Prediction of Antimicrobial Peptides from Metatranscriptomic Samples: A Case Study

By

Fatema Tuj-Johora

17336011

A thesis submitted to the Department of Mathematics and Natural Sciences in partial fulfillment of the requirements for the degree of Bachelor of Science in Biotechnology.

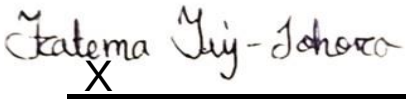
Department of Mathematics and Natural Sciences
BRAC University,

October, 2022

Declaration

I hereby declare that the work presented in this dissertation titled “Prediction of Antimicrobial Peptides from Metatranscriptomic Samples: A Case Study” has been completed by me as a prerequisite submission in requirement of the course “Biotech Project” coded as “BTE450” in the Biotechnology program of the Department of Mathematics and Natural Sciences of BRAC University, Dhaka. This document is my own original work for completing a degree at BRAC University. This document contains original work with some previously published or written materials appropriately acknowledged by the Vancouver format of standardized citation and referencing protocol. This thesis does not contain materials that has been accepted or submitted for any other degree or diploma at a university or other institution. All the primary sources of help have been rightfully acknowledged. The agreement of non-disclosure is made and entered into by and between BRAC University, Dhaka and the student, Fatema Tuj-Johora (ID:17336011), Biotechnology program, Department of Mathematics and Natural Sciences, BRAC University

Student Full Name, ID and Signature:



Fatema Tuj-Johora

ID: 17336011

Approval

The thesis titled "Prediction of Antimicrobial Peptides from Metatranscriptomic Samples: A Case Study" submitted by Fatema Tuj-Johora bearing the ID: 17336011 of Fall 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Biotechnology on the 22nd of September, 2022.

Examining Committee:

Supervisor:

Rafeed Rahman Turjya
Lecturer, Department of Mathematics and Natural Sciences
BRAC University

Program Coordinator:

Dr. Munima Haque
Associate Professor, Department of Mathematics and Natural Sciences
BRAC University

Departmental Head:

Prof. A. F. M. Yusuf Haider
Chairperson, Department of Mathematics and Natural Sciences
BRAC University

Acknowledgment

First and foremost, I would like to thank my research supervisor Mr. Rafeed Rahman Turjya, for his dedicated support and guidance. He continuously provided encouragement and was always willing and enthusiastic to assist in all the possible ways throughout the research project. Without his assistance and dedicated involvement in every step throughout the process, this project would never have been accomplished. I would like to thank you very much for your support and understanding over these past few months.

I would also like to express my gratitude to Dr. Munima Haque, Associate Professor and Director, Biotechnology program as well as our Professor Dr. A F M Yusuf Haider, Chairperson of the MNS department, for providing me the opportunity to complete my undergraduate thesis and for keeping their doors always open whenever I ran into trouble or had a question about my academic problems. They were always there to guide me through all the difficulties and queries that might arrive. Thank you so much for sharing expertise, sincere and valuable guidance and encouragement extended to me.

Getting through my dissertation required more than academic support, and I am forever grateful to my family and friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This dissertation stands as a testament to your constant motivation and encouragement. Above all, I show my utmost gratitude to the almighty who showered me with his mercy and uncountable blessing in all the steps of my life. I would like to conclude by acknowledging all the Department faculty members and university staff who directly or indirectly, have lent their hand in this venture.

Fatema Tuj-Johora

17336011

Abstract

The COVID-19 pandemic has shed new light on the ongoing antimicrobial resistance (AMR) global crisis as the use of antibiotics has dramatically increased to treat severely ill patients. In this post-pandemic era, the world is in urgent need of new and effective drugs that can fight multiple drug-resistant (MDR) bacterial infections. Natural antimicrobial peptides (AMP) are produced as a first line of defense in almost all living beings and they offer the most promising alternative to conventional antibiotics. These peptides are able to eliminate MDR bacteria through a set of killing mechanisms that do not induce bacterial resistance quickly. Moreover, they are effective against other classes of pathogens including virus, fungi, protozoa etc. as well as containing antitumor, anticancer and immunomodulatory properties. The comprehensive advantages of AMPs led to extensive ongoing research and trials to make AMP-based drugs commercially available for clinical purposes. Metagenomics is a culture-independent technique to study and characterize unculturable microbes, which can be applied to identify organisms and their bioactive components, including AMPs, from diverse environmental samples. This review provides a brief overview of the origins, functions, and potentials of AMPs, including the computational identification of putative AMPs, and also discusses the concepts and applications of metagenomics and Metatranscriptomics. Finally, a case study employs a workflow for predicting probable AMPs from the metatranscriptomic data of uncultured marine sediment microbiota, and characterizes the identified peptides both structurally and functionally. The establishment of such prediction pipelines makes way for discovering novel AMPs from the ever-increasing metagenomic data.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The World Health Organization (WHO) has declared multidrug resistance as a serious risk to public health across the globe [1]. The indiscriminate and prolonged use of antibiotics has largely contributed to the development and spread of drug-resistant microorganisms, especially in the developing countries. Approximately, 700,000 deaths per year worldwide are attributed to AMR, with a potential annual loss of up to $3.4 trillion by 2030. While we are already knocked down by the emergence of multi-drug resistant bacteria, the COVID-19 pandemic has drawn attention to the fact that we lack new and advanced therapeutics to tackle new pathogenic microbes. As antibiotics are used indiscriminately in both human and veterinary medicine, as well as in agriculture for disease prevention, these sectors are also suffering from the horrific effects of drug-resistance. Therefore, it is very urgent to search for alternatives to conventional antibiotics, with novel modes of action and less predisposed to bacterial resistance. While searching for a potential replacement of conventional antibiotics, the antimicrobial peptides (AMPs) have recently emerged as one of the most promising alternatives [2-4]. As a result, lots of research and trials are ongoing to develop antimicrobial peptide-based drugs which will effectively kill MDR bacteria.

AMPs are small cationic peptides naturally produced by all living organisms as crucial elements of their innate immune system. AMPs are effective against a wide range of both gram-positive and gram-negative bacteria, yeasts, fungi, viruses and parasites [5]. Moreover, they have immunomodulatory, anti-cancer and anti-tumor properties which adds to their therapeutic values. AMPs have multiple modes of action with a rapid onset of killing which altogether makes the possibilities of bacterial resistance very negligible. Thus, more emphasis is being given to develop AMP-based drugs into clinical trials to battle against rapidly increasing drug-resistance. Several AMPs have already been approved by FDA and widely being used as clinical therapeutic agents to treat various infections, while a large number of them are still ongoing clinical trial phases. The applications of AMP are not limited in clinical fields only, rather they have applications in multiple non-clinical sectors like food industries, animal husbandry, agriculture and aquaculture.

While manual identification of AMPs through wet-lab experiments is pretty expensive and time-consuming, *in silico* prediction is a time- and cost-effective approach for large-scale screening and detection of AMPs [6]. Thus, most of the researchers are using the in-silico approach for screening and modeling novel AMPs [7, 8]. There are many bioinformatic tools available online which are free to use and can predict AMP with absolute accuracy within minutes. Most popular sequence-based AMP prediction tools include CAMPR3 [9], AMP scanner [10], AMPA [11], AntiBP2[12], DBAASP [13], DBAMP [14] etc. These tools use different machine learning algorithms like random forest (RF), support vector machine (SVM), artificial neural network (ANN), and discriminant analysis (DA) to distinguish between AMPs and non-AMPs. Besides predicting AMPs, some of these tools also retrieve the structure, protein domain, family-membership, spectrum of activity and other important physicochemical properties of the peptides.

70% of the Earth's surface is covered by marine sediment and the marine sediment microbiota is a crucial component of the Earth system as it contributes substantially to global biomass. Despite the biodiversity and ecological significance of marine sediment microbiota, there is not much research or studies available on them as they are unculturable in routine laboratory procedures. However, recent advances in machine learning technologies have paved the way to identify novel AMPs from a variety of microbial samples including Taiwanese oolong tea [15], Penaeus [16] and Mexican teosinte [17]. Several metagenomic datasets from the marine sediment microbiota have been generated in the last few years which demonstrate the strong presence of industrially relevant enzymes [18-20], bioactive compounds [21, 22] and most recently, antimicrobial peptides with great therapeutic potentials in it [23]. As this sector has not been explored thoroughly, we chose to conduct our search for novel AMPs from this arena.

As we wanted to predict AMP from marine uncultured microbiota, we took the help of metagenomic and meta-transcriptomic analysis. Both of them are culture-independent techniques of analyzing microbial communities. The key difference between metagenomics and metatranscriptomic relies on the type of biomolecules studied in each area, while Metagenomics deals with the genomic DNA, Metatranscriptomics approaches focus only on the mRNA of an environmental sample. Metagenomics is applied for

exploring the genetic diversity and performing taxonomic analysis of a microbial community. On the other hand, meta-transcriptomics is applied to generate information on the functional gene expressions of a microbial community. For Amp prediction purposes, we only needed protein coding DNA sequences that might code for potential AMP. Thus, we chose metatranscriptomic data as our working sample.

Here, in this article, we offered a short overview of AMPs including their natural sources, structural classification, mechanism of action, clinical applications, the opportunities and challenges to develop AMP-based drugs for clinical applications, the innovative formulation strategies to improve their bioavailability and the machine learning tools for predicting AMP. In addition to that, we briefly discussed metagenomic and meta-transcriptomic approaches and combined the application of meta-transcriptomics and advanced computational analytic tools to identify and characterize AMP candidates from uncultured marine sediment microbiota.

Chapter 2. Antimicrobial Peptide

2.1 Definition

Antimicrobial peptides (AMP), also known as host defense peptides are naturally occurring polypeptides produced by all organisms as the first line of defense [24]. These peptides are short, cationic and amphipathic in nature which have a broad spectrum of activity against various pathogenic microbes including viruses, bacteria, fungi and protozoa [25]. While AMPs are produced in bacteria through non-ribosomal peptide synthesis, they are genetically coded and ribosomally synthesized in the rest of the species [26]. Non-ribosomal AMPs are well-known for their antimicrobial activities and widely being used as antibiotics for decades. However, the ribosomally synthesized AMPs are found to have a crucial role in the evolution and immunity (both adaptive and innate) of complex organisms. For the same reason, they are also known as 'host defense peptides' in higher eukaryotic species. Besides, they are found to be preserved as ancient evolutionary molecules in mammals [27].

2.2 History of Discovery

With the discovery of lysozyme in 1922, the beginning of contemporary innate immunity was ignited by Alexander Fleming. Since then, more and more AMPs have been explored and identified in both higher and lower organisms. These pioneering discoveries grabbed the attention of many other scientists and led to the purification and characterization of thousands more AMPs (Table 2.1). To preserve and keep track of all the newly discovered natural AMPs, several databases like APD [28, 29], DAMPD [30], CAMP [31], and YADAMP [32] were constructed.

