Figure 5.2C**). The repeated structures across the positive hits yield certain preferred construction parts that, combined in a determined order are the construction rules.

From any of the approaches shown in **Figure 5.1** (rational, single-protein engineering; semi-random combination of modules and variants screening; construction rule-based rational design of lysins), *ad hoc* engineered enzybiotics can be derived. Then, such enzybiotics may be directly applied as enzybiotics or be further engineered to generate 'third generation' drugs (encapsulated, deimmunized, etc.). The collective efforts in such a research scheme could lead to a 'tailor-made' paradigm in which (**Figure 5.3**): i) the 'problem' is precisely identified (a specific pathogen and its antibiotic susceptibility/resistance, a given niche, some application); ii) the construction rules are applied via a bioinformatically-aided procedure to decide on the protein architecture that better fits the predefined problem; iii) the protein-encoding gene is constructed by modular genetic engineering and then biotechnologically produced; iv) the final product and its defined formulation can be finally applied.

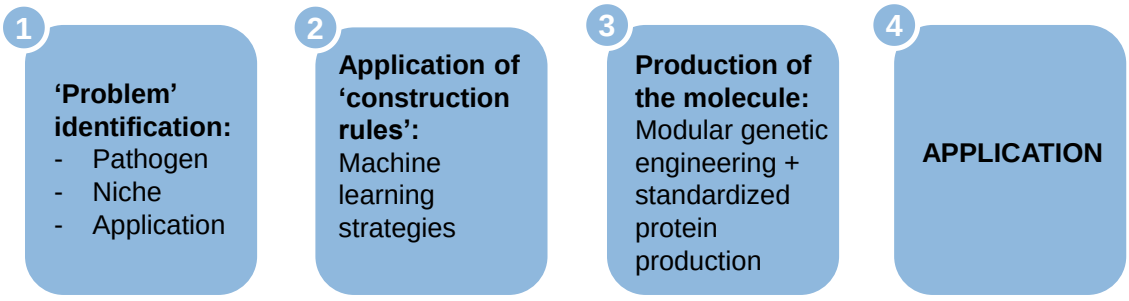


Figure 5.3. The possible tailor-made framework for lysin-based enzybiotics.

As a generic conclusion, the current advances in phage-derived antimicrobials science give many reasons to be optimistic about a future in which antibiotic resistance can be overcome by the market entrance of these kinds of products. Phage-derived therapies, as it has been shown, also threaten radically changing the way we understand infection treatment. The results and discussion presented in this thesis aim to be a part of this innovation.

VI. CONCLUSIONS

CONCLUSIONS

1. Through an exhaustive analysis of a database of phage lysin sequences, several correlations between lysins architecture and the features of the phage bacterial hosts have been uncovered. Specifically, it has been found that a relevant sub-population of Gram-negative phage lysins contains antimicrobial peptide-like elements in its C-terminal end.
2. The prediction of antimicrobial peptides within the sequence of Gram-negative phage lysins has been proved as a method for screening and selecting outer membrane-interacting lysins as enzybiotic candidates.
3. One of such candidates, Pae87, has a peptidoglycan degradation activity and an intrinsic antibacterial activity when exogenously applied against certain Gram-negative bacterial pathogens.
4. Pae87 harbors within its structure an unusually large catalytic center with two glutamic acid catalytic residues, a possible peptidoglycan binding region, and a C-terminal antimicrobial peptide-like region termed P87, which has an antibacterial activity on its own against a similar range of pathogenic bacteria, and it is proposed to be responsible for the intrinsic activity of Pae87.
5. The rational mutation of P87 to generate the derived peptide P88, with increased predicted net charge and hydrophobic moment, achieved an enhanced bactericidal efficiency without altering its antibacterial range nor dramatically increasing its cytotoxic effect, and showed *in vitro* synergy with some antibiotics.
6. The fusion of Pae87 to a cation-binding domain produced a chimeric protein (Pae87F) with an increased bactericidal effect *in vitro*, providing a novel source of enhancing modules for generating novel chimeric enzybiotics.
7. In this way, a complete workflow for the design and development of proteinaceous antimicrobial agents has been tested, taking advantage of both the many bioinformatic tools and source materials available and the experimentally-driven knowledge.
8. Chitosan-DEAE nanoparticles mimic the choline-binding protein affinity of the pneumococcal cell wall and efficiently bind these proteins due to the multivalent disposition of the DEAE moieties.
9. Such nanoparticles can be used as a scaffold for loading active choline-binding proteins with good stability of the nanoparticulate suspension due to the affinity-mediated interaction and such loaded proteins are then sustainedly released.

CONCLUSIONES

1. Mediante un análisis exhaustivo de una base de datos de secuencias de lisinas fágicas se han descubierto varias correlaciones entre la arquitectura de las lisinas y las características de los hospedadores bacterianos de sus fagos. Específicamente, se ha encontrado que una subpoblación relevante de lisinas de fagos de Gram-negativas contiene elementos similares a péptidos antimicrobianos en su extremo C-terminal.
2. Se ha demostrado que la predicción de péptidos antimicrobianos dentro de la secuencia de lisinas de fagos de bacterias Gram-negativas puede aplicarse para cribar y seleccionar lisinas que interactúen con la membrana externa como candidatos a enzibióticos.
3. Uno de dichos candidatos, Pae87, presenta actividad de degradación del peptidoglicano y una actividad antibacteriana intrínseca cuando se aplica exógenamente contra ciertos patógenos bacterianos Gram-negativos.
4. Pae87 comprende en su estructura un centro catalítico inusualmente grande con dos residuos catalíticos de ácido glutámico, una posible región de unión a peptidoglicano y una región similar a péptidos antimicrobianos en su extremo C-terminal que se propone como responsable de la actividad intrínseca de Pae87.
5. La mutación racional de P87 para generar el péptido derivado P88, con mayor carga neta y momento hidrofóbico predichos, logró una mejor eficiencia bactericida sin alterar el rango antibacteriano ni incrementar drásticamente su efecto citotóxico, y mostró sinergia *in vitro* con algunos antibióticos.
6. La fusión de Pae87 a un dominio de unión a cationes dio lugar a una proteína quimérica (Pae87F) con un efecto bactericida aumentado *in vitro*, proporcionando así una nueva fuente de módulos potenciadores de la actividad para generar nuevos enzibióticos quiméricos.
7. De este modo, se ha probado un flujo de trabajo completo de diseño y desarrollo de agentes antimicrobianos de naturaleza proteica aprovechando tanto las muchas herramientas e información bioinformáticas disponibles como el conocimiento impulsado experimentalmente.
8. Las nanopartículas de quitosano-DEAE imitan la afinidad por proteínas de unión a colina de la pared celular neumocócica y se unen a este tipo de proteínas de manera eficiente debido a la disposición multivalente de los residuos de DEAE.
9. Dichas nanopartículas pueden usarse como soporte para cargar proteínas de unión a colina activas con una buena estabilidad de la suspensión de

nanopartículas debido a la interacción mediada por afinidad, y dichas proteínas cargadas se liberan de las partículas de forma sostenida.

VII. REFERENCES

1. GBD 2016 Lower Respiratory Infections Collaborators. 2018. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis 18:1191-1210.
2. World Health Organization. 2018. Global health estimates 2016: deaths by cause, age, sex, by country and by region, 2000-2016. World Health Organization, Geneva.
3. World Health Organization. 2019. Pneumonia. World Health Organization,
4. Rodrigues CMC, Groves H. 2018. Community-acquired pneumonia in children: the challenges of microbiological diagnosis. J Clin Microbiol 56:e01318-01317.
5. Torres A, Cilloniz C, Blasi F, Chalmers JD, Gaillat J, Dartois N, Schmitt HJ, Welte T. 2018. Burden of pneumococcal community-acquired pneumonia in adults across Europe: A literature review. Respir Med 137:6-13.
6. Langford BJ, So M, Raybardhan S, Leung V, Westwood D, MacFadden DR, Soucy JR, Daneman N. 2020. Bacterial co-infection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis. Clin Microbiol Infect 26:1622-1629.
7. Bardi T, Pintado V, Gomez-Rojo M, Escudero-Sanchez R, Azzam Lopez A, Diez-Remesal Y, Martinez Castro N, Ruiz-Garbajosa P, Pestana D. 2021. Nosocomial infections associated to COVID-19 in the intensive care unit: clinical characteristics and outcome. Eur J Clin Microbiol Infect Dis 40:495-502.
8. Lanks CW, Musani AI, Hsia DW. 2019. Community-acquired pneumonia and hospital-acquired pneumonia. Med Clin North Am 103:487-501.
9. Jones RN. 2010. Microbial etiologies of hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia. Clin Infect Dis 51 Suppl 1:S81-87.
10. Kidd JM, Kuti JL, Nicolau DP. 2018. Novel pharmacotherapy for the treatment of hospital-acquired and ventilator-associated pneumonia caused by resistant gram-negative bacteria. Expert Opin Pharmacother 19:397-408.
11. Corrado RE, Lee D, Lucero DE, Varma JK, Vora NM. 2017. Burden of adult community-acquired, health-care-associated, hospital-acquired, and ventilator-associated pneumonia: New York city, 2010 to 2014. Chest 152:930-942.
12. Cilloniz C, Ewig S, Polverino E, Marcos MA, Esquinas C, Gabarrus A, Mensa J, Torres A. 2011. Microbial aetiology of community-acquired pneumonia and its relation to severity. Thorax 66:340-346.
13. Kollef MH, Chastre J, Fagon JY, Francois B, Niederman MS, Rello J, Torres A, Vincent JL, Wunderink RG, Go KW, Rehm C. 2014. Global prospective epidemiologic and surveillance study of ventilator-associated pneumonia due to *Pseudomonas aeruginosa*. Crit Care Med 42:2178-2187.
14. Abraham EP, Chain E. 1988. An enzyme from bacteria able to destroy penicillin. 1940. Rev Infect Dis 10:677-678.
15. Davies J. 1995. Vicious circles: looking back on resistance plasmids. Genetics 139:1465-1468.
16. European Centre for Disease Prevention and Control. 2020. Antimicrobial resistance in the EU/EEA (EARS-Net) - Annual Epidemiological Report 2019. European Centre for Disease Prevention and Control, Stockholm.
17. CDC. 2019. Antibiotic resistance threats in the United States. U.S. Department of Health and Human Services, CDC, Atlanta, GA.
18. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Carmeli Y, Ouellette M, Outterson K, Patel J, Cavaleri M, Cox EM, Houchens CR, Grayson ML, Hansen P, Singh N, Theuretzbacher U, Magrini N, WHO Pathogens Priority List Working Group. 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis 18:318-327.

19. O'Neill J. 2016. Tackling drug-resistant infections globally: final report and recommendations. The Review on Antimicrobial Resistance, United Kingdom.
20. Theuretzbacher U, Outterson K, Engel A, Karlen A. 2020. The global preclinical antibacterial pipeline. Nat Rev Microbiol 18:275-285.
21. de Kraker ME, Stewardson AJ, Harbarth S. 2016. Will 10 million people die a year due to antimicrobial resistance by 2050? PLoS Med 13:e1002184.
22. The World Bank. 2017. Drug-resistant infections: a threat to our economic future. The World Bank, Washington, DC.
23. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect 18:268-281.
24. Del Barrio-Tofino E, Zamorano L, Cortes-Lara S, Lopez-Causape C, Sanchez-Diener I, Cabot G, Bou G, Martinez-Martinez L, Oliver A, Group G-SRPs. 2019. Spanish nationwide survey on *Pseudomonas aeruginosa* antimicrobial resistance mechanisms and epidemiology. J Antimicrob Chemother 74:1825-1835.
25. Recio R, Mancheno M, Viedma E, Villa J, Orellana MA, Lora-Tamayo J, Chaves F. 2020. Predictors of mortality in bloodstream infections caused by *Pseudomonas aeruginosa* and impact of antimicrobial resistance and bacterial virulence. Antimicrob Agents Chemother 64:e01759-01719.
26. Hanquet G, Krizova P, Valentiner-Branth P, Ladhani SN, Nuorti JP, Lepoutre A, Mereckiene J, Knol M, Winje BA, Ciruela P, Ordobas M, Guevara M, McDonald E, Morfeldt E, Kozakova J, Slotved HC, Fry NK, Rinta-Kokko H, Varon E, Corcoran M, van der Ende A, Vestrheim DF, Munoz-Almagro C, Latasa P, Castilla J, Smith A, Henriques-Normark B, Whittaker R, Pastore Celentano L, Savulescu C, Sp IIMPG. 2019. Effect of childhood pneumococcal conjugate vaccination on invasive disease in older adults of 10 European countries: implications for adult vaccination. Thorax 74:473-482.
27. Sempere J, de Miguel S, Gonzalez-Camacho F, Yuste J, Domenech M. 2020. Clinical relevance and molecular pathogenesis of the emerging serotypes 22F and 33F of *Streptococcus pneumoniae* in Spain. Front Microbiol 11:309.
28. Linares J, Ardanuy C, Pallares R, Fenoll A. 2010. Changes in antimicrobial resistance, serotypes and genotypes in *Streptococcus pneumoniae* over a 30-year period. Clin Microbiol Infect 16:402-410.
29. Aguinagalde L, Corsini B, Domenech A, Domenech M, Camara J, Ardanuy C, Garcia E, Linares J, Fenoll A, Yuste J. 2015. Emergence of amoxicillin-resistant variants of Spain9V-ST156 pneumococci expressing serotype 11A correlates with their ability to evade the host immune response. PLoS One 10:e0137565.
30. Azarian T, Mitchell PK, Georgieva M, Thompson CM, Ghouila A, Pollard AJ, von Gottberg A, du Plessis M, Antonio M, Kwambana-Adams BA, Clarke SC, Everett D, Cornick J, Sadowy E, Hryniewicz W, Skoczynska A, Moisi JC, McGee L, Beall B, Metcalf BJ, Breiman RF, Ho PL, Reid R, O'Brien KL, Gladstone RA, Bentley SD, Hanage WP. 2018. Global emergence and population dynamics of divergent serotype 3 CC180 pneumococci. PLoS Pathog 14:e1007438.
31. Dahdouh E, Gomez-Gil R, Pacho S, Mingorance J, Daoud Z, Suarez M. 2017. Clonality, virulence determinants, and profiles of resistance of clinical *Acinetobacter baumannii* isolates obtained from a Spanish hospital. PLoS One 12:e0176824.
32. Ruiz J, Nunez ML, Perez J, Simarro E, Martinez-Campos L, Gomez J. 1999. Evolution of resistance among clinical isolates of *Acinetobacter* over a 6-year period. Eur J Clin Microbiol Infect Dis 18:292-295.