Table 2.1: Timeline of AMP discovery (adapted from [33])

Discovered (Year)	Name of AMP
1922	Lysozyme
1928	Nisin

Discovered (Year)	Name of AMP
1939	Gramicidin A, B and C
1942	plant Purothionins
1944	Gramicidin S
1947	Polymyxin/colistin
1967	Melittin (insects)
1970	Bombinin (amphibians)
1973	Plant Kalata B1
1974	purified alpha and beta Purothionins active (plants)
1979	Mastoparan
1981	silk moth Cecropins
1985	Human alpha defensins (HNP1), Microcin
1986	Enterocin AS-48, Pardaxins 1-3, Daptomycin
1987	frog Magainin
1988	Horseshoe crab Tachyplesin, Human histatin, Bactenecin
1989	Apidaecins, horseshoe crab Polyphemusin, HNP4, Patellamide
1990	bacterial Duramycin, Gamma-thionin
1991	pig cathelicidin PR-39, TAP
1992	human HD5, frog Brevinin, bovine Indolicidin, Mersacidin, bacterial Microcin J25
1993	human HD6, Bovine beta defensins, pig Protegrin, amphibian Caerin
1994	Amphibian Dermaseptin, Esculentin, Avian defensin, Protozoan amoebapore
1995	Human cathelicidin, hBD-1
1996	LL-37; Temporin, Thanatin, Buforin, Uperin, Mytilin, Pardaxin 4
1997	Clavanin, Styelin, hBD-2
1998	Lactoferrin, Ranaturin, Maculatin, Lycotoxin, Granulysin
1999	Circulin, RTD-1
2000	thrombocidin, Aurein, rCRAMP, Palustrin

Discovered (Year)	Name of AMP
2001	hBD-3, hBD-4, Ponericin, Dermcidin, Piscidin 1
2002	human RNase 7, Maximin, Hepcidin, Cupiennin, Retrocyclin, Halocidin
2003	human angiogenins, CCL20, Alo-1, Circularin A, PhD1, Halocin C8
2004	Arenicin, Brazzein, Ascaphin, Caenacin (Gly-rich)
2005	human psoriasin (S100A7), Plectasin, Phylloseptin
2006	human RNase 8; Latarcin, Cycloviolacin, Lividin
2007	Dybowskin, Pleurain, Hyposin, Odorranain, AvBD, Longicin
2008	human RNase 3; BTD, Vibi
2009	cn-AMP, Kassinatuerin, Lasioglossin, Lichennicidin, Myxinidin
2010	Caenopore-5, Lucifensin, Temporin-SHf, Centrocin, Beta-amyloid peptide
2011	LCI (no disulfide bond beta sheet), Glycocin F, frog cathelicidin-AL, PAM1
2012	Microcin S (103AA), Slerocin, fungal Eurocin, fungal Micasin
2013	human beta defensin 114 (DEFB114)
2014	2-residue lipopeptide Gageotetrin A, Copsin, human RNase 6, BacFL31
2015	Teixobactin, BnPRP1, Ep-AMP1, cOB1
2016	cPcAMP1/26, Lugdunin, MjPen-II, TLN-58
2017	Urumin, Rattusin
2018	Tur1A
2019	NEMURI, human resistin induced by vitamin A, Darobactin
2020	Cacaoidin, cyclic Cerecyclin
2021	Tryglysin A

2.3 Natural Source of AMPs

Plants, bacteria, fungi, archaea, protists, and mammals are the primary sources of natural AMPs (Table 2.2). According to the reports of Data Repository of Antimicrobial Peptides (DRAMP), 3791 AMPs from six kingdoms have been discovered till now, including 2519 animal, 824 plant, 431 bacterial, 7 protozoal, 6 fungal, and 4 archaeal AMPs [34].

Table 2.2: List of antimicrobial peptides from different natural sources (adapted from [35])

Sl. No.	Peptide Name	Source	Antimicrobial Activity	Ref
From Insects				
1	Acaloleptin	*Acalolepta luxuriosa*	G^+, G^-	[36]
2	Andropin	*Drosophila melanogaster*	G+	[37]
3	Apidaecin IA	*Apis mellifera*	G-	[38]
4	Cecropin	*Hyalophora cecropia*	G-	[39]
5	Defensin- α	*Aedes aegypti*	G^+, G^-	[40]
6	Drosomycin	*Drosophila melanogaster*	F	[41]
7	Holotricin	*Holotrichia diomphalia*	G^+, G^-	[42]
8	Sapecin- α	*Sarcophaga peregrine*	G^+, G^-	[43]
9	Tenicin 1	*Tenebrio molitor*	G^+, G^-	[44]
10	Thanatin	*Podisus maculiventris*	G^+, G^-	[45]
From Humans				
1	Cathelicidins	Human neutrophils	F, G^-, G^+	[46]
2	A Defensins	Human neutrophils	F, G^-, G^+	[47]
3	Human Histatin 8	*Homo sapiens*	F, G^-, G^+	[48]
4	LL37	Neutrophils (*Homo sapiens*)	F, G^-, G^+	[49]
From Animals				
1	Androctonin	*Androctonus australis*	F, G^-, G^+	[50]
2	Bactenecin	Bovine Neutrophils	G^-, G^+	[51]
3	Brevinin	*Rana brevipora porsa*	G^-, G^+	[52]
4	Buforin II	*Bufo gargarizans*	F, G^-, G^+	[53]
5	Cupiennin	*Cupiennius salei*	G^-, G^+	[54]
6	Dermaseptin S1	*Phyllomedusa sauvagii*	G^-, G^+	[55]

7	Lycotoxin	*Lycosa carolinensis*	G^-, G^+	[56]
8	Tachyplesins	*Tachypleus tridentatus*	G^-	[57]
From Plants				
1	Hevein	Latex of rubber trees	F	[58]
2	Purothionins	Wheat endosperm	G^+, G^-	[59]
From Microorganisms				
1	Nisin	*Lactococcus lactis*	G^+	[60]
2	Alamethicin	*Trichoderma viride*	G^+	[61]
3	Enterocin	*Enterococcus*	G^+, G^-	[62]
4	Hominicin	*Staphylococcus hominis* MBBL 2-9	G^+, G^-	[63]
5	Ericin S	*Bacillus subtilis*	G^+	[64]
6	Plantaricin A	*Lactobacillus plantarum*	G^+, G^-	[65]
7	Carnobacteriocin B2	*Carnobacterium piscicola*	G^+, G^-	[66]
8	Leucocin A	*Leuconostoc pseudomesenteroides*	G^+, G^-	[67]
9	Subtilin	*Bacillus subtilis*	G^+	[68]
10	Pyrularia thionin	*Pyrularia pubera*	G^+, G^-	[69]

2.3.1 Mammalian AMPs

AMPs are primarily found inside the granules of neutrophils, in the secretions of mammalian skin and mucosal epithelial cells [26]. Humans, cattle, sheep, and other vertebrates are good sources of mammalian AMPs (Table 2.2). Cathelicidins found in cattle, buffalo, horse, pig, chicken and fishes as well as Defensins found in human neutrophils are examples of two of the most studied mammalian AMPs. However, Defensins are among the fastest-evolving mammalian proteins of which hundreds of variants have already been discovered, there is notable diversity of this AMP even across the primate species. Dairy is another significant source of AMPs, as enzymatic hydrolysis of milk results in the production of AMP. For example, lactoferricin B has been produced

by the peptic digestion of food grade bovine lactoferrin [70].

2.3.2 Amphibian AMPs

Owing to the fact that frog skin contains more than 300 distinct AMPs, they are considered as the primary source of amphibian AMPs [71]. AMPs are found in abundance in the skin secretions of frogs belonging to the Pipidae family's genera *Xenopus*, *Silurana*, *Hymenochirus*, and *Pseudhymenochirus* [72]. Magainins and PGLa originated from the skin of *Xenopus laevis* are among the most well-studied AMPs till date. Marine amphibian *Rana cancrivora* are producers of Cancrin [73].

2.3.3 Insect AMPs

Insects produce a larger repertoire of AMPs than any other taxonomic group as their fat bodies and blood cells produce AMPs extensively to boost their adaptability of survival. Acalolepta luxuriosa, Apis mellifera, Bombyx mori, Galleria mellonella etc are the major source of insect AMPs. Cecropin-A, an insect AMP obtained from Galleria mellonella is effective against several inflammatory diseases and cancers [74]. Royalisin obtained from the royal jelly of *Apis mellifera*, can inhibit the parasite Leishmania major when present in lauric acid-conjugated form along with having antibacterial, antifungal and antiparasitic properties [75]. Coprisin from Copris tripartitus, Melittin from Apis mellifera and Defensin-1 from Apis mellifera are some of the well-established insect AMPs.

2.3.4 Bacterial AMPs

Microbes produce plenty of AMPs for self-defense purposes and to prevent the growth of other microorganisms. Bacterial and archaeal AMPs are also known as Bacteriocins. The majority of the well-established bacteriocins are produced by Gram-positive bacteria like *E. coli* and other enterobacteriaceae and they are also referred to as microcins (small peptides) or colicins (larger proteins) [76]. Bacteriocins are very potent antimicrobials but their effectiveness is limited to the species that are phylogenetically related to the bacteriocin-producing bacteria itself. Microbisporicin produced by Microbispora corallina, Nisin A/Z produced by Lactococcus lactis, Subtilin produced by Bacillus subtilis, Pep5

produced by Staphylococcus epidermidis are some of the bacterial AMPs [77].

2.3.5 Plant AMPs

Plants produce AMPs in its leaves, roots, seeds, flowers, and stems throughout its life as a part of its host defense mechanism [78]. AMPs play a crucial role in plant immunity by protecting them from biotic stress [79] and inhibiting different food spoilage bacteria, mould and yeasts from infecting. Moreover, plant AMPs significantly affect the growth and development of plants. Defensins are continuously produced in the seeds of the radish plant to protect the seedlings from pathogens [80]. Thionins, α-hairpinins, hevein-like peptides, snakins, knottins, cyclotides and lipid-transfer proteins are some of the major plant-derived AMP families.

2.3.6 Marine AMPs

In order to survive harsh conditions like high pressure and low temperatures, different pH and salinity and high level of pollution, marine organisms produce a wide range of novel bioactive compounds [81] and biocatalysts [82] of high therapeutic values. Marine AMPs are cationic and hydrophobic in nature while having structural differences from their terrestrial analogues. Myticusin-beta, a marine-derived AMP produced by Mytilus coruscus shows broad-spectrum antibacterial activity against both gram-positive and gram-negative bacteria such as B. cereus, B. subtilis, Streptococcus mutans, E. coli and Pseudomonas aeruginosa [83]. However, most of the marine AMPs discovered till now are only effective *in vitro*. Other reported marine AMPS include Clavanins A, B, C, D, and E from *Styela clava*, Dicynthaurin from *Halocynthia aurantium*, and Halocyntin from *Halocynthia papillosa* etc.