33. Goossens H, Ferech M, Vander Stichele R, Elseviers M, Group EP. 2005. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. Lancet 365:579-587.
34. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. 2010. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. BMJ 340:c2096.
35. Hutchings MI, Truman AW, Wilkinson B. 2019. Antibiotics: past, present and future. Curr Opin Microbiol 51:72-80.
36. Dyar OJ, Huttner B, Schouten J, Pulcini C, Esgap. 2017. What is antimicrobial stewardship? Clin Microbiol Infect 23:793-798.
37. Schweitzer VA, van Werkhoven CH, Rodriguez Bano J, Bielicki J, Harbarth S, Hulscher M, Huttner B, Islam J, Little P, Pulcini C, Savoldi A, Tacconelli E, Timsit JF, van Smeden M, Wolkewitz M, Bonten MJM, Walker AS, Llewelyn MJ, Joint Programming Initiative on Antimicrobial Resistance Working Group on Design of Antimicrobial Stewardship E. 2020. Optimizing design of research to evaluate antibiotic stewardship interventions: consensus recommendations of a multinational working group. Clin Microbiol Infect 26:41-50.
38. Pokharel S, Raut S, Adhikari B. 2019. Tackling antimicrobial resistance in low-income and middle-income countries. BMJ Glob Health 4:e002104.
39. Anomaly J. 2020. The Future of Phage: Ethical Challenges of Using Phage Therapy to Treat Bacterial Infections. Public Health Ethics 13:82-88.
40. Abedon ST, Thomas-Abedon C, Thomas A, Mazure H. 2011. Bacteriophage prehistory: Is or is not Hankin, 1896, a phage reference? Bacteriophage 1:174-178.
41. Bardell D. 1982. An 1898 report by Gamaleya for a lytic agent specific for *Bacillus anthracis*. J Hist Med Allied Sci 37:222-225.
42. Salmond GP, Fineran PC. 2015. A century of the phage: past, present and future. Nat Rev Microbiol 13:777-786.
43. Twort FW. 1915. An investigation on the nature of ultra-microscopic viruses. Lancet ii:1241-1243.
44. d'Herelle FH. 1917. Sur un microbe invisible antagoniste des bacilles dysenteriques. C R Acad Sci 165:373-375.
45. d'Herelle F. 1931. Bacteriophage as a treatment in acute medical and surgical infections. Bull N Y Acad Med 7:329-348.
46. Sulakvelidze A, Alavidze Z, Morris JG, Jr. 2001. Bacteriophage therapy. Antimicrob Agents Chemother 45:649-659.
47. Fleming A. 1929. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. Br J Exp Pathol 10:226-236.
48. Gordillo Altamirano FL, Barr JJ. 2019. Phage therapy in the postantibiotic era. Clin Microbiol Rev 32:e00066-00018.
49. Slopek S, Weber-Dabrowska B, Dabrowski M, Kucharewicz-Krukowska A. 1987. Results of bacteriophage treatment of suppurative bacterial infections in the years 1981-1986. Arch Immunol Ther Exp (Warsz) 35:569-583.
50. Wittig HJ, Raffetto JF, Bason R. 1966. Bacteriophage therapy in infective childhood asthma. JAMA 196:435.
51. Rhoads DD, Wolcott RD, Kuskowski MA, Wolcott BM, Ward LS, Sulakvelidze A. 2009. Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial. J Wound Care 18:237-238, 240-233.
52. Jault P, Leclerc T, Jennes S, Pirnay JP, Que YA, Resch G, Rousseau AF, Ravat F, Carsin H, Le Floch R, Schaal JV, Soler C, Fevre C, Arnaud I, Bretaudeau L, Gabard J. 2019. Efficacy and tolerability of a cocktail of bacteriophages to treat burn wounds infected by *Pseudomonas aeruginosa* (PhagoBurn): a randomised, controlled, double-blind phase 1/2 trial. Lancet Infect Dis 19:35-45.

53. Sarker SA, Berger B, Deng Y, Kieser S, Foata F, Moine D, Descombes P, Sultana S, Huq S, Bardhan PK, Vuillet V, Praplan F, Brussow H. 2017. Oral application of *Escherichia coli* bacteriophage: safety tests in healthy and diarrheal children from Bangladesh. Environ Microbiol 19:237-250.
54. Pirnay JP, Kutter E. 2021. Bacteriophages: it's a medicine, Jim, but not as we know it. Lancet Infect Dis 21:309-311.
55. Gindin M, Febvre HP, Rao S, Wallace TC, Weir TL. 2019. Bacteriophage for gastrointestinal health (PHAGE) study: evaluating the safety and tolerability of supplemental bacteriophage consumption. J Am Coll Nutr 38:68-75.
56. McCallin S, Sacher JC, Zheng J, Chan BK. 2019. Current state of compassionate phage therapy. Viruses 11.
57. Gibson SB, Green SI, Liu CG, Salazar KC, Clark JR, Terwilliger AL, Kaplan HB, Maresso AW, Trautner BW, Ramig RF. 2019. Constructing and characterizing bacteriophage libraries for phage therapy of human infections. Front Microbiol 10:2537.
58. Schooley RT, Biswas B, Gill JJ, Hernandez-Morales A, Lancaster J, Lessor L, Barr JJ, Reed SL, Rohwer F, Benler S, Segall AM, Taplitz R, Smith DM, Kerr K, Kumaraswamy M, Nizet V, Lin L, McCauley MD, Strathdee SA, Benson CA, Pope RK, Leroux BM, Picel AC, Mateczun AJ, Cilwa KE, Regeimbal JM, Estrella LA, Wolfe DM, Henry MS, Quinones J, Salka S, Bishop-Lilly KA, Young R, Hamilton T. 2017. Development and use of personalized bacteriophage-based therapeutic cocktails To treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. Antimicrob Agents Chemother 61.
59. Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, Gilmour KC, Soothill J, Jacobs-Sera D, Schooley RT, Hatfull GF, Spencer H. 2019. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. Nat Med 25:730-733.
60. Pirnay JP. 2020. Phage Therapy in the Year 2035. Front Microbiol 11:1171.
61. Pirnay JP, Verbeken G, Ceyssens PJ, Huys I, De Vos D, Ameloot C, Fauconnier A. 2018. The Magistral Phage. Viruses 10.
62. Redrejo-Rodriguez M, Garcia P. 2018. "FAGOMA: Spanish Network of Bacteriophages and Transducer Elements"-V meeting report. Viruses 10.
63. Schleifer KH, Kandler O. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev 36:407-477.
64. Wang IN, Smith DL, Young R. 2000. Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol 54:799-825.
65. Fisher JF, Mobashery S. 2020. Constructing and deconstructing the bacterial cell wall. Protein Sci 29:629-646.
66. Dams D, Briers Y. 2019. Enzybiotics: enzyme-based antibacterials as therapeutics. Adv Exp Med Biol 1148:233-253.
67. Garcia E. 1996. On the misuse of the term 'lysozyme'. Mol Microbiol 21:885.
68. Young R. 2014. Phage lysis: three steps, three choices, one outcome. J Microbiol 52:243-258.
69. Park T, Struck DK, Dankenbring CA, Young R. 2007. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. J Bacteriol 189:9135-9139.
70. Kongari R, Rajaure M, Cahill J, Rasche E, Mijalis E, Berry J, Young R. 2018. Phage spanins: diversity, topological dynamics and gene convergence. BMC Bioinformatics 19:326.
71. Berry J, Rajaure M, Pang T, Young R. 2012. The spanin complex is essential for lambda lysis. J Bacteriol 194:5667-5674.
72. Sutcliffe IC. 1998. Cell envelope composition and organisation in the genus *Rhodococcus*. Antonie Van Leeuwenhoek 74:49-58.
73. Payne KM, Hatfull GF. 2012. Mycobacteriophage endolysins: diverse and modular enzymes with multiple catalytic activities. PLoS One 7:e34052.

74. Fleming A. 1922. On a remarkable bacteriolytic element found in tissues and secretions. Proc R Soc B: Biol Sci 93:306-317.
75. Nelson D, Loomis L, Fischetti VA. 2001. Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. Proc Natl Acad Sci U S A 98:4107-4112.
76. Loeffler JM, Nelson D, Fischetti VA. 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science 294:2170-2172.
77. Fischetti VA. 2018. Development of phage lysins as novel therapeutics: a historical perspective. Viruses 10.
78. Pastagia M, Schuch R, Fischetti VA, Huang DB. 2013. Lysins: the arrival of pathogen-directed anti-infectives. J Med Microbiol 62:1506-1516.
79. Cheng M, Zhang Y, Li X, Liang J, Hu L, Gong P, Zhang L, Cai R, Zhang H, Ge J, Ji Y, Guo Z, Feng X, Sun C, Yang Y, Lei L, Han W, Gu J. 2017. Endolysin LysEF-P10 shows potential as an alternative treatment strategy for multidrug-resistant *Enterococcus faecalis* infections. Sci Rep 7:10164.
80. Diez-Martinez R, De Paz HD, Garcia-Fernandez E, Bustamante N, Euler CW, Fischetti VA, Menendez M, Garcia P. 2015. A novel chimeric phage lysin with high *in vitro* and *in vivo* bactericidal activity against *Streptococcus pneumoniae*. J Antimicrob Chemother 70:1763-1773.
81. Kusuma C, Jadanova A, Chanturiya T, Kokai-Kun JF. 2007. Lysostaphin-resistant variants of *Staphylococcus aureus* demonstrate reduced fitness *in vitro* and *in vivo*. Antimicrob Agents Chemother 51:475-482.
82. Gilmer DB, Schmitz JE, Euler CW, Fischetti VA. 2013. Novel bacteriophage lysin with broad lytic activity protects against mixed infection by *Streptococcus pyogenes* and methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 57:2743-2750.
83. Vázquez R GE, García P. 2018. Phage lysins for fighting bacterial respiratory infections: a new generation of antimicrobials. Front Immunol 9:2252.
84. Grishin AV, Karyagina AS, Vasina DV, Vasina IV, Gushchin VA, Lunin VG. 2020. Resistance to peptidoglycan-degrading enzymes. Crit Rev Microbiol 46:703-726.
85. Letrado P, Corsini B, Diez-Martinez R, Bustamante N, Yuste JE, Garcia P. 2018. Bactericidal synergism between antibiotics and phage endolysin Cpl-711 to kill multidrug-resistant pneumococcus. Future Microbiol 13:1215-1223.
86. Vázquez R, García P. 2019. Synergy between two chimeric lysins to kill *Streptococcus pneumoniae*. Front Microbiol 10:1251.
87. Jun SY, Jang IJ, Yoon S, Jang K, Yu KS, Cho JY, Seong MW, Jung GM, Yoon SJ, Kang SH. 2017. Pharmacokinetics and tolerance of the phage endolysin-based candidate drug SAL200 after a single intravenous administration among healthy volunteers. Antimicrob Agents Chemother 61:e02629-02616.
88. Loeffler JM, Fischetti VA. 2003. Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant *Streptococcus pneumoniae* strains. Antimicrob Agents Chemother 47:375-377.
89. Jado I, Lopez R, Garcia E, Fenoll A, Casal J, Garcia P, Spanish Pneumococcal Infection Study N. 2003. Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model. J Antimicrob Chemother 52:967-973.
90. Domenech M, Garcia E, Moscoso M. 2011. *In vitro* destruction of *Streptococcus pneumoniae* biofilms with bacterial and phage peptidoglycan hydrolases. Antimicrob Agents Chemother 55:4144-4148.
91. Loeffler JM, Djurkovic S, Fischetti VA. 2003. Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. Infect Immun 71:6199-6204.
92. Djurkovic S, Loeffler JM, Fischetti VA. 2005. Synergistic killing of *Streptococcus pneumoniae* with the bacteriophage lytic enzyme Cpl-1 and penicillin or gentamicin depends on the level of penicillin resistance. Antimicrob Agents Chemother 49:1225-1228.

93. McCullers JA, Karlstrom A, Iverson AR, Loeffler JM, Fischetti VA. 2007. Novel strategy to prevent otitis media caused by colonizing *Streptococcus pneumoniae*. PLoS Pathog 3:e28.
94. Witzenrath M, Schmeck B, Doehn JM, Tschernig T, Zahlten J, Loeffler JM, Zemlin M, Muller H, Gutbier B, Schutte H, Hippenstiel S, Fischetti VA, Suttorp N, Rosseau S. 2009. Systemic use of the endolysin Cpl-1 rescues mice with fatal pneumococcal pneumonia. Crit Care Med 37:642-649.
95. Vouillamoz J, Entenza JM, Giddey M, Fischetti VA, Moreillon P, Resch G. 2013. Bactericidal synergism between daptomycin and the phage lysin Cpl-1 in a mouse model of pneumococcal bacteraemia. Int J Antimicrob Agents 42:416-421.
96. Doehn JM, Fischer K, Reppe K, Gutbier B, Tschernig T, Hocke AC, Fischetti VA, Loffler J, Suttorp N, Hippenstiel S, Witzenrath M. 2013. Delivery of the endolysin Cpl-1 by inhalation rescues mice with fatal pneumococcal pneumonia. J Antimicrob Chemother 68:2111-2117.
97. Corsini B, Diez-Martinez R, Aguinagalde L, Gonzalez-Camacho F, Garcia-Fernandez E, Letrado P, Garcia P, Yuste J. 2018. Chemotherapy with phage lysins reduces pneumococcal colonization of the respiratory tract. Antimicrob Agents Chemother 62.
98. Rodriguez-Cerrato V, Garcia P, Huelves L, Garcia E, Del Prado G, Gracia M, Ponte C, Lopez R, Soriano F. 2007. Pneumococcal LytA autolysin, a potent therapeutic agent in experimental peritonitis-sepsis caused by highly beta-lactam-resistant *Streptococcus pneumoniae*. Antimicrob Agents Chemother 51:3371-3373.
99. Diez-Martinez R, de Paz HD, Bustamante N, Garcia E, Menendez M, Garcia P. 2013. Improving the lethal effect of cpl-7, a pneumococcal phage lysozyme with broad bactericidal activity, by inverting the net charge of its cell wall-binding module. Antimicrob Agents Chemother 57:5355-5365.
100. Blázquez B, Fresco-Taboada A, Iglesias-Bexiga M, Menéndez M, García P. 2016. PL3 amidase, a tailor-made lysin constructed by domain shuffling with potent killing activity against pneumococci and related species. Front Microbiol 7:1156.
101. Nelson D, Schuch R, Chahales P, Zhu S, Fischetti VA. 2006. PlyC: a multimeric bacteriophage lysin. Proc Natl Acad Sci U S A 103:10765-10770.
102. Shen Y, Barros M, Vennemann T, Gallagher DT, Yin Y, Linden SB, Heselpoth RD, Spencer DJ, Donovan DM, Moult J, Fischetti VA, Heinrich F, Losche M, Nelson DC. 2016. A bacteriophage endolysin that eliminates intracellular streptococci. Elife 5.
103. Lood R, Raz A, Molina H, Euler CW, Fischetti VA. 2014. A highly active and negatively charged *Streptococcus pyogenes* lysin with a rare D-alanyl-L-alanine endopeptidase activity protects mice against streptococcal bacteremia. Antimicrob Agents Chemother 58:3073-3084.
104. Cheng Q, Nelson D, Zhu S, Fischetti VA. 2005. Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. Antimicrob Agents Chemother 49:111-117.
105. Cheng Q, Fischetti VA. 2007. Mutagenesis of a bacteriophage lytic enzyme PlyGBS significantly increases its antibacterial activity against group B streptococci. Appl Microbiol Biotechnol 74:1284-1291.
106. Polak J, Della Latta P, Blackburn P. 1993. *In vitro* activity of recombinant lysostaphin-antibiotic combinations toward methicillin-resistant *Staphylococcus aureus*. Diagn Microbiol Infect Dis 17:265-270.
107. Climo MW, Ehlert K, Archer GL. 2001. Mechanism and suppression of lysostaphin resistance in oxacillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 45:1431-1437.

108. Kiri N, Archer G, Climo MW. 2002. Combinations of lysostaphin with beta-lactams are synergistic against oxacillin-resistant *Staphylococcus epidermidis*. Antimicrob Agents Chemother 46:2017-2020.
109. Wu JA, Kusuma C, Mond JJ, Kokai-Kun JF. 2003. Lysostaphin disrupts *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms on artificial surfaces. Antimicrob Agents Chemother 47:3407-3414.
110. Kokai-Kun JF, Walsh SM, Chanturiya T, Mond JJ. 2003. Lysostaphin cream eradicates *Staphylococcus aureus* nasal colonization in a cotton rat model. Antimicrob Agents Chemother 47:1589-1597.
111. Walsh S, Shah A, Mond J. 2003. Improved pharmacokinetics and reduced antibody reactivity of lysostaphin conjugated to polyethylene glycol. Antimicrob Agents Chemother 47:554-558.
112. Walsh S, Kokai-Kun J, Shah A, Mond J. 2004. Extended nasal residence time of lysostaphin and an anti-staphylococcal monoclonal antibody by delivery in semisolid or polymeric carriers. Pharm Res 21:1770-1775.
113. Kokai-Kun JF, Chanturiya T, Mond JJ. 2007. Lysostaphin as a treatment for systemic *Staphylococcus aureus* infection in a mouse model. J Antimicrob Chemother 60:1051-1059.
114. Becker SC, Foster-Frey J, Donovan DM. 2008. The phage K lytic enzyme LysK and lysostaphin act synergistically to kill MRSA. FEMS Microbiol Lett 287:185-191.
115. Aguinaga A, Francés ML, Del Pozo JL, Alonso M, Serrera A, Lasa I, Leiva J. 2011. Lysostaphin and clarithromycin: a promising combination for the eradication of *Staphylococcus aureus* biofilms. Int J Antimicrob Agents 37:585-587.
116. Schmelcher M, Shen Y, Nelson DC, Eugster MR, Eichenseher F, Hanke DC, Loessner MJ, Dong S, Pritchard DG, Lee JC, Becker SC, Foster-Frey J, Donovan DM. 2015. Evolutionarily distinct bacteriophage endolysins featuring conserved peptidoglycan cleavage sites protect mice from MRSA infection. J Antimicrob Chemother 70:1453-1465.
117. Hathaway H, Ajuebor J, Stephens L, Coffey A, Potter U, Sutton JM, Jenkins AT. 2017. Thermally triggered release of the bacteriophage endolysin CHAPK and the bacteriocin lysostaphin for the control of methicillin resistant *Staphylococcus aureus* (MRSA). J Control Release 245:108-115.
118. O'Flaherty S, Coffey A, Meaney W, Fitzgerald GF, Ross RP. 2005. The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant *Staphylococcus aureus*. J Bacteriol 187:7161-7164.
119. Filatova LY, Donovan DM, Becker SC, Lebedev DN, Priyma AD, Koudriachova HV, Kabanov AV, Klyachko NL. 2013. Physicochemical characterization of the staphylolytic LysK enzyme in complexes with polycationic polymers as a potent antimicrobial. Biochimie 95:1689-1696.
120. Horgan M, O'Flynn G, Garry J, Cooney J, Coffey A, Fitzgerald GF, Ross RP, McAuliffe O. 2009. Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci. Appl Environ Microbiol 75:872-874.
121. Fenton M, Casey PG, Hill C, Gahan CG, Ross RP, McAuliffe O, O'Mahony J, Maher F, Coffey A. 2010. The truncated phage lysin $CHAP_K$ eliminates *Staphylococcus aureus* in the nares of mice. Bioeng Bugs 1:404-407.
122. Fenton M, Ross RP, McAuliffe O, O'Mahony J, Coffey A. 2011. Characterization of the staphylococcal bacteriophage lysin $CHAP_K$. J Appl Microbiol 111:1025-1035.
123. Fenton M, Keary R, McAuliffe O, Ross RP, O'Mahony J, Coffey A. 2013. Bacteriophage-derived peptidase $CHAP_K$ eliminates and prevents staphylococcal biofilms. Int J Microbiol 2013:625341.

124. Daniel A, Euler C, Collin M, Chahales P, Gorelick KJ, Fischetti VA. 2010. Synergism between a novel chimeric lysin and oxacillin protects against infection by methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 54:1603-1612.
125. Jun SY, Jung GM, Son JS, Yoon SJ, Choi YJ, Kang SH. 2011. Comparison of the antibacterial properties of phage endolysins SAL-1 and LysK. Antimicrob Agents Chemother 55:1764-1767.
126. Jun SY, Jung GM, Yoon SJ, Oh MD, Choi YJ, Lee WJ, Kong JC, Seol JG, Kang SH. 2013. Antibacterial properties of a pre-formulated recombinant phage endolysin, SAL-1. Int J Antimicrob Agents 41:156-161.
127. Jun SY, Jung GM, Yoon SJ, Choi YJ, Koh WS, Moon KS, Kang SH. 2014. Preclinical safety evaluation of intravenously administered SAL200 containing the recombinant phage endolysin SAL-1 as a pharmaceutical ingredient. Antimicrob Agents Chemother 58:2084-2088.
128. Jun SY, Jung GM, Yoon SJ, Youm SY, Han HY, Lee JH, Kang SH. 2016. Pharmacokinetics of the phage endolysin-based candidate drug SAL200 in monkeys and its appropriate intravenous dosing period. Clin Exp Pharmacol Physiol 43:1013-1016.
129. Paul VD, Rajagopalan SS, Sundarrajan S, George SE, Asrani JY, Pillai R, Chikkamadaiah R, Durgaiah M, Sriram B, Padmanabhan S. 2011. A novel bacteriophage Tail-Associated Muralytic Enzyme (TAME) from Phage K and its development into a potent antistaphylococcal protein. BMC Microbiol 11:226.
130. Vipra AA, Desai SN, Roy P, Patil R, Raj JM, Narasimhaswamy N, Paul VD, Chikkamadaiah R, Sriram B. 2012. Antistaphylococcal activity of bacteriophage derived chimeric protein P128. BMC Microbiol 12:41.
131. George SE, Chikkamadaiah R, Durgaiah M, Joshi AA, Thankappan UP, Madhusudhana SN, Sriram B. 2012. Biochemical characterization and evaluation of cytotoxicity of antistaphylococcal chimeric protein P128. BMC Res Notes 5:280.
132. Drilling AJ, Cooksley C, Chan C, Wormald PJ, Vreugde S. 2016. Fighting sinus-derived *Staphylococcus aureus* biofilms *in vitro* with a bacteriophage-derived muralytic enzyme. Int Forum Allergy Rhinol 6:349-355.
133. Poonacha N, Nair S, Desai S, Tuppad D, Hiremath D, Mohan T, Vipra A, Sharma U. 2017. Efficient killing of planktonic and biofilm-embedded coagulase-negative staphylococci by bactericidal protein P128. Antimicrob Agents Chemother 61.
134. Nair S, Desai S, Poonacha N, Vipra A, Sharma U. 2016. Antibiofilm activity and synergistic inhibition of staphylococcus aureus biofilms by bactericidal protein P128 in combination with antibiotics. Antimicrob Agents Chemother 60:7280-7289.
135. Nair S, Poonacha N, Desai S, Hiremath D, Tuppad D, Mohan T, Chikkamadaiah R, Durgaiah M, Kumar S, Channabasappa S, Vipra A, Sharma U. 2018. Restoration of sensitivity of a diverse set of drug-resistant Staphylococcus clinical strains by bactericidal protein P128. J Med Microbiol 67:296-307.
136. Channabasappa S, Durgaiah M, Chikkamadaiah R, Kumar S, Joshi A, Sriram B. 2018. Efficacy of Novel Antistaphylococcal Ectolysin P128 in a Rat Model of Methicillin-Resistant Staphylococcus aureus Bacteremia. Antimicrob Agents Chemother 62.
137. Gu J, Zuo J, Lei L, Zhao H, Sun C, Feng X, Du C, Li X, Yang Y, Han W. 2011. LysGH15 reduces the inflammation caused by lethal methicillin-resistant *Staphylococcus aureus* infection in mice. Bioeng Bugs 2:96-99.
138. Gu J, Xu W, Lei L, Huang J, Feng X, Sun C, Du C, Zuo J, Li Y, Du T, Li L, Han W. 2011. LysGH15, a novel bacteriophage lysin, protects a murine bacteremia model efficiently against lethal methicillin-resistant *Staphylococcus aureus* infection. J Clin Microbiol 49:111-117.