2.4 Physicochemical Properties

All the natural AMPs have some basic physiochemical features (Table 2.3) like they are typically between 12 and 50 amino acids in length, cationic, amphipathic, containing two or more positively charged residues like arginine, lysine [84] and at least 30% of hydrophobic residues [85]. These physio-chemical features play a key role in the

antimicrobial potency, target organism and killing mechanism of AMPs. For example, the cationic nature of AMPs makes them selective towards anionic microbial surfaces and inert towards neutral mammalian cells. Besides, large numbers of cationic and hydrophobic amino acid residues in AMPs help them to obtain the characteristic amphipathic structure while the unique membrane bound conformation of AMPs are determined by their primary structure.

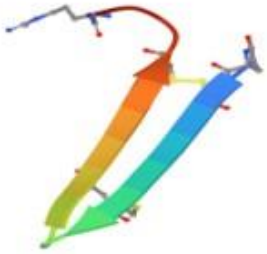
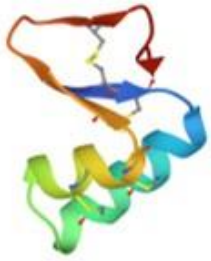
Table 2.3: Physiochemical properties of natural AMPs

Physiological Features	Value
Length	<100
Ion type	Highly cationic
Net charge	between + 2 and + 9
pH	7
Nature	Hydrophobic and amphipathic (In non-polar solvents)
Molecular mass	Between 1 and 5 kDa
Effective against	Gram+ bacteria, Gram- bacteria, virus, fungi, protozoa etc.
Water solubility	Soluble in water
Percentage of Hydrophobic residues	50%
Structure	Amphipathic (in biological membrane/ membrane mimetics)

2.5 Structural Classification:

Antimicrobial peptides are a unique and diverse group of molecules, which are divided into subgroups on the basis of their amino acid composition and structure. Based on the secondary structure, AMPs can be classified into four groups: α-helical peptides (Figure 2.1), β-sheet peptides (Figure 2.2), linear extension structure (Figure 2.4), and both α-helix and β-sheet peptides [86, 87] (Table 2.4 and Figure 2.3).

Table 2.4: Classification of AMPs based on their secondary structures

α-helical peptides	β-sheet peptides	Peptides with both α-helix and β-sheet	Linear extended peptides
 Figure 2.1: 3D structure of LL-37 **Description:** α - helical peptides are unstructured in aqueous solution, but adopt an amphipathic helical structure when exposed to a biological membrane [88]. **Example:** LL-37, human lactoferricin, Aurein 1–2, Mellitin	 Figure 2.2: 3D structure of Gomesin **Description:** β-sheet peptides contain β-hairpins stabilized by disulphide bonds [89]. Due to their rigid structure, the β-sheet peptides are more ordered in aqueous solution and do not undergo conformational changes upon membrane interaction [88]. **Example:** β-defensins, Gomesin, and Protegrin	 Figure 2.3: 3D structure of α1-purothionin **Description:** These peptides can adopt both alpha-helix and beta-sheet conformation. **Example:** α1-purothionin	 Figure 2.4: 3D structure of Indolicin **Description:** Linear Extended peptides lack secondary structure but fold into amphipathic structures after contact with a membrane [90]. Also, these peptides often contain a high content of arginine, proline, tryptophan, and/or histidine residues [86]. **Example:** Indolicidin, Tritrpticin, and Histatins are examples of such peptides.

2.6 Mechanism of Action

The mechanism of action of AMPs can be divided into two major classes: direct killing and immune modulation. The direct killing mechanism of action can be further divided

into membrane targeting and non-membrane targeting mechanism. In the membrane targeting mechanism, AMPs target the outer membrane of the microbes and form pores to kill them and in the non-membrane targeting mechanism, AMPs target the intracellular components like nucleic acid and protein synthesis pathways of the target organism. Besides, some AMPs kill their target in an indirect manner where they modulate the host immunity in various ways such as the activation, attraction, and differentiation of white blood cells, up and down regulation of anti and pro-inflammatory cytokines and stimulation of angiogenesis etc. (Figure 2.5).

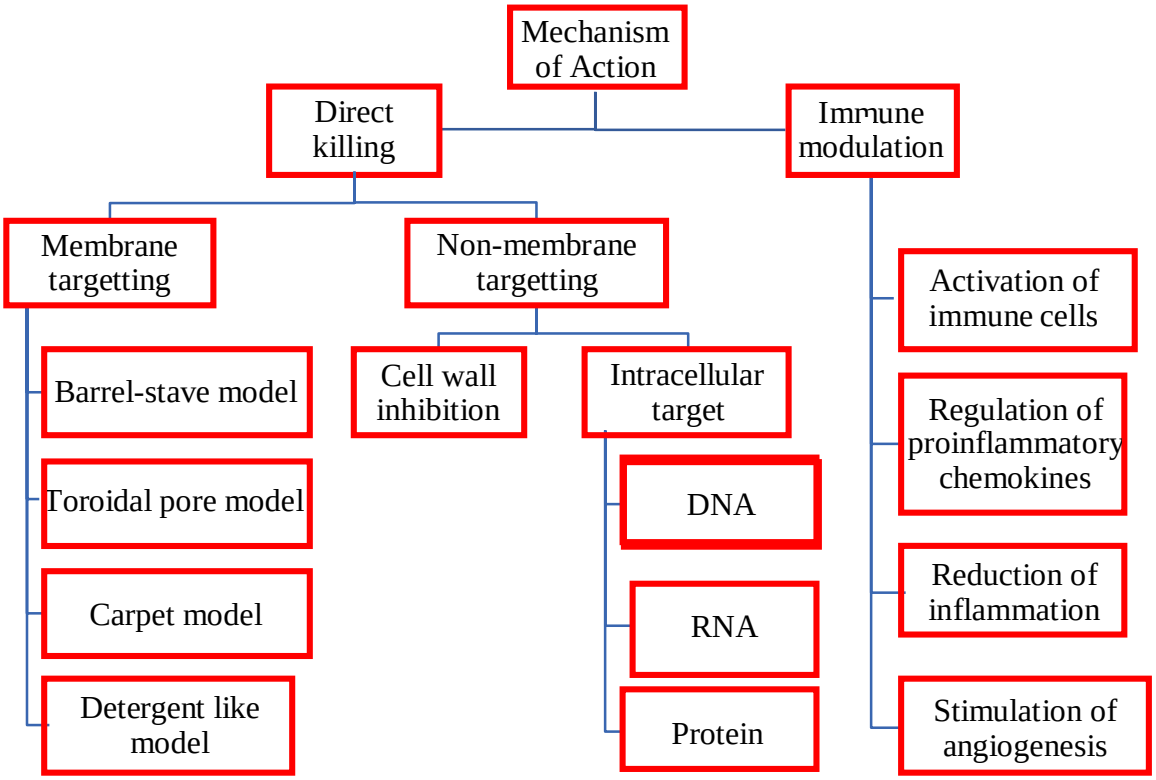


Figure 2.5: Modes of action of AMPs

2.6.1 Membrane Targeting Mechanism via Pore Formation

The presence of teichoic and lipopolysaccharides make the outer surface of Gram-positive and Gram-negative bacteria anionic, initiating electrostatic interaction with cationic AMPs. After attaching themselves to the bacterial membrane, AMPs undergo necessary conformational changes to interact with the cell membrane and form pores. The pores allow

water and ions to flow across the membrane, disintegrate the transmembrane electrochemical gradient and ultimately rupture and lyse the cell. The pores can me formed in four different manners namely the barrel-stave model, the carpet model, the toroidal-pore model and the aggregated channel model (Table 2.5 and Figure 2.6)

Table 2.5: Different pore forming models of AMPs

Method	Description
Barrel-stave model	AMPs bind to the surface of the cell membrane, aggregate with each other to form multimers and insert perpendicularly [92] in the bilayer in a way that their hydrophobic groups are embedded inside the cell membranes to form a pore structure (Figure 2.7 a). The formed channels result in the cytoplasmic outflow of bacterial cell contents and death. Alamethicin, gramicidin, alamethicin and pardaxin form pores in this model.
Toroidal pore model	AMPs aggregate, insert inside the cell membranes, and induce change of the bacterial phospholipid monolayer until a ring hole of 1 to 2 nm in diameter is formed (Figure 2.7 b), ultimately resulting in bacterial death. Magainin 2, melittin and protegrins form pore in this model [93, 94].
Carpet model	Peptides align parallel to the membrane based on electrostatic interactions and cover the cell surface like a carpet (Figure 2.7 c). After a threshold concentration is reached, the detergent-like activity breaks up regions of the membrane and leads to micelle formation. This requires the rotation of phospholipids in a way that the hydrophilic regions face towards the solvent and the hydrophobic regions face towards the membrane. Human cathelicidin LL-37 and melittin are reported to form pore in this way [95].
Detergent like model	AMPs aggregate randomly on the membrane surface, regardless of orientation, and form channels in the membrane of variable sizes by generating micelles (Figure 2.7 d) that result in bacterial death. The AMP maculatin is reported to use this mechanism [95].

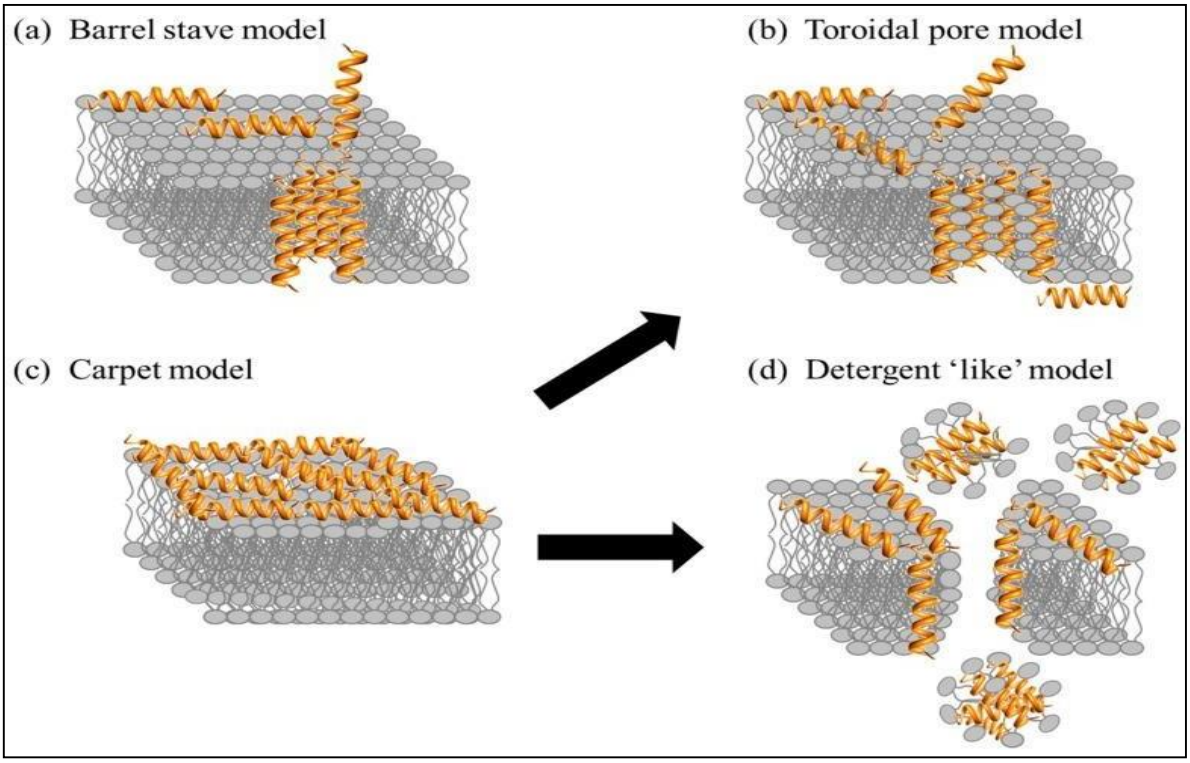


Figure 2.6: Different pore forming models of AMPs (adapted from [91])

2.6.2 Non-membrane Targeting Mechanisms

The non-membrane targeting AMPs interfere with the intracellular components by translocating through the membrane via direct insertion or endocytosis. These AMPs can be divided into two broad categories: AMPs that target the bacterial cell wall and those that have intracellular targets. Some AMPs can have multiple intracellular targets such as inhibition of protein/nucleic acid synthesis and disruption of enzymatic/protein activity [96] (Table 2.6 and Figure 2.7). AMPs that target the bacterial cell wall, bind to various precursor molecules essential for cell wall synthesis. For instance, human β defensin-3[97] and α defensin-1 [98] selectively bind to lipid II that is crucial for cell wall synthesis [97]. AMPs that have an intracellular target, bind to different intracellular organelles like DNA, RNA, ribosomes, proteins and other important cytosolic enzymes to block critical cellular pathways. Buforin II which binds to the DNA and RNA of E. coli [99], indolicidin [100], human α defensin-1[101] and β defensin-4 [102] are examples of AMPs having intracellular targets.

Table 2.6: Intracellular targeting AMPs with their target (adapted from [35])

Sl. No.	AMPs	Intracellular Target	Ref
1	Buforin II, tachyplesin	Binds to DNA	[104]
2	Pleurocidin, dermaseptin, PR-39, HNP-1, HNP-2, Indolicidin	Inhibits DNA, RNA and protein synthesis	[105]
3	Histatins, pyrrhocoricin, Drosocin, Apidaecin	Inhibits enzymatic activity	[106]
4	*N*-acetylmuramoyl-L-alanine, Amidase	Activation of autolysin	[107]
5	PR-39, PR-26, indolicidin, microcin 25	Alters cytoplasmic membrane (inhibits septum formation)	[108]
6	Mersacidin	Inhibits cell-wall Synthesis	[109]

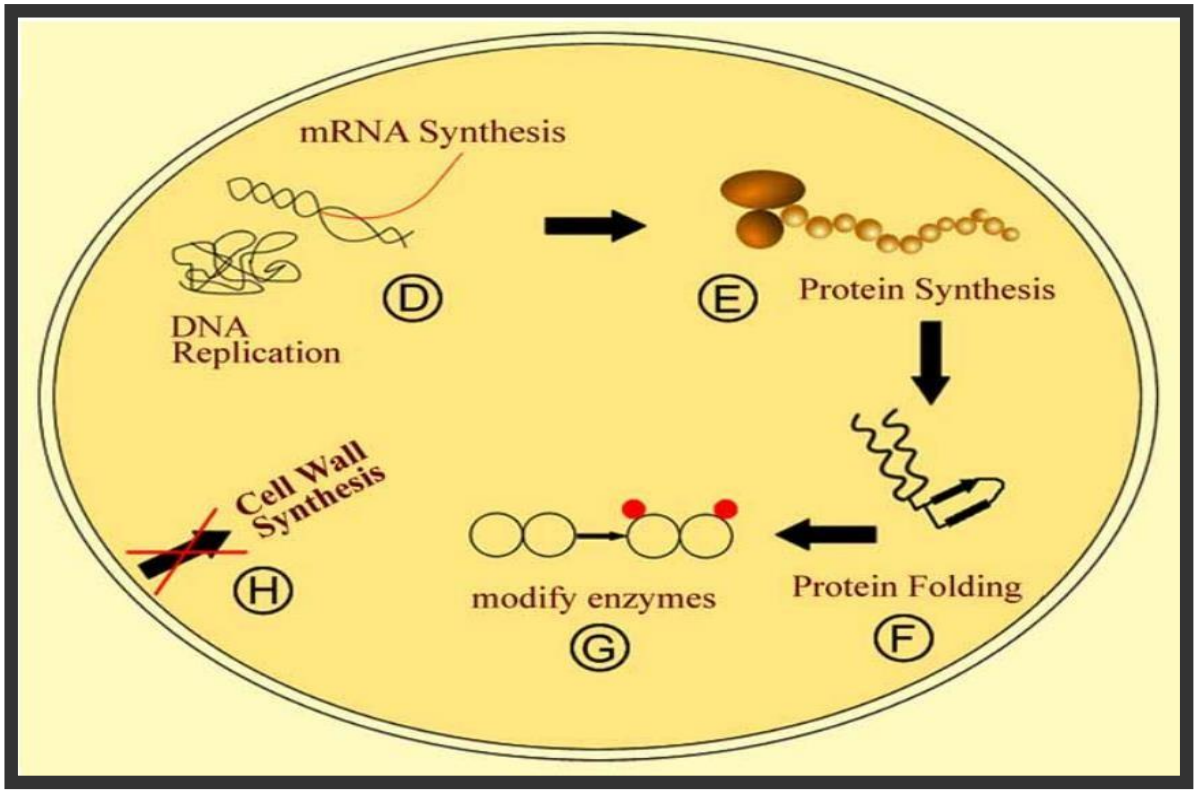


Figure 2.7: Various Intracellular targets of AMPs: (D) Inhibition of DNA replication and RNA synthesis (E) inhibition of protein synthesis, (F) impaired protein folding (G) interaction with enzymes (H) inhibition of cell wall synthesis, (adapted from [103])

2.7 Immunomodulatory Properties

Some AMPs, instead of directly killing the pathogen, help in boosting the host immunity for rapid clearance of infection. AMPs conduct a wide range of immune modulatory activity to enhance the innate and adaptive immune response including activation and recruitment of macrophages, regulation of neutrophil and epithelial cell apoptosis[110], regulation of dendritic cell differentiation and activation, stimulation of epithelial cell migration, production of chemokines and stimulation of chemotaxis[111], suppression of toll-like receptors (TLR), regulation of inflammation by releasing pro/anti-inflammatory cytokines, promotion of angiogenesis, neutralization of endotoxin[112] and enhance wound healing [113] (Figure 2.8 and Table 2.7).

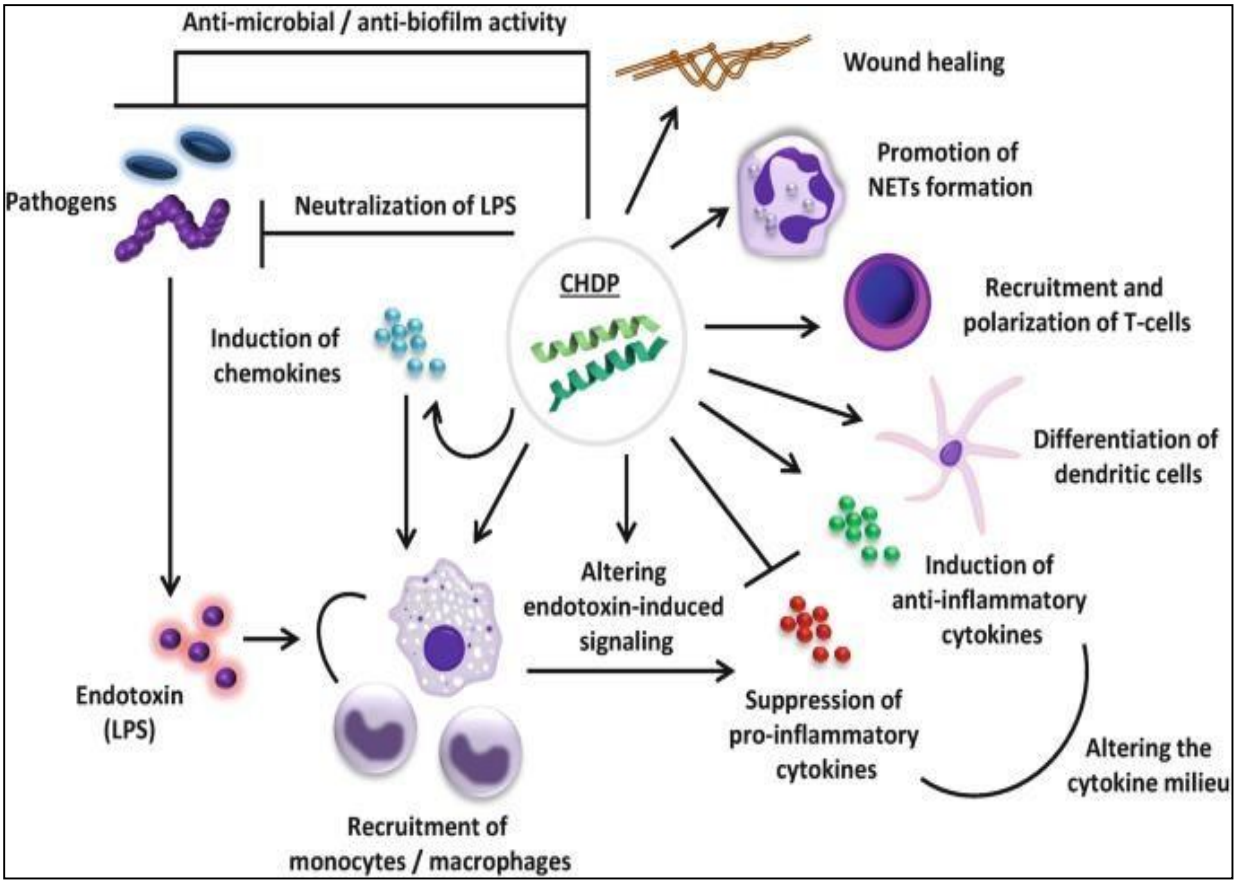


Figure 2.8: Immune modulation by AMPs [114]