139. Xia F, Li X, Wang B, Gong P, Xiao F, Yang M, Zhang L, Song J, Hu L, Cheng M, Sun C, Feng X, Lei L, Ouyang S, Liu ZJ, Li X, Gu J, Han W. 2016. Combination therapy of LysGH15 and apigenin as a new strategy for treating pneumonia caused by *Staphylococcus aureus*. Appl Environ Microbiol 82:87-94.
140. Zhang L, Li D, Li X, Hu L, Cheng M, Xia F, Gong P, Wang B, Ge J, Zhang H, Cai R, Wang Y, Sun C, Feng X, Lei L, Han W, Gu J. 2016. LysGH15 kills *Staphylococcus aureus* without being affected by the humoral immune response or inducing inflammation. Sci Rep 6:29344.
141. Zhang Y, Cheng M, Zhang H, Dai J, Guo Z, Li X, Ji Y, Cai R, Xi H, Wang X, Xue Y, Sun C, Feng X, Lei L, Han W, Gu J. 2018. Antibacterial Effects of Phage Lysin LysGH15 on Planktonic Cells and Biofilms of Diverse Staphylococci. Appl Environ Microbiol 84:e00886-00818.
142. Schuch R, Lee HM, Schneider BC, Sauve KL, Law C, Khan BK, Rotolo JA, Horiuchi Y, Couto DE, Raz A, Fischetti VA, Huang DB, Nowinski RC, Wittekind M. 2014. Combination therapy with lysin CF-301 and antibiotic is superior to antibiotic alone for treating methicillin-resistant *Staphylococcus aureus*-induced murine bacteremia. J Infect Dis 209:1469-1478.
143. Schuch R, Khan BK, Raz A, Rotolo JA, Wittekind M. 2017. Bacteriophage lysin CF-301, a potent antistaphylococcal biofilm agent. Antimicrob Agents Chemother 61.
144. Yang H, Zhang H, Wang J, Yu J, Wei H. 2017. A novel chimeric lysin with robust antibacterial activity against planktonic and biofilm methicillin-resistant *Staphylococcus aureus*. Sci Rep 7:40182.
145. Grover N, Paskaleva EE, Mehta KK, Dordick JS, Kane RS. 2014. Growth inhibition of *Mycobacterium smegmatis* by mycobacteriophage-derived enzymes. Enzyme Microb Technol 63:1-6.
146. Gil F, Grzegorzewicz AE, Catalao MJ, Vital J, McNeil MR, Pimentel M. 2010. Mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. Microbiology (Reading) 156:1497-1504.
147. Lai MJ, Liu CC, Jiang SJ, Soo PC, Tu MH, Lee JJ, Chen YH, Chang KC. 2015. Antimycobacterial Activities of Endolysins Derived From a Mycobacteriophage, BTCU-1. Molecules 20:19277-19290.
148. Diaz E, Lopez R, Garcia JL. 1991. Chimeric pneumococcal cell wall lytic enzymes reveal important physiological and evolutionary traits. J Biol Chem 266:5464-5471.
149. De Maesschalck V, Gutierrez D, Paeshuyse J, Lavigne R, Briers Y. 2020. Advanced engineering of third-generation lysins and formulation strategies for clinical applications. Crit Rev Microbiol 46:548-564.
150. Vázquez R, Domenech M, Iglesias-Bexiga M, Menéndez M, García P. 2017. Csl2, a novel chimeric bacteriophage lysin to fight infections caused by *Streptococcus suis*, an emerging zoonotic pathogen. Sci Rep 7:16506.
151. Maestro B, Sanz JM. 2016. Choline binding proteins from *Streptococcus pneumoniae*: a dual role as enzybiotics and targets for the design of new antimicrobials. Antibiotics 5:21.
152. Bustamante N, Iglesias-Bexiga M, Bernardo-Garcia N, Silva-Martin N, Garcia G, Campanero-Rhodes MA, Garcia E, Uson I, Buey RM, Garcia P, Hermoso JA, Bruix M, Menendez M. 2017. Deciphering how Cpl-7 cell wall-binding repeats recognize the bacterial peptidoglycan. Sci Rep 7:16494.
153. Low LY, Yang C, Perego M, Osterman A, Liddington R. 2011. Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins. J Biol Chem 286:34391-34403.
154. Gerstmans H, Grimon D, Gutierrez D, Lood C, Rodriguez A, van Noort V, Lammertyn J, Lavigne R, Briers Y. 2020. A VersaTile-driven platform for rapid hit-to-lead development of engineered lysins. Sci Adv 6:eaaz1136.

155. Duyvejonck L, Gerstmans, H., Stock, M., Grimon, D., Lavigne, R., Briers, Y. 2021. Rapid and high-throughput evaluation of diverse configurations of engineered lysins using the VersaTile technique. Antibiotics 10:293.
156. Briers Y, Lavigne R. 2015. Breaking barriers: expansion of the use of endolysins as novel antibacterials against Gram-negative bacteria. Future Microbiol 10:377-390.
157. Briers Y, Walmagh M, Lavigne R. 2011. Use of bacteriophage endolysin EL188 and outer membrane permeabilizers against *Pseudomonas aeruginosa*. J Appl Microbiol 110:778-785.
158. Oliveira H, Thiagarajan V, Walmagh M, Sillankorva S, Lavigne R, Neves-Petersen MT, Kluskens LD, Azeredo J. 2014. A thermostable *Salmonella* phage endolysin, Lys68, with broad bactericidal properties against gram-negative pathogens in presence of weak acids. PLoS One 9:e108376.
159. Briers Y, Walmagh M, Van Puyenbroeck V, Cornelissen A, Cenens W, Aertsen A, Oliveira H, Azeredo J, Verween G, Pirnay JP, Miller S, Volckaert G, Lavigne R. 2014. Engineered endolysin-based "Artilysins" to combat multidrug-resistant Gram-negative pathogens. mBio 5:e01379-01314.
160. Heselpoth RD, Euler CW, Schuch R, Fischetti VA. 2019. Lysocins: bioengineered antimicrobials that deliver lysins across the outer membrane of Gram-negative bacteria. Antimicrob Agents Chemother 63:e00342-00319.
161. Zampara A, Sorensen MCH, Grimon D, Antenucci F, Vitt AR, Bortolaia V, Briers Y, Brondsted L. 2020. Exploiting phage receptor binding proteins to enable endolysins to kill Gram-negative bacteria. Sci Rep 10:12087.
162. Morita M, Tanji Y, Mizoguchi K, Soejima A, Orito Y, Unno H. 2001. Antibacterial activity of *Bacillus amyloliquefaciens* phage endolysin without holin conjugation. J Biosci Bioeng 91:469-473.
163. Morita M, Tanji Y, Orito Y, Mizoguchi K, Soejima A, Unno H. 2001. Functional analysis of antibacterial activity of *Bacillus amyloliquefaciens* phage endolysin against Gram-negative bacteria. FEBS Lett 500:56-59.
164. Orito Y, Morita M, Hori K, Unno H, Tanji Y. 2004. *Bacillus amyloliquefaciens* phage endolysin can enhance permeability of *Pseudomonas aeruginosa* outer membrane and induce cell lysis. Appl Microbiol Biotechnol 65:105-109.
165. Ibrahim HR HS, Koketsu M, Juneja LR, Kim M, Yamamoto T, Sugimoto Y, Aoki T. 1996. Partially unfolded lysozyme at neutral pH agglutinates and kills Gram-negative and Gram-positive bacteria through membrane damage mechanism. J Agric Food Chem 44:3799-3806.
166. During K, Porsch P, Mahn A, Brinkmann O, Gieffers W. 1999. The non-enzymatic microbicidal activity of lysozymes. FEBS Lett 449:93-100.
167. Rotem S, Radzishevsky I, Inouye RT, Samore M, Mor A. 2006. Identification of antimicrobial peptide regions derived from genomic sequences of phage lysins. Peptides 27:18-26.
168. Lim JA, Shin H, Heu S, Ryu S. 2014. Exogenous lytic activity of SPN9CC endolysin against gram-negative bacteria. J Microbiol Biotechnol 24:803-811.
169. Thandar M, Lood R, Winer BY, Deutsch DR, Euler CW, Fischetti VA. 2016. Novel engineered peptides of a phage lysin as effective antimicrobials against multidrug-resistant *Acinetobacter baumannii*. Antimicrob Agents Chemother 60:2671-2679.
170. Guo M, Feng C, Ren J, Zhuang X, Zhang Y, Zhu Y, Dong K, He P, Guo X, Qin J. 2017. A novel antimicrobial endolysin, LysPA26, against *Pseudomonas aeruginosa*. Front Microbiol 8:293.
171. Peng SY, You RI, Lai MJ, Lin NT, Chen LK, Chang KC. 2017. Highly potent antimicrobial modified peptides derived from the *Acinetobacter baumannii* phage endolysin LysAB2. Sci Rep 7:11477.
172. Ghose C, Euler CW. 2020. Gram-Negative Bacterial Lysins. Antibiotics 9:74.

173. Maciejewska B, Zrubek K, Espaillat A, Wisniewska M, Rembacz KP, Cava F, Dubin G, Drulis-Kawa Z. 2017. Modular endolysin of *Burkholderia* AP3 phage has the largest lysozyme-like catalytic subunit discovered to date and no catalytic aspartate residue. Sci Rep 7:14501.
174. Sykilinda NN, Nikolaeva AY, Shneider MM, Mishkin DV, Patutin AA, Popov VO, Boyko KM, Klyachko NL, Miroshnikov KA. 2018. Structure of an *Acinetobacter* broad-range prophage endolysin reveals a C-terminal alpha-helix with the proposed role in activity against live bacterial cells. Viruses 10.
175. Le CF, Fang CM, Sekaran SD. 2017. Intracellular targeting mechanisms by antimicrobial peptides. Antimicrob Agents Chemother 61:e02340-02316.
176. Magana M, Pushpanathan M, Santos AL, Leanse L, Fernandez M, Ioannidis A, Giulianotti MA, Apidianakis Y, Bradfute S, Ferguson AL, Cherkasov A, Seleem MN, Pinilla C, de la Fuente-Nunez C, Lazaridis T, Dai T, Houghten RA, Hancock REW, Tegos GP. 2020. The value of antimicrobial peptides in the age of resistance. Lancet Infect Dis 20:e216-e230.
177. Briers Y, Volckaert G, Cornelissen A, Lagaert S, Michiels CW, Hertveldt K, Lavigne R. 2007. Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages phiKZ and EL. Mol Microbiol 65:1334-1344.
178. Briers Y, Cornelissen A, Aertsen A, Hertveldt K, Michiels CW, Volckaert G, Lavigne R. 2008. Analysis of outer membrane permeability of *Pseudomonas aeruginosa* and bactericidal activity of endolysins KZ144 and EL188 under high hydrostatic pressure. FEMS Microbiol Lett 280:113-119.
179. Walmagh M, Briers Y, dos Santos SB, Azeredo J, Lavigne R. 2012. Characterization of modular bacteriophage endolysins from Myoviridae phages OBP, 201phi2-1 and PVP-SE1. PLoS One 7:e36991.
180. Briers Y, Walmagh M, Grymonprez B, Biebl M, Pirnay JP, Defraine V, Michiels J, Cenens W, Aertsen A, Miller S, Lavigne R. 2014. Art-175 is a highly efficient antibacterial against multidrug-resistant strains and persisters of *Pseudomonas aeruginosa*. Antimicrob Agents Chemother 58:3774-3784.
181. Defraine V, Schuermans J, Grymonprez B, Govers SK, Aertsen A, Fauvart M, Michiels J, Lavigne R, Briers Y. 2016. Efficacy of Artilysin Art-175 against Resistant and Persistent *Acinetobacter baumannii*. Antimicrob Agents Chemother 60:3480-3488.
182. Lai MJ, Lin NT, Hu A, Soo PC, Chen LK, Chen LH, Chang KC. 2011. Antibacterial activity of *Acinetobacter baumannii* phage varphiAB2 endolysin (LysAB2) against both gram-positive and gram-negative bacteria. Appl Microbiol Biotechnol 90:529-539.
183. Thummeepak R, Kitti T, Kunthalert D, Sitthisak S. 2016. Enhanced antibacterial activity of *Acinetobacter baumannii* bacteriophage ØABP-01 endolysin (LysABP-01) in combination with colistin. Front Microbiol 7:1402.
184. Huang G, Shen X, Gong Y, Dong Z, Zhao X, Shen W, Wang J, Hu F, Peng Y. 2014. Antibacterial properties of *Acinetobacter baumannii* phage Abp1 endolysin (PlyAB1). BMC Infect Dis 14:681.
185. Lood R, Winer BY, Pelzek AJ, Diez-Martinez R, Thandar M, Euler CW, Schuch R, Fischetti VA. 2015. Novel phage lysin capable of killing the multidrug-resistant gram-negative bacterium *Acinetobacter baumannii* in a mouse bacteremia model. Antimicrob Agents Chemother 59:1983-1991.
186. Lai MJ, Soo PC, Lin NT, Hu A, Chen YJ, Chen LK, Chang KC. 2013. Identification and characterisation of the putative phage-related endolysins through full genome sequence analysis in *Acinetobacter baumannii* ATCC 17978. Int J Antimicrob Agents 42:141-148.
187. Lv M, Wang S, Yan G, Sun C, Feng X, Gu J, Han W, Lei L. 2015. Genome sequencing and analysis of an *Escherichia coli* phage vB_EcoM-ep3 with a novel lysin, Lysep3. Virus Genes 50:487-497.