Table 2.7: Immunomodulatory AMPs and their functions

AMP	Immune-modulation
β-Defensins	1. Activates primary macrophages and enhance pro-inflammatory responses [115] 2. Chemoattract mast cells [116], leukocytes [117] and dendritic cells [118]
β-Defensin 2	1. Activates immature dendritic cells through interaction with TLR4 [119] 2. Up-regulates IL-1β, IL-6, IL-8, IL-10, MCP-1, MIP-1β, and RANTES in PBMCs [120]
β-Defensin 131	Enhances the expression of cytokines IL-1, IL-6, and IL-12 & chemokines CCL20, CCL22, and CXL8 in human prostate cancer cells [121]
Human β-defensin-3	Activate APCs via TLR1- and TLR2-mediated signaling [122]
human α-defensins (1-3)	1. Increases the expression of the pro-inflammatory cytokines TNF-α and IL-1 β in human monocytes [123] 2. Induces the migration of macrophages and mast cells [124] 3. Chemoattracts various types of immune cells including monocytes, immature DCs, and naïve CD4+ T cells [125]
Apidaecin	1. At high concentration: upregulates the production of the T-cell costimulatory molecule CD80 & cytokines and chemokines in macrophages 2. In low concentration: inhibits IL-6, TNF-α, FGF, and eotaxin in monocytes [126]
Tiger17	1. Promotes remarkable wound healing via recruiting macrophages to the wound site during the inflammatory reaction stage 2. Activates mitogen-activated protein kinase signaling pathways 3. Aids tissue formation and tissue remodeling through the release of transforming TGF-β1 and IL-6 in macrophages 4. Promotes migration and proliferation of both keratinocytes and fibroblasts leading to re-epithelialization and granulation.[127]

IDR-1 and IDR-1018	1. Enhances anti-inflammatory functions 2. Facilitates necessary pro-inflammatory activities needed for the resolution of infection, by driving macrophage differentiation towards an intermediate M1–M2 phenotype [128] 3. Suppress pro-inflammatory cytokines in mice infection models [129, 130]
LL-37	1. Suppresses the LTA (TLR2)- and LPS (TLR4)-induced production of TNF-α, IL-1β, IL-6, and IL-8 in primary monocytes [131] 2. Chemoattracts immune cells like mast cells [116], leukocytes [117] and dendritic cells [118]
IDR-1002	1. Induces chemokine production in human peripheral blood mononuclear cells 2. Induces neutrophil and monocyte recruitment *in vivo* [132]

2.8. Therapeutic Applications of AMP

AMPs have a broad spectrum of activity against a wide range of microorganisms including bacteria, viruses, fungi, protozoa etc. Additionally, some AMPs have antitumor and anticancer properties [133-135]. The therapeutic applications of AMPs are briefly discussed below:

2.8.1 Antibacterial Activity

As discussed earlier, cationic AMPs have strong electrostatic attraction towards bacterial membranes due to the presence of the unique anionic components in their plasma membrane. Antibacterial peptides are effective against a wide range common pathogenic bacteria including VRE, *Acinetobacter baumannii*, MRSA, *S. aureus*, *Listeria monocytogenes*, *E. coli*, *Salmonella* and *Vibrio parahaemolyticus*. AMPs kill target bacteria in two ways, either by forming pores on the outer membrane or targeting the intracellular components of bacteria. Bacteriocins are the subset of ABPs and can be of two types: lantibiotics and non-lantibiotics. Bacteriocin like nisin, cecropins and defensins are highly effective against a wide range of Gram-positive bacteria and Gram-negative

bacteria (Table 2.8) [136].

Table 2.8: Antibacterial AMPs with their target bacterial strain (adapted from [137])

AMP	Source	Target bacteria	Ref
ZmD32	Corn	*E. coli, Bacillus subtilis, P. aeruginosa,* and *S. aureus*	[138]
LL-37	Human cathelicidin hCAP18	*Methicillin-resistant Staphylococcus aureus (MRSA), methicillin-susceptible S. aureus, Vancomycin Intermediate Staphylococcus aureus (VISA)* and Vancomycin Resistant *Staphylococcus aureus (VRSA)*	[139]
Melimine and Mel4		*P. aeruginosa*	[140]
Cecropin A	Moth	Uropathogenic *E. coli* (UPEC)	[141]
BING	Japanese medaka plasma	Broad spectrum including *E. coli, Enterococcus faecalis, S. aureus and P. aeruginosa A*	[142]
D-Cateslytin	Human	*Methicillin-susceptible Staphylococcus aureus, Methicillin-resistant Staphylococcus aureus, Pseudomonas micra, Pseudomonas intermedia* and *F. nucleatum*	[143]
Guavanin 2	guava	*E. coli, Listeria ivanovii* and *Candida parapsilosis*	[144]
Thanatin		*E. coli*and*K. pneumoniae*	[145]
Temporin B	Frog skin	*Staphylococcus epidermidis*	[146]
Oncocin	Milkweed bug	*P. aeruginosa, E. coli* and *Acinetobacter baumannii*	[147]

2.8.2 Antibiofilm Activity

Biofilm refers to immobile and surface attached microbial colonies embedded in self-produced extracellular matrix composed of polysaccharides, DNA, and other components [148]. Biofilm formation is a coping mechanism applied by bacteria to escape external

harsh conditions and it makes the microbe resistant to antibiotics [149]. Biofilms are responsible for 80% of all bacterial infections in humans [150, 151]. Fortunately, some AMPs are highly effective against multidrug resistant biofilm bacteria (Table 2.9) [152]. AMPs prevent biofilm formation and destroy pre-formed biofilms via multiple mechanisms such as suppressing extracellular matrix biosynthesis genes, intercellular quorum sensing signaling and c-di-GMP signaling molecules which are crucial for biofilm formation. Also, AMPs target PpGpp, an essential element for the initiation and maintenance of biofilms. All of these result in the breakage of extracellular polymeric matrices of biofilms as well as make bacteria more susceptible to conventional antibiotics.

Table 2.9: Antibiofilm peptides and their MOA (adapted from [153])

Antibiofilm peptide	Source	Active against	MOA	Ref
LL-37	Human	*Pseudomonas aeruginosa*	downregulates the genes required for biofilm formation and influences QS system	[154]
IDR-1018	*De novo*	*Pseudomonas aeruginosa*	Decrease intracellular (p) PpGpp	[155]
P1	*Calliphora vicina*	*Escherichia coli, Staphylococcus aureus, Acinetobacter baumannii*	Degrades biofilm matrix	[156]
Human β-defensin 3	Human	*Stahyloccocus epidermidis*	Targets icaA, icaD and icaR genes	[157]
Nal-P-113	*De novo*	*Porphyromonas gingivalis*	Down regulates genes related to transport and binding proteins	[158]
Nisin A	*De novo*	*Staphylococcus aureus*	Depolarizes cell membrane	[159]

Antibiofilm peptide	Source	Active against	MOA	Ref
Piscidin-3	Fish	*Pseudomonas aeruginosa*	Degrades eDNA	[160]
lacticin Q	*De novo*	*Staphylococcus aureus*	Depolarizes cell membrane	[159]
Japonicin-2LF	frog skin secretion	*S. aureus;* MRSA	membrane permeabilization, eradication of planktonic as well as sessile pathogens	[161]
melittin	bee venom	*P. aeruginosa*	Membrane disintegration	[162]
Moronecidin	Seahorse	*S. aureus*	Inhibition of surface attachment	[163]
Capsicumicine	Red pepper	*S. epidermidis*	Inhibits establishment of biofilm via matrix anti-assembly mechanism	[164]

2.8.3 Anti-Parasitic Activity

Malaria, leishmaniasis, taeniasis, trypanosomiasis, and schistosomiasis are some of the fetal parasitic diseases caused by *Plasmodium Spp., Leishmania Spp., Trypanosoma* etc. that risk the lives of millions of people across the world. AMPs from various organisms, mostly amphibians and insects are reported to show strong antiparasitic activities against Leishmania, Plasmodium, and Trypanosoma. A (Table 2.10). The outer leaflet of the Protozoan membrane consists of anionic phospholipids which make AMPs more specific towards parasites over host cells [165]. The main mechanism by which AMPs act on parasites is the membrane disruption. AMPs destroy the cellular membranes of protozoa, disrupt the electrochemical gradient and thus disturb the protozoan homeostasis resulting in osmotic shock in pathogen cells [166]. AMPs can also act via intracellular targets and interfere with the key pathways in the parasite metabolism.

Table 2.10: Antiparasitic peptides and their MOA (adapted from [167, 168])

AMP	Effective Against	MOA	Ref
Cecropin B	*Plasmodium spp.*	inhibition of oocyst development	[169]
Defensin A	*P. gallinaceum*	inhibition of oocyst proliferation	[170]
PGLa	*P. falciparum, P. caudatum, T. pyriformis, A. castellanii*	inhibition of RBC reinvasion, cell disruption	[171]
Xenopsin	*P. caudatum, T. Pyriformis, A. castellanii*	cell disruption	[172]
Magainin 2	*P. caudatum, T. pyriformis, A. castellanii, P. falciparum C. parvum*	cell disruption, Inhibition of RBC reinvasion	[172], [171, 173]
Ranalexin-1CB	*C. parvum*	inhibition of intracellular growth in A549 cells	[174]
Dermaseptin-O1	*L. amazonensis, T. cruzi*	Biphasic killing	[175, 176]
Bombinin H2 & H4	*L. donovani, L. pifanoi*	inhibition of cell proliferation	[177]
Temporin A, B	*L. donovani, L. pifanoi*	inhibition of cell proliferation	[178]
Phylloseptin O1 & O2	*T. cruzi*	inhibition of cell proliferation	[179]

2.8.4 Antifungal Activity

Peptides with antifungal properties are mainly produced by bacteria, fungi, and actinomycetes and have a broad spectrum of antifungal effect. Antifungal peptides destroy fungal cells in many ways like forming pores on the fungal membrane or inhibiting essential components of the fungal cell wall like β-Glucan and Chitin. For example, Caspofungin, an antifungal peptide used in the treatment of invasive aspergillosis, acts by disrupting the fungal cell wall through inhibition of beta-(1,3)-glucan synthase [180]. Other antifungal AMPS might target nucleic acid biosynthesis and metabolism or cause apoptosis

by generating reactive oxygen species inside the cell. A great number of AFPs are proven effective against common pathogenic fungi like Candida albicans, filamentous fungi (e.g., Aspergillus flavus), yeast, and other mold found in food stuff and agriculture (Table 2.11). AFPs like Caspofungin, Anidulafungin, and Micafungin are some of the widely used antifungal peptides approved by the FDA and widely used for the treatment of systemic candidemia and candidiasis since the early 2000s.

Table 2.11: List of Antifungal peptides (adapted from [181])

Antifungal Peptide	Source	Fungal Species	Ref
Flagellin	*Bacillus*	*Aspergillus niger, Pythium, Botrytis cinerea, Fusarium oxysporum*	[182]
Metabolites BMME-1	*Bacillus*	*Alternaria solani*	[183]
Iturin A	*Bacillus*	*Candida, Hyphomyces cerevisiae, Fusarium* and *Aspergillus*	[184]
Subtilin, Iturin	*Bacillus*	*Beauveria bassiana*	[185]
Chitinase, chitosanase, protease	*Bacillus*	*Fusarium oxysporum, Fusarium solani, Pythium ultimum*	[186]
Chitin-binding protein CBP24	*Bacillus*	*Fusarium, Rhizoctonia subtilis*	[187]
Serine protease	*Bacillus*	*Botrytis cinerea*	[188]
P-1	*Bacillus*	*Trichothecium roseum*	[189]
Fusaricidin A	*Paenibacillus*	*Fusarium oxysporum, Aspergillus niger, Saccharomyces cerevisiae, Magnaporthe grisea,* etc.	[190]
Syringostatin A, syringostatin E	*Pseudomonas*	*Yeasts, filamentous fungi*	[191]
HP 2-20	*Spirillum*	*Candida albicans, Hyphomyces burnetii*	[192]
EntV	*Enterococcus*	*C. albicans, C. tropicalis, C. paraplanatus,* etc.	[193]
Echinocandin B	*Aspergillus*	*Candida*	[194]

Antifungal Peptide	Source	Fungal Species	Ref
AcAFP	*Aspergillus*	*Fusarium oxysporum, Aspergillus niger, Botrytis cinerea,* etc.	[195]
PcPAF	*Penicillium*	*Trichoderma viride, Fusarium oxysporum, Paecilomyces variotii, and Alternaria longipes*	[196]
Aureobasidin A(AbA)	*Aureobasidium*	*Candida, Cryptococcus neoformans, Blastomyces dermatitis,* etc.	[197]
VL-2397	*Acremonium*	*Aspergillus, Cryptococcus neoformans, Candida glabrata,* etc.	[198]
Chitinase	*Streptomyces*	*Aspergillus niger, Candida albicans*	[199]
Chandrananimycin A	*Cladothrix Actinomyces*	*M. miehei*	[200]
Polyoxin D	*Streptomyces*	*Candida albicans, Cryptococcus neoformans,* etc.	[201]
Nikkomycin Z	*Streptomyces*	*Glomus, Aspergillus fumigatus,* etc.	[202, 203]

2.8.5 Antiviral Activity

Antiviral peptides (AVPs) are usually 8–40 amino acid long polycationic peptides having a wide range of antiviral activities. AVPs can manifest their antiviral activities from the initial to the final phase of the viral cycle including blocking viral entry, attachment, fusion, gene expression, and adsorption of viral proteins virions to the host cells [204] (Figure 2.9 and Table 2.12). Some AVPs compete with the viral spike protein for cellular binding sites present in the host cell surface and block viral attachment to the host cell [205]. For example, the Tat antiviral peptide interacts with CXCR4, a co-receptor (protein) of HIV-1, and inhibits the replication of HIV-1[206]. Besides, some AVPs interfere with the cellular pathways including DNA replication and protein synthesis to stop viral gene expression and translation of viral proteins [207]. Also, AVPs can stimulate various immune modulatory activities that stop viral infection [208].

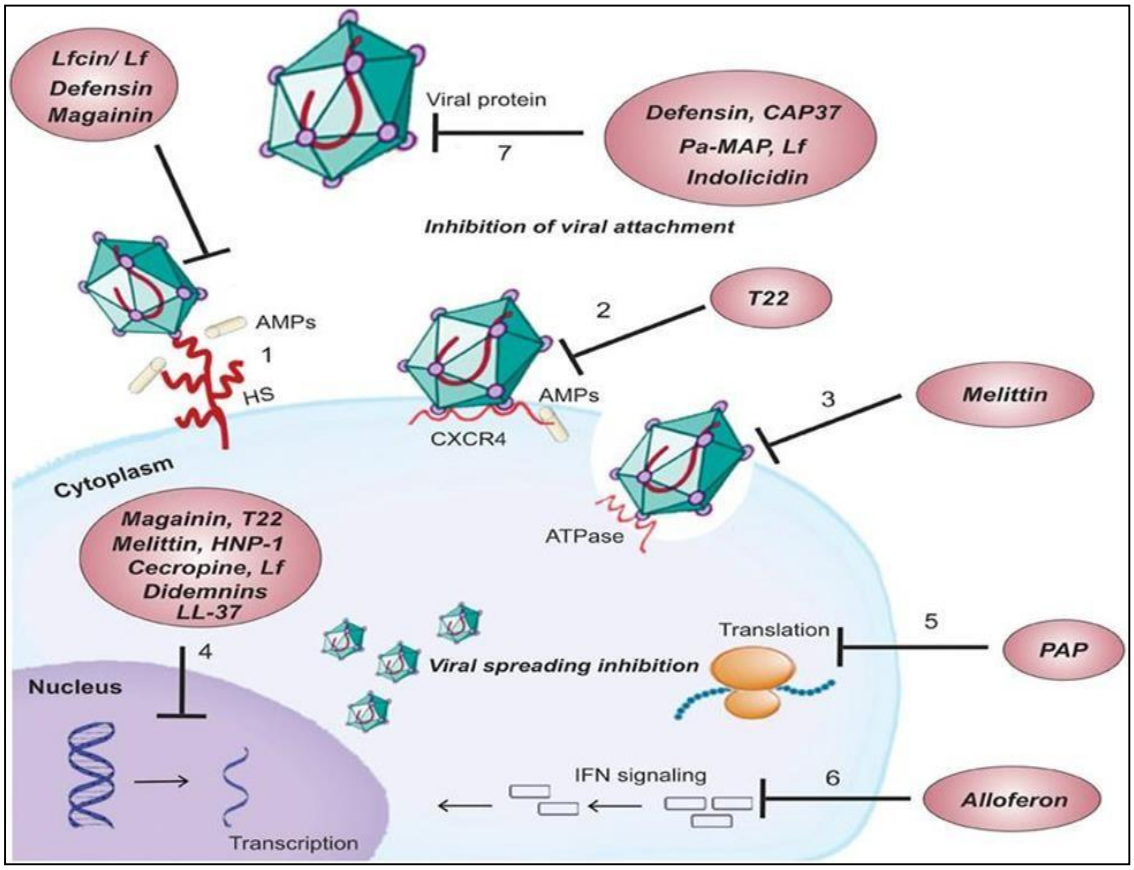


Figure 2.9: Various mechanisms of actions of antiviral peptides: (1) Interacting with glycosaminoglycan present on the cell surface and competing with the virus for cellular binding sites, (2) Binding to the viral CXCR4 co-receptor to block viral entry into the cell, (3) Interfering with ATPase protein to suppress cell fusion, (4) Suppressing viral gene expression, (5) ribosome inactivation to halt peptide chain elongation, (6) Inducing NK and IFN to activate immune modulatory pathways, (7) Inhibition of adsorption/virus-cell fusion by binding to viral proteins. (Adapted from [209])

Table 2.12: Antiviral peptides and their MOA (adapted from [210])

Peptide	Source	Targeted virus	Mode of action/ activity	Ref
Kalata B1	Leaves of *Oldenlandia affinis*	HIV	inhibition of viral-host membrane fusion	[211]
Alloferon 1 and 2		Influenza virus	inhibition by activation of natural	[212]

Peptide	Source	Targeted virus	Mode of action/ activity	Ref
	Hemolymph of blowfly: *Calliphora vicina*		killer cells and release of interferon	
		HSV-1	inhibition of viral replication	
Cecropin A	*Hyalophora cecropia*	HIV; HSV-1 and 2; JUNV	suppression of viral gene expression	[213, 214]
Melittin	*Apis mellifera*	HIV-1; HSV-1 and 2; JUNV	inhibition of cell entry by disruption of envelope	[204, 215]
Magainin I and II	*Xenopus laevis*	HSV-1 and 2	inhibition at cell entry step by disrupting the structure of viral envelope proteins; virucidal activity	------
Temporin B	*Rana temporaria*	HSV-1	inhibition at entry step by interfering in cell-to-cell spread of the virus	--------
Urumin	*Hydrophylax bahuvistara*	Influenza-H1N1 and H1N2	inhibition at cell entry step by targeting cell receptors	[216]
α-Defensin HNPs 1, 2 and 4	Human neutrophil	HIV-1	Inhibition at cell entry step	[217]
β-defensins hDB-2 and 3	Epithelial cells	HPV; VV VZV; HIV	inhibition at cell entry as well as viral replication by late reverse transcripts and nuclear import	[218, 219]

Peptide	Source	Targeted virus	Mode of action/ activity	Ref
Cathelicidin LL-37	Human neutrophil granules	VZV; VV; HSV-1; HIV; RSV; Influenza A; HCV; DENV; ZIKV; VEEV	inhibition of cell entry by disruption of envelope	[220]
		Adenovirus; Aichi virus; Rhinovirus	inhibition of cell entry	[221]
Indolicidin	Bovine neutrophils	HIV	inhibition by membrane-disruption	[222]
Lactoferrin	Mammals' milk	CMV; HSV-1and 2; Adenovirus; Rotavirus; Poliovirus; RSV; HIV; Influenza; HCV; HBV DENV; CHIKV; ZIKV	inhibition at cell entry as well as viral replication	[223]
Protegrin-1	White blood cells of swine	DENV	inhibition of viral replication by binding to viral protease	[224]

AVPs against COVID-19:

Coronaviruses also known as COVID-19 are enveloped RNA viruses that affect the liver, intestinal, respiratory, and nervous system of animals and humans. According to WHO, more than 570M confirmed cases of COVID and 6.39M deaths have been reported so far. Treating COVID patients with traditional antiviral drugs resulted in several side effects, which is why researchers are opting for designing newer and safer drugs to tackle this pandemic. Interestingly, some AVPs have demonstrated prophylactic and therapeutic effects against multiple coronavirus strains (Table 2.13) such as Middle East respiratory syndrome (MERS) virus, severe acute respiratory syndrome (SARS) virus, severe acute

respiratory syndrome coronavirus 2 (SARS-COV-2 or SARS-nCOV19) and other respiratory viruses making them a promising area of research for COVID 19 drug development.