188. Wang S, Gu J, Lv M, Guo Z, Yan G, Yu L, Du C, Feng X, Han W, Sun C, Lei L. 2017. The antibacterial activity of *E. coli* bacteriophage lysin Lysep3 is enhanced by fusing the *Bacillus amyloliquefaciens* bacteriophage endolysin binding domain D8 to the C-terminal region. J Microbiol 55:403-408.
189. Yan G, Liu J, Ma Q, Zhu R, Guo Z, Gao C, Wang S, Yu L, Gu J, Hu D, Han W, Du R, Yang J, Lei L. 2017. The N-terminal and central domain of colicin A enables phage lysin to lyse *Escherichia coli* extracellularly. Antonie Van Leeuwenhoek 110:1627-1635.
190. Shavrina MS, Zimin AA, Molochkov NV, Chernyshov SV, Machulin AV, Mikoulinskaia GV. 2016. *In vitro* study of the antibacterial effect of the bacteriophage T5 thermostable endolysin on *Escherichia coli* cells. J Appl Microbiol 121:1282-1290.
191. Larpin Y, Oechslin F, Moreillon P, Resch G, Entenza JM, Mancini S. 2018. *In vitro* characterization of PlyE146, a novel phage lysin that targets Gram-negative bacteria. PLoS One 13:e0192507.
192. Walmagh M, Boczkowska B, Grymonprez B, Briers Y, Drulis-Kawa Z, Lavigne R. 2013. Characterization of five novel endolysins from Gram-negative infecting bacteriophages. Appl Microbiol Biotechnol 97:4369-4375.
193. Maciejewska B, Roszniowski B, Espaillat A, Kesik-Szeloch A, Majkowska-Skrobek G, Kropinski AM, Briers Y, Cava F, Lavigne R, Drulis-Kawa Z. 2017. *Klebsiella* phages representing a novel clade of viruses with an unknown DNA modification and biotechnologically interesting enzymes. Appl Microbiol Biotechnol 101:673-684.
194. Oliveira H, Pinto G, Oliveira A, Oliveira C, Faustino MA, Briers Y, Domingues L, Azeredo J. 2016. Characterization and genome sequencing of a *Citrobacter freundii* phage CfP1 harboring a lysin active against multidrug-resistant isolates. Appl Microbiol Biotechnol 100:10543-10553.
195. Dong H, Zhu C, Chen J, Ye X, Huang YP. 2015. Antibacterial activity of *Stenotrophomonas maltophilia* endolysin P28 against both Gram-positive and Gram-negative bacteria. Front Microbiol 6:1299.
196. Eisenberg D, Weiss RM, Terwilliger TC. 1982. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. Nature 299:371-374.
197. Lee EY, Wong GCL, Ferguson AL. 2018. Machine learning-enabled discovery and design of membrane-active peptides. Bioorg Med Chem 26:2708-2718.
198. Wang G. 2015. Improved methods for classification, prediction, and design of antimicrobial peptides. Methods Mol Biol 1268:43-66.
199. Porto WF, Irazazabal L, Alves ESF, Ribeiro SM, Matos CO, Pires AS, Fensterseifer ICM, Miranda VJ, Haney EF, Humblot V, Torres MDT, Hancock REW, Liao LM, Ladram A, Lu TK, de la Fuente-Nunez C, Franco OL. 2018. *In silico* optimization of a guava antimicrobial peptide enables combinatorial exploration for peptide design. Nat Commun 9:1490.
200. Montesinos E, Bardaji E. 2008. Synthetic antimicrobial peptides as agricultural pesticides for plant-disease control. Chem Biodivers 5:1225-1237.
201. Liu Q, Yao S, Chen Y, Gao S, Yang Y, Deng J, Ren Z, Shen L, Cui H, Hu Y, Ma X, Yu S. 2017. Use of antimicrobial peptides as a feed additive for juvenile goats. Sci Rep 7:12254.
202. Rai M, Pandit R, Gaikwad S, Kovics G. 2016. Antimicrobial peptides as natural bio-preservative to enhance the shelf-life of food. J Food Sci Technol 53:3381-3394.
203. Gaglione R, Pane K, Dell'Olmo E, Cafaro V, Pizzo E, Olivieri G, Notomista E, Arciello A. 2019. Cost-effective production of recombinant peptides in *Escherichia coli*. N Biotechnol 51:39-48.
204. Cao J, de la Fuente-Nunez C, Ou RW, Torres MT, Pande SG, Sinskey AJ, Lu TK. 2018. Yeast-based synthetic biology platform for antimicrobial peptide production. ACS Synth Biol 7:896-902.

205. Jiang Z, Vasil AI, Hale JD, Hancock RE, Vasil ML, Hodges RS. 2008. Effects of net charge and the number of positively charged residues on the biological activity of amphipathic alpha-helical cationic antimicrobial peptides. Biopolymers 90:369-383.
206. Zhang SK, Song JW, Gong F, Li SB, Chang HY, Xie HM, Gao HW, Tan YX, Ji SP. 2016. Design of an alpha-helical antimicrobial peptide with improved cell-selective and potent anti-biofilm activity. Sci Rep 6:27394.
207. Wieprecht T, Dathe M, Epand RM, Beyermann M, Krause E, Maloy WL, MacDonald DL, Bienert M. 1997. Influence of the angle subtended by the positively charged helix face on the membrane activity of amphipathic, antibacterial peptides. Biochemistry 36:12869-12880.
208. de la Fuente-Nunez C, Reffuveille F, Mansour SC, Reckseidler-Zenteno SL, Hernandez D, Brackman G, Coenye T, Hancock RE. 2015. D-enantiomeric peptides that eradicate wild-type and multidrug-resistant biofilms and protect against lethal *Pseudomonas aeruginosa* infections. Chem Biol 22:196-205.
209. Zharkova MS, Orlov DS, Golubeva OY, Chakchir OB, Eliseev IE, Grinchuk TM, Shamova OV. 2019. Application of antimicrobial peptides of the Innate Immune system in combination with conventional antibiotics-A novel way to combat antibiotic resistance? Front Cell Infect Microbiol 9:128.
210. Gondil VS, Dube T, Panda JJ, Yennamalli RM, Harjai K, Chhibber S. 2020. Comprehensive evaluation of chitosan nanoparticle based phage lysin delivery system; a novel approach to counter *S. pneumoniae* infections. Int J Pharm 573:118850.
211. Grishin AV, Lavrova NV, Lyashchuk AM, Strukova NV, Generalova MS, Ryazanova AV, Shestak NV, Boksha IS, Polyakov NB, Galushkina ZM, Soboleva LA, Vetchinin SS, Pavlov VM, Karyagina AS, Lunin VG. 2019. The influence of dimerization on the pharmacokinetics and activity of an antibacterial enzyme lysostaphin. Molecules 24.
212. Resch G, Moreillon P, Fischetti VA. 2011. A stable phage lysin (Cpl-1) dimer with increased antipneumococcal activity and decreased plasma clearance. Int J Antimicrob Agents 38:516-521.
213. Becker SC, Dong S, Baker JR, Foster-Frey J, Pritchard DG, Donovan DM. 2009. LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. FEMS Microbiol Lett 294:52-60.
214. Grishin AV, Shestak NV, Lavrova NV, Lyashchuk AM, Popova LI, Strukova NV, Generalova MS, Ryazanova AV, Polyakov NB, Galushkina ZM, Soboleva LA, Boksha IS, Karyagina AS, Lunin VG. 2019. Fusion of lysostaphin to an albumin binding domain prolongs Its half-life and bactericidal activity in the systemic circulation. Molecules 24.
215. Szweda P, Gorczyca G, Tylingo R, Kurlenda J, Kwiecinski J, Milewski S. 2014. Chitosan-protein scaffolds loaded with lysostaphin as potential antistaphylococcal wound dressing materials. J Appl Microbiol 117:634-642.
216. Johnson CT, Wroe JA, Agarwal R, Martin KE, Guldberg RE, Donlan RM, Westblade LF, Garcia AJ. 2018. Hydrogel delivery of lysostaphin eliminates orthopedic implant infection by *Staphylococcus aureus* and supports fracture healing. Proc Natl Acad Sci U S A 115:E4960-E4969.
217. Nithya S, Nimal TR, Baranwal G, Suresh MK, C PA, Anil Kumar V, Gopi Mohan C, Jayakumar R, Biswas R. 2018. Preparation, characterization and efficacy of lysostaphin-chitosan gel against *Staphylococcus aureus*. Int J Biol Macromol 110:157-166.
218. Nour El-Din HT, Elhosseiny NM, El-Gendy MA, Mahmoud AA, Hussein MMM, Attia AS. 2020. A rapid lysostaphin production approach and a convenient novel lysostaphin loaded nano-emulgel; as a sustainable low-cost methicillin-resistant *Staphylococcus aureus* combating platform. Biomolecules 10.

219. Hajiahmadi F, Alikhani MY, Shariatifar H, Arabestani MR, Ahmadvand D. 2019. The bactericidal effect of lysostaphin coupled with liposomal vancomycin as a dual combating system applied directly on methicillin-resistant *Staphylococcus aureus* infected skin wounds in mice. Int J Nanomedicine 14:5943-5955.
220. Miao J, Pangule RC, Paskaleva EE, Hwang EE, Kane RS, Linhardt RJ, Dordick JS. 2011. Lysostaphin-functionalized cellulose fibers with antistaphylococcal activity for wound healing applications. Biomaterials 32:9557-9567.
221. Shah A, Mond J, Walsh S. 2004. Lysostaphin-coated catheters eradicate Staphylococccus aureus challenge and block surface colonization. Antimicrob Agents Chemother 48:2704-2707.
222. Yeroslavsky G, Girshevitz O, Foster-Frey J, Donovan DM, Rahimipour S. 2015. Antibacterial and antibiofilm surfaces through polydopamine-assisted immobilization of lysostaphin as an antibacterial enzyme. Langmuir 31:1064-1073.
223. Huang CY, Hsu JT, Chung PH, Cheng WT, Jiang YN, Ju YT. 2013. Site-specific N-glycosylation of caprine lysostaphin restricts its bacteriolytic activity toward *Staphylococcus aureus*. Anim Biotechnol 24:129-147.
224. Resch G, Moreillon P, Fischetti VA. 2011. PEGylating a bacteriophage endolysin inhibits its bactericidal activity. AMB Express 1:29.
225. Gokarn YR, McLean M, Laue TM. 2012. Effect of PEGylation on protein hydrodynamics. Mol Pharm 9:762-773.
226. Becker SC, Roach DR, Chauhan VS, Shen Y, Foster-Frey J, Powell AM, Bauchan G, Lease RA, Mohammadi H, Harty WJ, Simmons C, Schmelcher M, Camp M, Dong S, Baker JR, Sheen TR, Doran KS, Pritchard DG, Almeida RA, Nelson DC, Marriott I, Lee JC, Donovan DM. 2016. Triple-acting lytic enzyme treatment of drug-resistant and intracellular *Staphylococcus aureus*. Sci Rep 6:25063.
227. Rohrig C, Huemer M, Lorge D, Luterbacher S, Phothaworn P, Schefer C, Sobieraj AM, Zinsli LV, Mairpady Shambat S, Leimer N, Keller AP, Eichenseher F, Shen Y, Korbsrisate S, Zinkernagel AS, Loessner MJ, Schmelcher M. 2020. Targeting hidden pathogens: cell-penetrating enzybiotics eradicate intracellular drug-resistant *Staphylococcus aureus*. mBio 11:e00209-00220.
228. Nileback L, Widhe M, Seijsing J, Bysell H, Sharma PK, Hedhammar M. 2019. Bioactive silk coatings reduce the adhesion of *Staphylococcus aureus* while supporting growth of osteoblast-like cells. ACS Appl Mater Interfaces 11:24999-25007.
229. Seijsing F, Nileback L, Ohman O, Pasupuleti R, Stahl C, Seijsing J, Hedhammar M. 2020. Recombinant spider silk coatings functionalized with enzymes targeting bacteria and biofilms. Microbiologyopen 9:e993.
230. Abouhmad A, Mamo G, Dishisha T, Amin MA, Hatti-Kaul R. 2016. T4 lysozyme fused with cellulose-binding module for antimicrobial cellulosic wound dressing materials. J Appl Microbiol 121:115-125.
231. Abouhmad A, Dishisha T, Amin MA, Hatti-Kaul R. 2017. Immobilization to positively charged cellulose nanocrystals enhances the antibacterial activity and stability of hen egg white and T4 lysozyme. Biomacromolecules 18:1600-1608.
232. Hernandez-Rocamora VM, Maestro B, de Waal B, Morales M, Garcia P, Meijer EW, Merkx M, Sanz JM. 2009. Multivalent choline dendrimers as potent inhibitors of pneumococcal cell-wall hydrolysis. Angew Chem Int Ed Engl 48:948-951.
233. Sanchez-Puelles JM, Sanz JM, Garcia JL, Garcia E. 1992. Immobilization and single-step purification of fusion proteins using DEAE-cellulose. Eur J Biochem 203:153-159.
234. Maestro B, Gonzalez A, Garcia P, Sanz JM. 2007. Inhibition of pneumococcal choline-binding proteins and cell growth by esters of bicyclic amines. FEBS J 274:364-376.

235. Ribes S, Riegelmann J, Redlich S, Maestro B, de Waal B, Meijer EW, Sanz JM, Nau R. 2013. Multivalent choline dendrimers increase phagocytosis of *Streptococcus pneumoniae* R6 by microglial cells. Chemotherapy 59:138-142.
236. de Gracia Retamosa M, Diez-Martinez R, Maestro B, Garcia-Fernandez E, de Waal B, Meijer EW, Garcia P, Sanz JM. 2015. Aromatic esters of bicyclic amines as antimicrobials against *Streptococcus pneumoniae*. Angew Chem Int Ed Engl 54:13673-13677.
237. Elieh-Ali-Komi D, Hamblin MR. 2016. Chitin and chitosan: production and application of versatile biomedical nanomaterials. Int J Adv Res (Indore) 4:411-427.
238. Drzyzga O, Prieto A. 2019. Plastic waste management, a matter for the 'community'. Microb Biotechnol 12:66-68.
239. Younes I, Rinaudo M. 2015. Chitin and chitosan preparation from marine sources. Structure, properties and applications. Mar Drugs 13:1133-1174.
240. Dua K, Bebawy M, Awasthi R, Tekade RK, Tekade M, Gupta G, De Jesus Andreoli Pinto T, Hansbro PM. 2017. Application of chitosan and its derivatives in nanocarrier based pulmonary drug delivery systems. Pharm Nanotechnol 5:243-249.
241. Dias AM, Dos Santos Cabrera MP, Lima AMF, Taboga SR, Vilamaior PSL, Tiera MJ, de Oliveira Tiera VA. 2018. Insights on the antifungal activity of amphiphilic derivatives of diethylaminoethyl chitosan against *Aspergillus flavus*. Carbohydr Polym 196:433-444.
242. Huang Y, Niu B, Gao Y, Fu L, Li W. 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680-682.
243. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636-W641.
244. Vázquez R, García, E., García, P. 2020. Curated phage lysins collection including identifiers, amino acid sequences, functional domain predictions, architectures and physicochemical properties calculations doi:http://dx.doi.org/10.20350/digitalCSIC/12674, Digital.CSIC.doi:http://dx.doi.org/10.20350/digitalCSIC/12674.
245. Osorio D R-VP, Torres R. 2015. Peptides: a package for data mining of antimicrobial peptides. The R Journal 7:4-14.
246. Dawson RMC. 1986. Data for biochemical research, 3rd ed. Clarendon Press, Oxford.
247. Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105-132.
248. Vázquez R, García, E., García, P. 2021. Sequence-function relationships in phage-encoded bacterial cell wall lytic enzymes and their implications for phage-derived products design. J Virol doi:https://doi.org/10.1128/JVI.00321-21.
249. Kuhn M. 2008. Building predictive models in R using the caret package. Journal of Statistical Software 28.
250. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12:77.
251. Liaw AW, M. 2002. Classification and regression by randomForest. R News 2:18-22.
252. Wang G, Li X, Wang Z. 2016. APD3: the antimicrobial peptide database as a tool for research and education. Nucleic Acids Res 44:D1087-1093.
253. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H. 2012. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40:W597-603.