Table 2.13: List of AMPs effective against COVID viruses (adapted from [225])

AMP	Source	Target	MOA	Ref
Mucroporin-M1	Lychas mucronatus	SARS-CoV, MERS-CoV, and influenza H5N1 viruses	disruption of the viral envelope	[226]
EK1	*De novo*	SARS-CoV, MERS-CoV	inhibition of viral fusion entry	[227]
EK1C4	Derivative of EK1	SARS-CoV-2, MERS-CoV	inhibition of viral fusion	[228]
HR2P-M2	Denovo	MERS-CoV	blocks the S protein-mediated membrane fusion	[229]
P9	Derivative of mouse β-defensins-4	SARS-CoV, MERS-CoV, influenza viruses H1N1, H3N2, H5N1, H7N7, and H7N9.	Inhibition of endosomal membrane fusion	[230]
HD5	human defensins-5 (HD5)	SARS-CoV-2	protects the host cells from viral recognition and infection.	[231]
RTD-1	rhesus macaque leukocytes	SARS-CoV	Immunomodulation by reducing proinflammatory cytokine response	[232]

2.8.6 Anticancer Activity

Tumor cell membranes have relatively higher expression of anionic molecules like sialic acid-rich glycoproteins, phosphatidylserine (PS) or heparan sulfate etc. which make them highly anionic and facilitates their preferential binding to cationic AMPs. AMPs are highly

effective against any type of tumor cell let it be the metabolically active ones or slow-growing ones, even the multidrug-resistant ones (Table 2.14). Besides fighting the MDR cancer cells, AMPs show minimal side effects compared to other available chemotherapeutic agents. Also, AMPs are unaffected by chemotherapy resistance mutations, can show synergy with classical chemotherapy and can be used as combination therapy with other therapeutics. AMP target and kill tumor cells in multiple ways including membrane permeabilization [233], apoptosis [234], immune modulation [235], membrane receptor binding [236], inhibition/regulation of DNA synthesis [237] and inhibition of angiogenesis [238] (Figure 2.10).

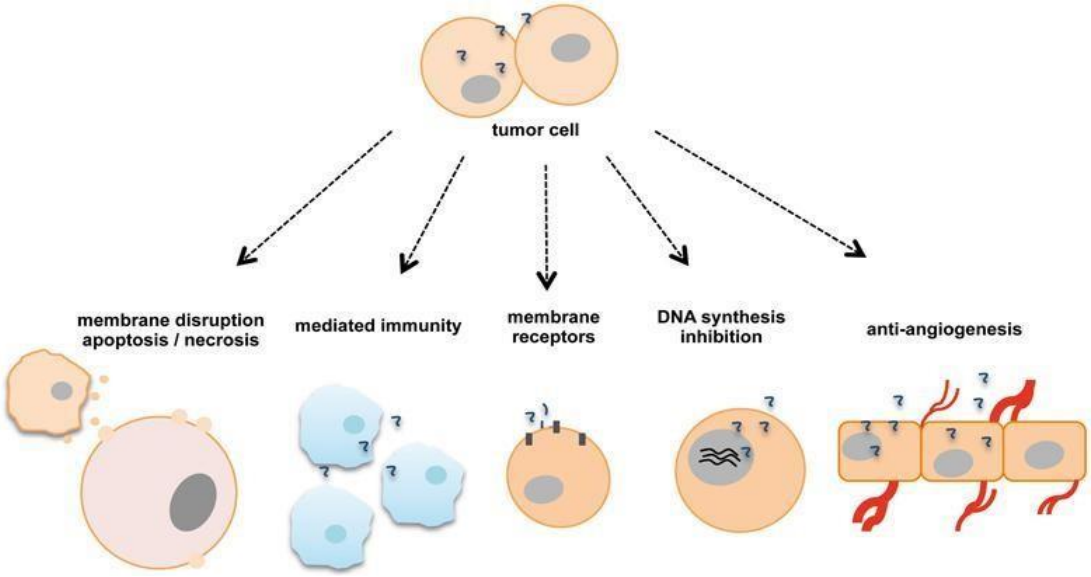


Figure 2.10: Mechanism of Action of ACPs (adapted from [239])

Table 2.14: Anticancer peptides and their MOA

ACPs	Cancer Type	Mechanism	Ref
LL-37	human oral squamous cell carcinoma (OSCC) cells	toroidal pore formation	[240]
HNP-1, HNP-2, and HNP-3	human myeloid leukemia cell line (U937), human erythroleukemic cell line (K562), and lymphoblastoid B cells (IM-9 and WIL-2).	cytolytic activity	[241]

ACPs	Cancer Type	Mechanism	Ref
Human β defensin-3	HeLa, Jurkat and U937 cancer cell lines	cytolysis	[242]
Bovine lactoferricin LTX-315	drug-resistant and drug-sensitive cancer cells	cytolysis and immunogenicity	[243]
Gomesin	murine and human cancer cell lines along with melanoma and leukemia	pore formation via Carpet model	[244]
Mastoparan-C	lung cancer H157, melanocyte MDA-MB-435S, human prostate carcinoma PC-3, human glioblastoma astrocytoma U251MG and human breast cancer MCF-7 cell lines	Induction of apoptosis and activation of phospholipase, selective inhibition of ATPase activity	[245]
Cecropin B1	NSCLC cell line	Inhibiting the growth of tumor via pore formation and apoptosis	[246]
Magainin 2	human lung cancer cells A59 and in Ehrlich's murine ascites cells	Pore formation	[247]
Bufforin IIb	Leukemia, breast, prostate, and colon cancer	cell destruction via mitochondrial apoptosis	[248]
Brevinin 2R	T-cell leukemia Jurkat, B-cell lymphoma BJAB, colon carcinoma HT29/219 and SW742, fibrosarcoma L929, breast adenocarcinoma MCF-7, and lung carcinoma A549 cells	lysosomal death pathway (LDP) and autophagy-like cell death	[249]

ACPs	Cancer Type	Mechanism	Ref
Limnonectes fujianensis br evinvin (LFB)	lung cancer H460, melanoma cell, glioblastoma U251MG, colon cancer HCT116 cell lines	penetrating the lipidic bilayer	[250]
Phylloseptin-PHa	breast cancer cells MCF-7, breast epithelial cells MCF10A	penetrating the lipidic bilayer	[251]
Ranatuerin-2PLx	prostate cancer cell PC-3	cell apoptosis using caspase-3	[252]
Chrysophsin-1, -2 and-3	human fibrosarcoma HT-1080, histiocytic lymphoma U937, and cervical carcinoma HeLa cell lines	disruption of the plasma membrane	[253]
Ss-arasin	human cervical carcinoma HeLa and colon carcinoma HT-29	induction of cytotoxicity	[254]
Turgencin A and B	melanoma cancer cells A2058 and the human fibroblast cell line MRC-5	pore formation and internalization of the lipid bilayer	[255]
Dusquetide (SGX942)	neck and head cancer	binding to p62 to cause membrane damage	[249]

2.8.7 Battling Antibiotic-resistance

The World Health Organization (WHO) has claimed antibiotic resistance to be the biggest challenge in treating infectious diseases today. Conventional Antibiotics are becoming increasingly ineffective as more microbes are becoming drug-resistant leading to persistent infections and death. The main reason microbes easily develop resistance against antibiotics is that antibiotics have fixed targets and they need a longer time to kill microbes. New and effective antibacterial agents are urgently needed to tackle this global concern. AMPs can be the potential solution to this problem as they are able to bypass the common resistance mechanisms that are nullifying the effectiveness of conventional antibiotics. Due

to their multiple modes of action and rapid onset of killing, microbes grow absolutely low/slow resistance against AMPs. Also, AMPs are used in combination with other immune compounds inside the host cells to kill microbes to ensure fewer resistant variants can emerge. A good number of AMPs have shown antibacterial activity against multidrug resistant bacteria *in vivo*. For example, Nisin and mersacidin produced by Lactococcus lactis and Bacillus sp, respectively are active against antibiotic-resistant Gram-positive bacteria [256], G3KL (synthetic) is effective against multidrug-resistant and extensively drug-resistant *Acinetobacter baumannii* and *Pseudomonas aeruginosa*[257], synthetic defensin from *Tribolium castaneum* is effective against drug resistant *Staphylococcus aureus* both *in vitro* and *in vivo*[258] and LL-37 exhibits significant antimicrobial activity against multidrug-resistant *Acinetobacter baumannii*[259].

2.9 Non-Therapeutic Applications

Beside the pharmaceutical applications, AMPs are also used in the food, agricultural and aquaculture industries, as discussed in the following sections.

2.9.1 Food Industry

In the food industries, AMPs are used in multiple ways such as food additives, antimicrobial agents and as components of packaging materials. For example, Lactoferrin is approved and used as an antimicrobial agent in meat products in the USA. ε-Polylysine originated from *Streptomyces albulus,* has been approved by FDA as a food preservative [260]. Natamycin, produced by *Streptomyces* species, is applied on the surface of cheese and salami-type sausages to inhibit fungal growth on them [261]. For antimicrobial packaging of foods, AMPs are incorporated in the packaging materials via adsorption or immobilization on polymer surfaces. AMP-coated food packaging material helps the gradual release of AMPs from packaging material to the food surface and helps to kill the microbes that might be present on food. For example, Dermaseptin K4K20-S4, has been incorporated into different food coatings and showed significant inhibition effects against mould and aerobic bacteria [262].

2.9.2 Animal Husbandry

Antimicrobial peptides as we discussed earlier, are very potent alternatives for treating antibiotic-resistant microbes thus can also be used as an alternative to conventional antibiotic feed additives for improving the growth performance and health of the animals and achieve sustainable livestock production. For example, using Nisin in the dietary supplement of broiler chickens showed a modulating effect on the gut microbial ecology and reduced Bacteroides and Enterobacteriaceae count in ileum digesta [263]. Also, adding colisin E1 [264], cipB-lactoferricin-lactoferrampin [265], and Cecropin AD [266] in the pig diets resulted in increased immune function and reduced intestinal pathogens. Besides, transgenic expression of AMPs in the livestock also resulted in increased protection against various pathogenic infections. For example, mammary gland expression of bovine lactoferricin and human lactoferrin in transgenic goats conferred a wide spectrum of antimicrobial activity against several pathogens [267].

2.9.3 Aquaculture

A number of AMPs are found to be effective against a wide range of fish pathogenic bacteria and viruses, thus they can be used to disinfect aqua environments. For instance, synthetic AMP epinecidin-1 showed effectiveness against a group of bacteria including *E. coli, Pasturella multocida, Aeromonas sobrio, A. hydrophila, Morganella morganii, V. parahaemolyticus, V. vulnificus* etc. [268]. Also, it was found that co-incubation of native cecropin B and a synthetic analogue CF17 with some important fish viral pathogens decreased viral titres upto 10^4 folds [269]. Moreover, a recent study by León et al. (2020) demonstrated that the NKL-24, a truncated peptide derived from zebrafish NK-lysin, effectively killed V. parahaemolyticus via membrane permeabilization [270]. Recently, EmPis-1 was shown to be effective against antibiotic- resistant *E. coli Top10, S. aureus and V. parahaemolyticus OS4* which cause pathogenesis in various aquatic farmed products such as scallops, shrimps and shellfish [271].