254. Bairoch A. 2013. Amino acid composition (%) in the UniProtKB/Swiss-Prot data bank. https://web.expasy.org/protscale/pscale/A.A.Swiss-Prot.html. Accessed August 2020.
255. Vázquez R B-GS, Ruiz S, García P. 2021. Mining of Gram-negative surface-active enzybiotic candidates by sequence-based calculation of physicochemical properties. Front Microbiol doi:10.3389/fmicb.2021.660403.
256. Wickham H. 2016. ggplot2 elegant graphics for data analysis, 2nd ed doi:10.1007/978-3-319-24277-4. Springer International Publishing.
257. Wilcox RR. 2005. Introduction to robust estimation and hypothesis testing, Second ed.
258. Wilcox RR TT. 2011. Measuring effect size: a robust heteroscedastic approach for two or more groups. Journal of Applied Statistics 38:1359-1368.
259. Mair P WR. 2020. Robust statistical methods in R using the WRS2 package. Behavior Research Methods 52:464-488.
260. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189-1191.
261. Sayers EW, Beck J, Brister JR, Bolton EE, Canese K, Comeau DC, Funk K, Ketter A, Kim S, Kimchi A, Kitts PA, Kuznetsov A, Lathrop S, Lu Z, McGarvey K, Madden TL, Murphy TD, O'Leary N, Phan L, Schneider VA, Thibaud-Nissen F, Trawick BW, Pruitt KD, Ostell J. 2020. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 48:D9-D16.
262. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29-37.
263. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47:D427-D432.
264. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567-580.
265. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37:420-423.
266. Kall L, Krogh A, Sonnhammer EL. 2007. Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. Nucleic Acids Res 35:W429-432.
267. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652-1662.
268. Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43:W389-394.
269. Gautier R, Douguet D, Antonny B, Drin G. 2008. HELIQUEST: a web server to screen sequences with specific alpha-helical properties. Bioinformatics 24:2101-2102.
270. Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16:276-277.
271. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112:531-552.
272. Holm L. 2020. DALI and the persistence of protein shape. Protein Sci 29:128-140.
273. Zallot R, Oberg N, Gerlt JA. 2019. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. Biochemistry 58:4169-4182.

274. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498-2504.
275. Barbier M, Oliver A, Rao J, Hanna SL, Goldberg JB, Alberti S. 2008. Novel phosphorylcholine-containing protein of *Pseudomonas aeruginosa* chronic infection isolates interacts with airway epithelial cells. J Infect Dis 197:465-473.
276. Barbier M, Martinez-Ramos I, Townsend P, Alberti S. 2012. Surfactant protein A blocks recognition of *Pseudomonas aeruginosa* by CKAP4/P63 on airway epithelial cells. J Infect Dis 206:1753-1762.
277. Martinez-Ramos I, Mulet X, Moya B, Barbier M, Oliver A, Alberti S. 2014. Overexpression of MexCD-OprJ reduces *Pseudomonas aeruginosa* virulence by increasing its susceptibility to complement-mediated killing. Antimicrob Agents Chemother 58:2426-2429.
278. Lopez-Causape C, Rojo-Molinero E, Mulet X, Cabot G, Moya B, Figuerola J, Togores B, Perez JL, Oliver A. 2013. Clonal dissemination, emergence of mutator lineages and antibiotic resistance evolution in *Pseudomonas aeruginosa* cystic fibrosis chronic lung infection. PLoS One 8:e71001.
279. Lacks S, Hotchkiss RD. 1960. A study of the genetic material determining an enzyme in Pneumococcus. Biochim Biophys Acta 39:508-518.
280. Green MR, Sambrook J, Sambrook J. 2012. Molecular cloning : a laboratory manual, 4th ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
281. Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557-580.
282. Sanz JM, Lopez R, Garcia JL. 1988. Structural requirements of choline derivatives for 'conversion' of pneumococcal amidase. A new single-step procedure for purification of this autolysin. FEBS Lett 232:308-312.
283. Bustamante N, Campillo NE, Garcia E, Gallego C, Pera B, Diakun GP, Saiz JL, Garcia P, Diaz JF, Menendez M. 2010. Cpl-7, a lysozyme encoded by a pneumococcal bacteriophage with a novel cell wall-binding motif. J Biol Chem 285:33184-33196.
284. Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.
285. Zhou R, Chen S, Recsei P. 1988. A dye release assay for determination of lysostaphin activity. Anal Biochem 171:141-144.
286. Yunck R, Cho H, Bernhardt TG. 2016. Identification of MltG as a potential terminase for peptidoglycan polymerization in bacteria. Mol Microbiol 99:700-718.
287. Alvarez L, Hernandez SB, de Pedro MA, Cava F. 2016. Ultra-sensitive, high-resolution liquid chromatography methods for the high-throughput quantitative analysis of bacterial cell wall chemistry and structure. Methods Mol Biol 1440:11-27.
288. CLSI. 2018. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically, 11th Edn. Clinical and Laboratory Standards Institute, Wayne, PA.
289. CLSI. 1999. Methods for determining bactericidal activity of antimicrobial agents, 1st Edn. Clinical and Laboratory Standards Institute, Wayne, PA.
290. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, 3rd, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci 27:293-315.
291. Li Q, Cheng W, Morlot C, Bai XH, Jiang YL, Wang W, Roper DI, Vernet T, Dong YH, Chen Y, Zhou CZ. 2015. Full-length structure of the major autolysin LytA. Acta Crystallogr D Biol Crystallogr 71:1373-1381.

292. Fujinami Y, Hirai Y, Sakai I, Yoshino M, Yasuda J. 2007. Sensitive detection of *Bacillus anthracis* using a binding protein originating from gamma-phage. Microbiol Immunol 51:163-169.
293. Kikkawa H, Fujinami Y, Suzuki S, Yasuda J. 2007. Identification of the amino acid residues critical for specific binding of the bacteriolytic enzyme of gamma-phage, PlyG, to *Bacillus anthracis*. Biochem Biophys Res Commun 363:531-535.
294. Kelly G, Prasannan S, Daniell S, Fleming K, Frankel G, Dougan G, Connerton I, Matthews S. 1999. Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. Nat Struct Biol 6:313-318.
295. Hermoso JA, Monterroso B, Albert A, Galan B, Ahrazem O, Garcia P, Martinez-Ripoll M, Garcia JL, Menendez M. 2003. Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. Structure 11:1239-1249.
296. Willing SE, Candela T, Shaw HA, Seager Z, Mesnage S, Fagan RP, Fairweather NF. 2015. *Clostridium difficile* surface proteins are anchored to the cell wall using CWB2 motifs that recognise the anionic polymer PSII. Mol Microbiol 96:596-608.
297. Brand S, Niehaus K, Puhler A, Kalinowski J. 2003. Identification and functional analysis of six mycolyltransferase genes of *Corynebacterium glutamicum* ATCC 13032: the genes *cop1*, *cmt1*, and *cmt2* can replace each other in the synthesis of trehalose dicorynomycolate, a component of the mycolic acid layer of the cell envelope. Arch Microbiol 180:33-44.
298. Mesnage S, Dellarole M, Baxter NJ, Rouget JB, Dimitrov JD, Wang N, Fujimoto Y, Hounslow AM, Lacroix-Desmazes S, Fukase K, Foster SJ, Williamson MP. 2014. Molecular basis for bacterial peptidoglycan recognition by LysM domains. Nat Commun 5:4269.
299. Pei J, Grishin NV. 2005. COG3926 and COG5526: a tale of two new lysozyme-like protein families. Protein Sci 14:2574-2581.
300. Korndorfer IP, Danzer J, Schmelcher M, Zimmer M, Skerra A, Loessner MJ. 2006. The crystal structure of the bacteriophage PSA endolysin reveals a unique fold responsible for specific recognition of *Listeria* cell walls. J Mol Biol 364:678-689.
301. Loessner MJ, Kramer K, Ebel F, Scherer S. 2002. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. Mol Microbiol 44:335-349.
302. Lu JZ, Fujiwara T, Komatsuzawa H, Sugai M, Sakon J. 2006. Cell wall-targeting domain of glycylglycine endopeptidase distinguishes among peptidoglycan cross-bridges. J Biol Chem 281:549-558.
303. Grundling A, Schneewind O. 2006. Cross-linked peptidoglycan mediates lysostaphin binding to the cell wall envelope of *Staphylococcus aureus*. J Bacteriol 188:2463-2472.
304. Beaussart A, Rolain T, Duchene MC, El-Kirat-Chatel S, Andre G, Hols P, Dufrene YF. 2013. Binding mechanism of the peptidoglycan hydrolase Acm2: low affinity, broad specificity. Biophys J 105:620-629.
305. Yahashiri A, Jorgenson MA, Weiss DS. 2015. Bacterial SPOR domains are recruited to septal peptidoglycan by binding to glycan strands that lack stem peptides. Proc Natl Acad Sci U S A 112:11347-11352.
306. Alcorlo M, Dik DA, De Benedetti S, Mahasenan KV, Lee M, Dominguez-Gil T, Hesek D, Lastochkin E, Lopez D, Boggess B, Mobashery S, Hermoso JA. 2019. Structural basis of denuded glycan recognition by SPOR domains in bacterial cell division. Nat Commun 10:5567.
307. Chen Y, Simmonds RS, Timkovich R. 2013. Proposed docking interface between peptidoglycan and the target recognition domain of zoocin A. Biochem Biophys Res Commun 441:297-300.

308. van Straaten KE, Dijkstra BW, Vollmer W, Thunnissen AM. 2005. Crystal structure of MltA from *Escherichia coli* reveals a unique lytic transglycosylase fold. J Mol Biol 352:1068-1080.
309. Lee M, Artola-Recolons C, Carrasco-Lopez C, Martinez-Caballero S, Hesek D, Spink E, Lastochkin E, Zhang W, Hellman LM, Boggess B, Hermoso JA, Mobashery S. 2013. Cell-wall remodeling by the zinc-protease AmpDh3 from *Pseudomonas aeruginosa*. J Am Chem Soc 135:12604-12607.
310. Martinez-Caballero S, Lee M, Artola-Recolons C, Carrasco-Lopez C, Hesek D, Spink E, Lastochkin E, Zhang W, Hellman LM, Boggess B, Mobashery S, Hermoso JA. 2013. Reaction products and the X-ray structure of AmpDh2, a virulence determinant of *Pseudomonas aeruginosa*. J Am Chem Soc 135:10318-10321.
311. Buttner FM, Zoll S, Nega M, Gotz F, Stehle T. 2014. Structure-function analysis of *Staphylococcus aureus* amidase reveals the determinants of peptidoglycan recognition and cleavage. J Biol Chem 289:11083-11094.
312. Garcia P, Garcia E, Ronda C, Lopez R, Tomasz A. 1983. A phage-associated murein hydrolase in *Streptococcus pneumoniae* infected with bacteriophage Dp-1. J Gen Microbiol 129:489-497.
313. Bateman A, Rawlings ND. 2003. The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. Trends Biochem Sci 28:234-237.
314. Rigden DJ, Jedrzejas MJ, Galperin MY. 2003. Amidase domains from bacterial and phage autolysins define a family of gamma-D,L-glutamate-specific amidohydrolases. Trends Biochem Sci 28:230-234.
315. Gu J, Feng Y, Feng X, Sun C, Lei L, Ding W, Niu F, Jiao L, Yang M, Li Y, Liu X, Song J, Cui Z, Han D, Du C, Yang Y, Ouyang S, Liu ZJ, Han W. 2014. Structural and biochemical characterization reveals LysGH15 as an unprecedented "EF-hand-like" calcium-binding phage lysin. PLoS Pathog 10:e1004109.
316. Rossi P, Aramini JM, Xiao R, Chen CX, Nwosu C, Owens LA, Maglaqui M, Nair R, Fischer M, Acton TB, Honig B, Rost B, Montelione GT. 2009. Structural elucidation of the Cys-His-Glu-Asn proteolytic relay in the secreted CHAP domain enzyme from the human pathogen *Staphylococcus saprophyticus*. Proteins 74:515-519.
317. Hakulinen N, Tenkanen M, Rouvinen J. 2000. Three-dimensional structure of the catalytic core of acetylxylan esterase from *Trichoderma reesei*: insights into the deacetylation mechanism. J Struct Biol 132:180-190.
318. Martinez C, De Geus P, Lauwereys M, Matthyssens G, Cambillau C. 1992. *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. Nature 356:615-618.
319. Xu W, Chooi YH, Choi JW, Li S, Vederas JC, Da Silva NA, Tang Y. 2013. LovG: the thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. Angew Chem Int Ed Engl 52:6472-6475.
320. Inagaki N, Iguchi A, Yokoyama T, Yokoi KJ, Ono Y, Yamakawa A, Taketo A, Kodaira K. 2009. Molecular properties of the glucosaminidase AcmA from *Lactococcus lactis* MG1363: mutational and biochemical analyses. Gene 447:61-71.
321. Stojkovic EA, Rothman-Denes LB. 2007. Coliphage N4 N-acetylmuramidase defines a new family of murein hydrolases. J Mol Biol 366:406-419.
322. Oliveira H, Melo LD, Santos SB, Nobrega FL, Ferreira EC, Cerca N, Azeredo J, Kluskens LD. 2013. Molecular aspects and comparative genomics of bacteriophage endolysins. J Virol 87:4558-4570.
323. Davies G, Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. Structure 3:853-859.

324. Szewczyk B, Bienkowska-Szewczyk K, Kozloff LM. 1986. Identification of T4 gene 25 product, a component of the tail baseplate, as a 15K lysozyme. Mol Gen Genet 202:363-367.
325. Li Y, Jin K, Setlow B, Setlow P, Hao B. 2012. Crystal structure of the catalytic domain of the *Bacillus cereus* SleB protein, important in cortex peptidoglycan degradation during spore germination. J Bacteriol 194:4537-4545.
326. Rodriguez-Rubio L, Gerstmans H, Thorpe S, Mesnage S, Lavigne R, Briers Y. 2016. DUF3380 domain from a *Salmonella* phage endolysin shows potent *N*-acetylmuramidase activity. Appl Environ Microbiol 82:4975-4981.
327. Anantharaman V, Aravind L. 2003. Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. Genome Biol 4:R11.
328. Xu Q, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Cai X, Carlton D, Chen C, Chiu HJ, Chiu M, Clayton T, Das D, Deller MC, Duan L, Ellrott K, Farr CL, Feuerhelm J, Grant JC, Grzechnik A, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kozbial P, Krishna SS, Kumar A, Lam WW, Marciano D, Miller MD, Morse AT, Nigoghossian E, Nopakun A, Okach L, Puckett C, Reyes R, Tien HJ, Trame CB, van den Bedem H, Weekes D, Wooten T, Yeh A, Hodgson KO, Wooley J, Elsliger MA, Deacon AM, Godzik A, Lesley SA, Wilson IA. 2010. Structure of the gamma-D-glutamyl-L-diamino acid endopeptidase YkfC from *Bacillus cereus* in complex with L-Ala-gamma-D-Glu: insights into substrate recognition by NlpC/P60 cysteine peptidases. Acta Crystallogr Sect F Struct Biol Cryst Commun 66:1354-1364.
329. Ginalski K, Kinch L, Rychlewski L, Grishin NV. 2004. BTLCP proteins: a novel family of bacterial transglutaminase-like cysteine proteinases. Trends Biochem Sci 29:392-395.
330. Bochtler M, Odintsov SG, Marcyjaniak M, Sabala I. 2004. Similar active sites in lysostaphins and D-Ala-D-Ala metallopeptidases. Protein Sci 13:854-861.
331. Fukushima T, Yao Y, Kitajima T, Yamamoto H, Sekiguchi J. 2007. Characterization of new L,D-endopeptidase gene product CwlK (previous YcdD) that hydrolyzes peptidoglycan in *Bacillus subtilis*. Mol Genet Genomics 278:371-383.
332. Schneewind O, Fowler A, Faull KF. 1995. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. Science 268:103-106.
333. Grabowska M, Jagielska E, Czapinska H, Bochtler M, Sabala I. 2015. High resolution structure of an M23 peptidase with a substrate analogue. Sci Rep 5:14833.
334. Horsburgh GJ, Atrih A, Foster SJ. 2003. Characterization of LytH, a differentiation-associated peptidoglycan hydrolase of *Bacillus subtilis* involved in endospore cortex maturation. J Bacteriol 185:3813-3820.
335. Patzer SI, Albrecht R, Braun V, Zeth K. 2012. Structural and mechanistic studies of pesticin, a bacterial homolog of phage lysozymes. J Biol Chem 287:23381-23396.
336. Mooers BH, Matthews BW. 2006. Extension to 2268 atoms of direct methods in the *ab initio* determination of the unknown structure of bacteriophage P22 lysozyme. Acta Crystallogr D Biol Crystallogr 62:165-176.
337. Xiang Y, Morais MC, Cohen DN, Bowman VD, Anderson DL, Rossmann MG. 2008. Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage phi29 tail. Proc Natl Acad Sci U S A 105:9552-9557.
338. Iyer LM, Burroughs AM, Aravind L. 2006. The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like beta-grasp domains. Genome Biol 7:R60.
339. Williams AH, Wheeler R, Thiriau C, Haouz A, Taha MK, Boneca IG. 2017. Bulgecin A: the key to a broad-spectrum Inhibitor that targets lytic transglycosylases. Antibiotics 6:8.

340. Sultana R, Tanneeru K, Guruprasad L. 2011. The PE-PPE domain in *Mycobacterium* reveals a serine alpha/beta hydrolase fold and function: an *in silico* analysis. PLoS One 6:e16745.
341. Oliveira H, Sampaio M, Melo LDR, Dias O, Pope WH, Hatfull GF, Azeredo J. 2019. Staphylococci phages display vast genomic diversity and evolutionary relationships. BMC Genomics 20:357.
342. Zheng J, Wittouck S, Salvetti E, Franz C, Harris HMB, Mattarelli P, O'Toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Ganzle MG, Lebeer S. 2020. A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. Int J Syst Evol Microbiol 70:2782-2858.
343. Brune D, Andrade-Navarro MA, Mier P. 2018. Proteome-wide comparison between the amino acid composition of domains and linkers. BMC Res Notes 11:117.
344. Wang Z, Wang G. 2004. APD: the Antimicrobial Peptide Database. Nucleic Acids Res 32:D590-592.
345. Jenssen H, Hamill P, Hancock RE. 2006. Peptide antimicrobial agents. Clin Microbiol Rev 19:491-511.
346. Kathuria SV, Chan YH, Nobrega RP, Ozen A, Matthews CR. 2016. Clusters of isoleucine, leucine, and valine side chains define cores of stability in high-energy states of globular proteins: Sequence determinants of structure and stability. Protein Sci 25:662-675.
347. Ikai A. 1980. Thermostability and aliphatic index of globular proteins. J Biochem 88:1895-1898.
348. Lata S, Sharma BK, Raghava GP. 2007. Analysis and prediction of antibacterial peptides. BMC Bioinformatics 8:263.
349. Xiao X, Wang P, Lin WZ, Jia JH, Chou KC. 2013. iAMP-2L: a two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Anal Biochem 436:168-177.
350. Meher PK, Sahu TK, Saini V, Rao AR. 2017. Predicting antimicrobial peptides with improved accuracy by incorporating the compositional, physico-chemical and structural features into Chou's general PseAAC. Sci Rep 7:42362.
351. Bhadra P, Yan J, Li J, Fong S, Siu SWI. 2018. AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. Sci Rep 8:1697.
352. Oliveira H, Vilas Boas D, Mesnage S, Kluskens LD, Lavigne R, Sillankorva S, Secundo F, Azeredo J. 2016. Structural and enzymatic characterization of ABgp46, a novel phage endolysin with broad anti-Gram-negative bacterial activity. Front Microbiol 7:208.
353. Garbe J, Bunk B, Rohde M, Schobert M. 2011. Sequencing and characterization of *Pseudomonas aeruginosa* phage JG004. BMC Microbiol 11:102.
354. Rios AC VM, Lima R, Del Fiol FS, Tubino M, Teixeira JA, Balcão VM. 2018. Structural and functional stabilization of bacteriophage particles within the aqueous core of a W/O/W multiple emulsion: A potential biotherapeutic system for the inhalational treatment of bacterial pneumonia. Process Biochemistry 64:177-192.
355. Sun WT, Yinling; Jia, Ming; Hu, Xiaomei; Rao, Xiancai; Hu, Fuquan. 2010. Functional characterization of the endolysin gene encoded by Pseudomonas aeruginosa bacteriophage PaP1. African Journal of Microbiology Research 5:933-939.
356. Briers Y, Lavigne R, Volckaert G, Hertveldt K. 2007. A standardized approach for accurate quantification of murein hydrolase activity in high-throughput assays. J Biochem Biophys Methods 70:531-533.

357. Roig-Molina E, Sanchez-Angulo M, Seele J, Garcia-Asencio F, Nau R, Sanz J, Maestro B. 2020. Searching for antipneumococcal targets: choline-binding modules as phagocytosis enhancers. ACS Infect Dis 6:954-974.
358. Krissinel E. 2015. Stock-based detection of protein oligomeric states in jsPISA. Nucleic Acids Res 43:W314-319.
359. Yang H, Wang DB, Dong Q, Zhang Z, Cui Z, Deng J, Yu J, Zhang XE, Wei H. 2012. Existence of separate domains in lysin PlyG for recognizing *Bacillus anthracis* spores and vegetative cells. Antimicrob Agents Chemother 56:5031-5039.
360. McCarter JD, Withers SG. 1994. Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struct Biol 4:885-892.
361. Hudson KL, Bartlett GJ, Diehl RC, Agirre J, Gallagher T, Kiessling LL, Woolfson DN. 2015. Carbohydrate-aromatic interactions in proteins. J Am Chem Soc 137:15152-15160.
362. Munoz-Munoz J, Cartmell A, Terrapon N, Henrissat B, Gilbert HJ. 2017. Unusual active site location and catalytic apparatus in a glycoside hydrolase family. Proc Natl Acad Sci U S A 114:4936-4941.
363. Litzinger S, Fischer S, Polzer P, Diederichs K, Welte W, Mayer C. 2010. Structural and kinetic analysis of *Bacillus subtilis* N-acetylglucosaminidase reveals a unique Asp-His dyad mechanism. J Biol Chem 285:35675-35684.
364. Mhlongo NN, Skelton AA, Kruger G, Soliman ME, Williams IH. 2014. A critical survey of average distances between catalytic carboxyl groups in glycoside hydrolases. Proteins 82:1747-1755.
365. Eckert C, Lecerf M, Dubost L, Arthur M, Mesnage S. 2006. Functional analysis of AtlA, the major N-acetylglucosaminidase of *Enterococcus faecalis*. J Bacteriol 188:8513-8519.
366. Loh B, Grant C, Hancock RE. 1984. Use of the fluorescent probe 1-N-phenylnaphthylamine to study the interactions of aminoglycoside antibiotics with the outer membrane of *Pseudomonas aeruginosa*. Antimicrob Agents Chemother 26:546-551.
367. Helander IM, Mattila-Sandholm T. 2000. Fluorometric assessment of gram-negative bacterial permeabilization. J Appl Microbiol 88:213-219.
368. Radovic-Moreno AF, Lu TK, Puscasu VA, Yoon CJ, Langer R, Farokhzad OC. 2012. Surface charge-switching polymeric nanoparticles for bacterial cell wall-targeted delivery of antibiotics. ACS Nano 6:4279-4287.
369. Simmen HP, Blaser J. 1993. Analysis of pH and pO2 in abscesses, peritoneal fluid, and drainage fluid in the presence or absence of bacterial infection during and after abdominal surgery. Am J Surg 166:24-27.
370. Poschet J, Perkett E, Deretic V. 2002. Hyperacidification in cystic fibrosis: links with lung disease and new prospects for treatment. Trends Mol Med 8:512-519.
371. de Groot NS, Castillo V, Grana-Montes R, Ventura S. 2012. AGGRESCAN: method, application, and perspectives for drug design. Methods Mol Biol 819:199-220.
372. Abdelkader K, Gutierrez D, Grimon D, Ruas-Madiedo P, Lood C, Lavigne R, Safaan A, Khairalla AS, Gaber Y, Dishisha T, Briers Y. 2020. Lysin LysMK34 of Acinetobacter baumannii Bacteriophage PMK34 Has a Turgor Pressure-Dependent Intrinsic Antibacterial Activity and Reverts Colistin Resistance. Appl Environ Microbiol 86.
373. Zelezetsky I, Tossi A. 2006. Alpha-helical antimicrobial peptides--using a sequence template to guide structure-activity relationship studies. Biochim Biophys Acta 1758:1436-1449.
374. Matsuzaki K. 2009. Control of cell selectivity of antimicrobial peptides. Biochim Biophys Acta 1788:1687-1692.
375. Dosler S, Karaaslan E. 2014. Inhibition and destruction of *Pseudomonas aeruginosa* biofilms by antibiotics and antimicrobial peptides. Peptides 62:32-37.

376. Picoli T, Peter CM, Zani JL, Waller SB, Lopes MG, Boesche KN, Vargas GDA, Hubner SO, Fischer G. 2017. Melittin and its potential in the destruction and inhibition of the biofilm formation by *Staphylococcus aureus*, *Escherichia coli* and *Pseudomonas aeruginosa* isolated from bovine milk. Microb Pathog 112:57-62.
377. Sierra JM, Fuste E, Rabanal F, Vinuesa T, Vinas M. 2017. An overview of antimicrobial peptides and the latest advances in their development. Expert Opin Biol Ther 17:663-676.
378. Giacometti A, Cirioni O, Del Prete MS, Paggi AM, D'Errico MM, Scalise G. 2000. Combination studies between polycationic peptides and clinically used antibiotics against Gram-positive and Gram-negative bacteria. Peptides 21:1155-1160.
379. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. 2018. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 46:D1074-D1082.
380. Buyck JM, Plesiat P, Traore H, Vanderbist F, Tulkens PM, Van Bambeke F. 2012. Increased susceptibility of *Pseudomonas aeruginosa* to macrolides and ketolides in eukaryotic cell culture media and biological fluids due to decreased expression of oprM and increased outer-membrane permeability. Clin Infect Dis 55:534-542.
381. Clifton LA, Skoda MW, Le Brun AP, Ciesielski F, Kuzmenko I, Holt SA, Lakey JH. 2015. Effect of divalent cation removal on the structure of gram-negative bacterial outer membrane models. Langmuir 31:404-412.
382. Calvo P, Remuñán-López, C., Vila-Jato, J.L., Alonso, M.J. 1998. Novel hydrophilic chitosan-polyethylene oxide nanoparticles as protein carriers. J Appl Pol Sci 63:125-132.
383. Briese T, Hakenbeck R. 1985. Interaction of the pneumococcal amidase with lipoteichoic acid and choline. Eur J Biochem 146:417-427.
384. Song M, Li, L., Zhang, Y., Chen, K., Wang, H., Gong, R. 2017. Carboxymethyl-β-cyclodextrin grafted chitosan nanoparticles as oral delivery carrier of protein drugs. Reactive and Functional Polymers 117:10-15.
385. Liu Q, Sacco P, Marsich E, Furlani F, Arib C, Djaker N, Lamy de la Chapelle M, Donati I, Spadavecchia J. 2018. Lactose-modified chitosan gold(III)-PEGylated complex-bioconjugates: from synthesis to interaction with targeted galectin-1 protein. Bioconjug Chem 29:3352-3361.
386. Kadri T, Cuprys, A., Rouissi, T., Brar, S.K., Daghrir, R., Lauzon, J.-M. 2018. Nanoencapsulation and release study of enzymes from *Alkanivorax borkumensis* in chitosan-tripolyphosphate formulation. Biochem Eng J 137:1-10.
387. World Health Organization. 2017. Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis. World Health Organization, Geneva.
388. Abdelkader K, Gerstmans H, Saafan A, Dishisha T, Briers Y. 2019. The preclinical and clinical progress of bacteriophages and their lytic enzymes: the parts are easier than the whole. Viruses 11.
389. Gerstmans H, Criel B, Briers Y. 2018. Synthetic biology of modular endolysins. Biotechnol Adv 36:624-640.
390. Guillen D, Sanchez S, Rodriguez-Sanoja R. 2010. Carbohydrate-binding domains: multiplicity of biological roles. Appl Microbiol Biotechnol 85:1241-1249.
391. Rodriguez-Rubio L, Martinez B, Rodriguez A, Donovan DM, Gotz F, Garcia P. 2013. The phage lytic proteins from the *Staphylococcus aureus* bacteriophage vB_SauS-phiIPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance. PLoS One 8:e64671.
392. Oechslin F, Menzi C, Moreillon P, Resch G. 2021. The multi-domain architecture of a bacteriophage endolysin enables intramolecular synergism and regulation of bacterial lysis. J Biol Chem doi:10.1016/j.jbc.2021.100639:100639.

393. Son B, Kong M, Ryu S. 2018. The auxiliary role of the amidase domain in cell wall binding and exolytic activity of staphylococcal phage endolysins. Viruses 10:284.
394. Cloutier I, Paradis-Bleau C, Giroux AM, Pigeon X, Arseneault M, Levesque RC, Auger M. 2010. Biophysical studies of the interactions between the phage varphiKZ gp144 lytic transglycosylase and model membranes. Eur Biophys J 39:263-276.
395. Holt A CJ, Ramsey J, O'Leary C, Moreland R, Martin C, Galbadage DT, Sharan R, Sule P, Bettridge K, Xiao J, Cirillo J, Young R. 2019. Phage-encoded cationic antimicrobial peptide used for outer membrane disruption in lysis. bioRxiv doi:https://doi.org/10.1101/515445.
396. Morita Y, Tomida J, Kawamura Y. 2014. Responses of *Pseudomonas aeruginosa* to antimicrobials. Front Microbiol 4:422.
397. Fratzl P. 2007. Biomimetic materials research: what can we really learn from nature's structural materials? J R Soc Interface 4:637-642.
398. Kean T, Thanou M. 2010. Biodegradation, biodistribution and toxicity of chitosan. Adv Drug Deliv Rev 62:3-11.
399. Islam N, Dmour I, Taha MO. 2019. Degradability of chitosan micro/nanoparticles for pulmonary drug delivery. Heliyon 5:e01684.
400. Fischer D, Li Y, Ahlemeyer B, Krieglstein J, Kissel T. 2003. *In vitro* cytotoxicity testing of polycations: influence of polymer structure on cell viability and hemolysis. Biomaterials 24:1121-1131.
401. Su C, Liu Y, Li R, Wu W, Fawcett JP, Gu J. 2019. Absorption, distribution, metabolism and excretion of the biomaterials used in Nanocarrier drug delivery systems. Adv Drug Deliv Rev 143:97-114.
402. Choi M, Cho M, Han BS, Hong J, Jeong J, Park S, Cho MH, Kim K, Cho WS. 2010. Chitosan nanoparticles show rapid extrapulmonary tissue distribution and excretion with mild pulmonary inflammation to mice. Toxicol Lett 199:144-152.
403. Islam N, Ferro V. 2016. Recent advances in chitosan-based nanoparticulate pulmonary drug delivery. Nanoscale 8:14341-14358.
404. Ngo CC, Massa HM, Thornton RB, Cripps AW. 2016. Predominant bacteria detected from the middle ear fluid of children experiencing otitis media: a systematic review. PLoS One 11:e0150949.

VIII. ANNEXES

Annex 1: **Articles and patents elaborated during the research time of this doctoral thesis.**

Published or accepted for publication:

1. **Vázquez R**, Domenech M, Iglesias-Bexiga M, Menéndez M, García P*. 2017. Csl2, a novel chimeric bacteriophage lysin to fight infections caused by *Streptococcus suis*, an emerging zoonotic pathogen. Sci Rep 7:16506.
2. **Vázquez R**, García E, García P*. 2018. Phage lysins for fighting bacterial respiratory infections: a new generation of antimicrobials. Front Immunol 9:2252.
3. **Vázquez R**, García P*. 2019. Synergy between two chimeric lysins to kill *Streptococcus pneumoniae*. Front Microbiol 10:1251.
4. **Vázquez R**, Blanco-Gañán S, Ruiz S; García P*. 2021. Mining of Gram-negative surface-active enzybiotic candidates by sequence-based calculation of physicochemical properties. Front Microbiol doi: https://doi.org/10.3389/fmicb.2021.660403.
5. **Vázquez R***, García, E, García, P. 2021. Sequence-function relationships in phage-encoded bacterial cell wall lytic enzymes and their implications for phage-derived products design. J Virol doi: https://doi.org/10.1128/JVI.00321-21 .

In preparation (the finally published titles or authors may change):

6. **Vázquez R**, Seoane-Blanco M, Rivero-Buceta V, Susana Ruiz, van Raaij MJ, Pedro García*. 2021. *Pseudomonas aeruginosa* phage JG004 lysozyme (Pae87) contains a bacterial surface-active antimicrobial peptide-like region and an independent substrate binding subdomain.
7. **Vázquez R**, Ruiz S, García P. 2021. Pae87 lysin as a scaffold for the design of novel antimicrobials.
8. **Vázquez R**†, Caro-León FJ†, Nakal A, Ruiz S, Doñoro C, García L, Vázquez-Lasa B, San Román J, Sanz J, García P*, Aguilar MR*. 2021. DEAE-chitosan nanoparticles as a pneumococcus-biomimetic material for the development of antipneumococcal therapeutics.
9. Caro-León J, López Donaire ML, **Vázquez R**, García L, García P, Huerta-Madroñal M, Aguilar MR, Vázquez-Lasa B, San Román J. 2021. Chitosan–diethylaminoethyl-catechol conjugates with antioxidant and antimicrobial properties.
10. Blanco FG, **Vázquez R**, Cañadas O, García P, Prieto MA*. Functionalization of plastic materials with antipneumococcal lysin Cpl-711 using a polyhydroxyalkanoate affinity tag.

*Corresponding author
†Equal contribution

Patent:

Vázquez R, García E, García P. 2020. Polypeptides with antibacterial activity. No. 20382666.4-1118.

Annex 2: **DNA and aa sequences of the genes, proteins and peptides used in this work.** 6 His tags are underlined and catalytic residues in blue. Linkers are shaded in grey and the calcium binding domain of Pae87F is in bold.

>*pae87*
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGG
CTCTGACCGAGCAAGACTTCCAATCGGCTGCCGATGATCTGGGCGTCGATGTTGCCAGTGTAAA
AGCCGTCACCAAAGTAGAGAGTCGTGGGAGCGGCTTTCTGCTGTCTGGCGTCCCGAAAATTCTG
TTCGAACGCCACTGGATGTTCAAACTGCTGAAACGCAAACTGGGTCATGATCCGGAAATTAACG
ACGTTTGCAACCCGAAAGCTGGCGGCTACCTGGGCGGCCAAGCGGAGCACGAACGTCTGGATAA
AGCAGTCAAAATGGATCGCGACTGCGCACTGCAAAGTGCCTCTTGGGGCCTGTTCCAGATTATG
GGCTTCCATTGGGAGGCACTGGGTTATGCGAGTGTTCAGGCATTTGTGAATGCCCAGTATGCTA
GCGAAGGCTCGCAACTGAACACCTTTGTTCGCTTCATCAAAATCAATCCGGCAATCCACAAAGC
TCTGAAATCCAAAAACTGGGCAGAATTCGCAAAACGCTATAACGGGCCGGATTACAAAAAAAAC
AACTACGATGTTAAACTGGCAGAAGCCTATCAATCCTTCAAATAA

>Pae87
MGSSHHHHHHSSGLVPRGSHMALTEQDFQSAADDLGVDVASVKAVTKVESRGSGFLLSGVPKIL
FERHWMFKLLKRKLGHDPEINDVCNPKAGGYLGGQAEHERLDKAVKMDRDCALQSASWGLFQIM
GFHWEALGYASVQAFVNAQYASEGSQLNTFVRFIKINPAIHKALKSKNWAEFAKRYNGPDYKKN
NYDVKLAEAYQSFK

>P87
LNTFVRFIKINPAIHKALKSKNWAEFAKR

>P88
LNAFVRFKKIAPAIFKALKSKNWAQFAKR

>*pae87F*
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGG
CGTTGACCGAGCAAGACTTTCAAAGCGCTGCTGACGATCTGGGGGGTAGACGTTGCGAGCGTCAA
GGCCGTAACCAAAGTAGAATCTCGCGGTAGTGGCTTCTTATTGTCGGGCGTCCCAAAAATCCTG
TTTGAACGTCATTGGATGTTCAAGCTCCTCAAGCGGAAACTTGGTCACGATCCTGAGATTAACG
ATGTTTGTAACCCTAAGGCAGGTGGTTACCTTGGGGGTCAGGCCGAGCACGAGCGTCTCGATAA
GGCTGTCAAGATGGATCGGGACTGCGCATTGCAAAGCGCATCCTGGGGGTTGTTTCAGATTATG
GGCTTCCATTGGGAGGCTCTGGGTTATGCCAGTGTCCAAGCATTTGTCAACGCCCAATACGCCA
GCGAGGGCTCTCAACTCAACACATTTGTACGCTTTATCAAGATCAACCCGGCTATCCATAAGGC
CCTTAAATCTAAAAATTGGGCGGAATTTGCCAAACGGTACAATGGGCCAGACTATAAAAAGAAC
AATTATGATGTAAAGTTGGCAGAGGCATATCAATCTTTCAAGGGCAGCGGCAGCGGCAGCACCT
TTAAAAATTGCACCGAACTGCGCAAAAAATATCCGAATGGCGTGCCGAGCAGCCATCCGGCGTA
TCAGAGCAAAATGGATCGCGATCATGATAATTATGCGTGCGAACGCTAAGCTT

>Pae87F
MGSSHHHHHHSSGLVPRGSHMALTEQDFQSAADDLGVDVASVKAVTKVESRGSGFLLSGVPKIL
FERHWMFKLLKRKLGHDPEINDVCNPKAGGYLGGQAEHERLDKAVKMDRDCALQSASWGLFQIM
GFHWEALGYASVQAFVNAQYASEGSQLNTFVRFIKINPAIHKALKSKNWAEFAKRYNGPDYKKN
NYDVKLAEAYQSFKGSGSGS**TFKNCTELRKKYPNGVPSSHPAYQSKMDRDHDNYACER**

>pae87RF
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGG
CCCTGACAGAGCAGGATTTTCAGAGCGCCGCCGACGATCTTGGTGTTGATGTGGCCTCAGTTAA
AGCGGTTACGAAGGTTGAGAGTCGCGGCAGCGGCTTCCTGCTCTCAGGCGTGCCGAAAATCTTG
TTCGAGCGTCATTGGATGTTCAAATTACTCAAGCGCAAACTGGGTCATGACCCTGAAATCAATG
ATGTGTGTAACCCTAAGGCCGGGGGGTATTTAGGGGGGCAGGCAGAGCACGAACGGTTAGACAA
GGCCGTGAAGATGGACCGTGACTGCGCCCTGCAGTCGGCATCCTGGGGTTTGTTTCAAATTATG
GGTTTTCATTGGGAGGCTTTAGGGTATGCATCGGTTCAGGCATTTGTCAACGCGCAATATGCAT
CAGAAGGGTCTCAACTCAACACCTTTGTACGGTTTATTAAGATCAACCCTGCGATCCATAAAGC
ATTAAAGAGCAAAAATTGGGCTGAGTTCGCTAAACGGTACAACGGTCCTGACTATAAGAAGAAT
AACTACGACGTGAAGCTGGCTGAGGCTTACCAAAGTTTTAAAGGGAGTGGTTCCGGCTCGGATC
GGGTATCGAAATATCCAAAGGGCCCATTCGACTACGCCAAGAATCAGGAGTCGAACTGCCACAT
GAGCAAAAACGCCACAACTGAATGCCGCCGTTTACCTGACCACTACTAA

>Pae87RF
MGSSHHHHHHSSGLVPRGSHMALTEQDFQSAADDLGVDVASVKAVTKVESRGSGFLLSGVPKIL
FERHWMFKLLKRKLGHDPEINDVCNPKAGGYLGGQAEHERLDKAVKMDRDCALQSASWGLFQIM
GFHWEALGYASVQAFVNAQYASEGSQLNTFVRFIKINPAIHKALKSKNWAEFAKRYNGPDYKKN
NYDVKLAEAYQSFKGSGSGSDRVSKYPKGPFDYAKNQESNCHMSKNATTECRRLPDHY