2.9.4 Agriculture

AMPs can be used as an environment-friendly alternative of harmful chemical pesticides

against phytopathogens in various fields of agriculture to prevent environment pollution and poisonous side effects on human and animal health. For instance, Short cecropin A-melittin hybrid peptides synthesized by Ferré et al. in 2006 showed antifungal and antibacterial activities against pathogens causing fire blight, halo blight and bacterial spot in crops [272]. Also, recombinant expression of AMPs in plant bodies resulted in building resistance to phytopathogens in various transgenic plants. For example, Mj-AMP1 jalapa defensin provided protection against *Alternaria solani* in transgenic tomatoes and the expression of horseshoe crab-derived tachyplesin I, in tobacco plants provide resistance to the fungal pathogen *Verticullum dahliae* and the phytopathogen *Erwinia carotovora[273]*. More recently, the effectiveness of AMPs has also shown promising results for controlling post-harvest decay caused by various phytopathogens. For instance, the O3TR and C12O3TR peptides successfully controlled P. digitatum in citrus plants both *in vitro* and *in vivo* [274].

2.10 Pros & Cons of AMPs in Clinical Trial

Pharmacological properties of AMPs such as their high cationic charge, hydrophobicity, low molecular mass, relatively simple structures, ease of absorption, low risk for inducing multi-drug resistance etc. have made them highly potential candidates for therapeutics (Figure 2.11(A)). Despite vigorous amounts of research and clinical trials being made to make AMP drugs commercially available, only a few of them have been proven successful and achieved FDA approval to make their way to the market. Currently, nisin, gramicidin, polymyxins, daptomycin and melittin are in clinical use and others are on different trial phases (Table 2.15). However, all of these commercially available AMPs are limited to topical and intravenous applications only in order to prevent any potential toxicity of administration. There are some major drawbacks of AMPs (mentioned in Figure 2.11(B)) which needs to be worked upon to make more AMP drugs clinically available.

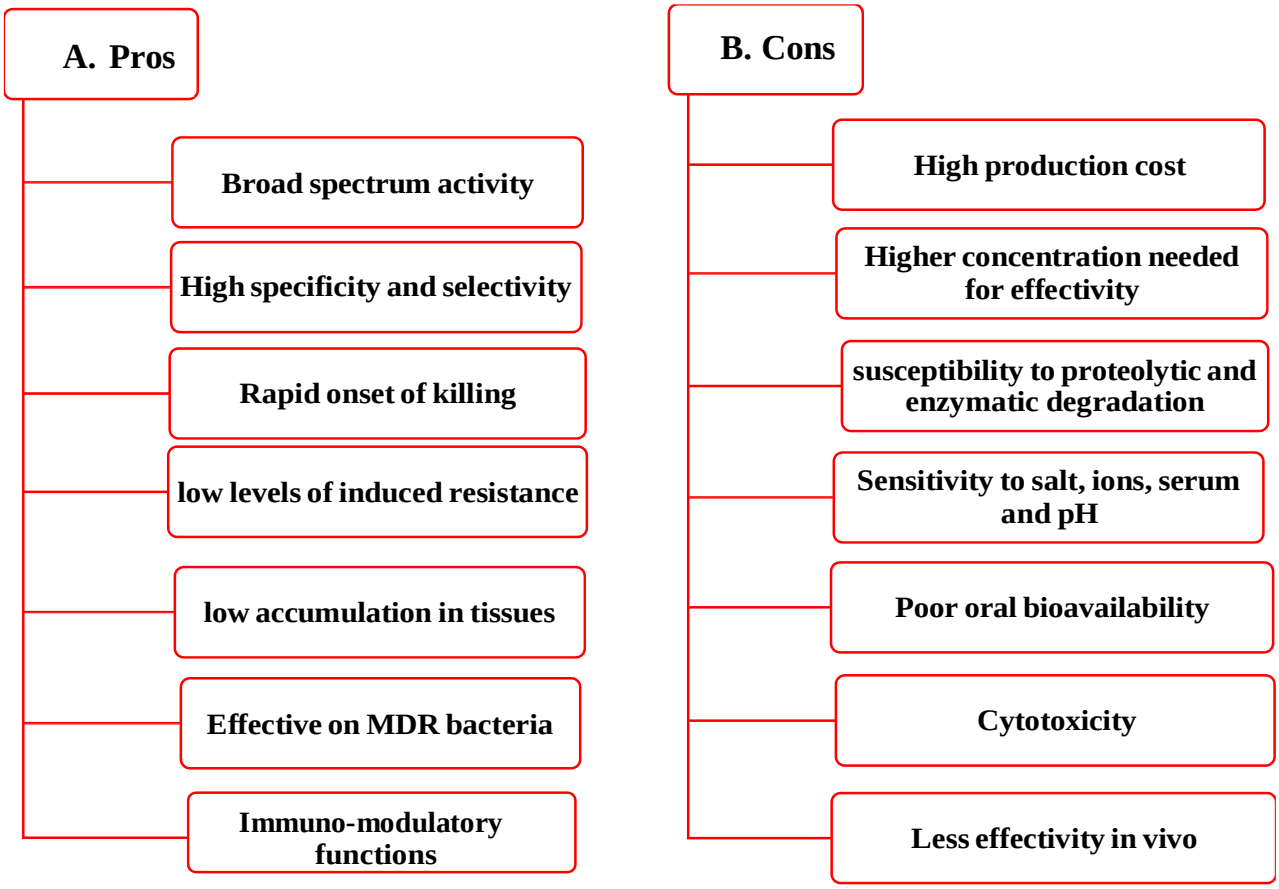


Figure 2.11: Advantages and disadvantages of AMPs in clinical tria

Table 2.15: List of AMP-based Drugs undergoing clinical trials (adapted from [275])

AMP	Description	Activity	Medical use	Dev. Stage
Histatin	Using a variant of histatins, which are naturally occurring cationic peptides in saliva	Antifungal	Antimicrobial-peptide-containing mouthwash for the treatment of oral candidiasis (gingivitis and periodontal diseases)	Phase II-III
hLF1-11	An 11-mer peptide from the N terminus of human lactoferrin	Antibacterial, Antifungal	LPS-mediated diseases and fungal infections	Phase I (completed)

AMP	Description	Activity	Medical use	Dev. Stage
IDR-1	Derivative of bactenecin from bovine neutrophils	Chemokine induction and reduction of pro-inflammatory cytokines	Prevention of infections in the immune compromised	Phase I
Opebacan	21-amino-acid peptide derivative of bactericidal/permeability-increasing protein	Antibacterial, Antiviral	Endotoxemia in hematopoietic stem cell transplant recipients	Phase I/II
AP-214	Synthetic derivative from α-melanocyte-stimulating hormone	Antibacterial	Sepsis and post-surgical organ failure	Phase II (completed)
CZEN-002	Synthetic 8-mer derived from α-melanocyte-stimulating hormone	Anticandidal	Vulvovaginal candidiasis	Phase IIb
EA-230	A derivative peptide from the human pregnancy hormone	Anti-inflammatories; Antiseptics	Sepsis	Phase II
Ghrelin	Endogenous host-defense peptide, synthetic construct	Anti-inflammatory	Airway inflammation, chronic respiratory infection and cystic fibrosis	Phase II (completed)
IMX942	Synthetic cationic host defense peptide, derivative of IDR-1 and indolicidin	Antibacterial	Nosocomial infections, febrile, neutropenia	Phase II
MX-594AN	Indolicidin based antimicrobial peptide variant	Antibacterial, Antifungal	The treatment of catheter-related infections and acne	Phase IIb (completed)

AMP	Description	Activity	Medical use	Dev. Stage
Novexatin (NP213)	Cyclic cationic peptide derived from Nova Biotics arginine peptide platform	Antifungal	Treatment of dermatophyte fungal infections such as onychomycosis	Phase IIb
OP-145	Synthetic 24-mer peptide derived from LL-37	Antibacterial	Chronic bacterial middle ear infection.	Phase II (completed)
P113	A 12 amino acid fragment of histatin 5	Antifungal	HIV	Phase II (completed)
PAC113	A 12 amino-acid antimicrobial peptide derived from histatin	Antifungal	Oral candidiasis	Phase IIb
XOMA-629	9-amino-acid peptide derivative of bactericidal/permeability-increasing protein	Antibacterial	Impetigo	Phase IIA
XMP 629	A 9-amino-acid peptide derived from bactericidal/permeability-increasing protein (BPI)	Antibacterial	Acne	Phase III
Mycoprex	Extracted from insects	Antifungal	Fungal infections	Phase III
Murepavadin (POL7080)	A synthetic analogue of protegrin I	Antibacterial	Treatment of nosocomial pneumonia and ventilator-associated bacterial pneumonia (VABP)	Phase III
Omiganan (MBI-226)	A synthetic analogue of indolicidine	Antibacterial	Treatment of rosacea	Phase III (completed)

2.11 Optimization Strategies for AMPs

As we discussed earlier, natural AMPs have some drawbacks that hinders their clinical availability (Figure 2.16). In order to overcome their limitations and make them suitable for clinical applications, scientists have designed several peptide modification/ formulation strategies and delivery technologies that make AMPs more stable, increase their effectivity and bioavailability *in vivo*, reduce their cytotoxicity, increase their serum half-life and site-directed delivery etc. (Table 2.17).

Table 2.16: Optimization strategies for AMPs

Optimization Strategy	Description
Truncated AMPs	The long primary sequence of natural AMPs makes their commercial production costly and difficult along with making them prone to enzymatic degradation and immunogenicity. Therefore, trimming or removing the minimally active or unwanted regions of AMP reduce their production cost and ease their production process without compromising the antimicrobial potency. For example, LZ1 derived from truncated cathelicidin-BF showed promising antimicrobial activities against *acne vulgaris* causing bacteria including *Propionibacterium acnes, Staphylococcus epidermidis, and Staphylococcus aureus[276].*
Cyclization	Cyclization of peptides with one or multiple disulfide bridges helps to form a secondary or higher structure than their linear form and provides structural rigidity. Reduced flexibility protects AMPs from proteolytic degradation, increases stability, antimicrobial potency and selectivity. For example, cyclic peptide (ZY4) stabilized by a disulfide bridge showed high stability in vivo (half-life of 1.8 h) [277] and excellent activity against Pseudomonas aeruginosa and Acinetobacter baumannii, including standard and clinical MDR strains [277].
Amino acid	Only a small number of key amino acids in AMPs are essential for antimicrobial activity and many other residues can be exchanged without

Optimization Strategy	Description
substitution	influencing their function. Substitution of neutral and non-polar amino acids with charged and polar amino acids enhances the amphipathicity, net charge, helicity and hydrophobicity of the AMPs which altogether lead to improved antimicrobial activity of the peptide. For example, R-BP100 which has all its Lys residues replaced with Arg resulted in a two-fold decrease in MIC against E. coli and a fivefold decrease in MIC against S. aureus [278].
Incorporation of unusual amino acids	Unusual amino acids are used to increase the activity or selectivity and plasma stability of peptides. Incorporation of d- or other chemically modified unusual amino acids alter the stereochemistry of the AMP and makes it unrecognizable by proteolytic enzymes. Thus, it reduces peptide susceptibility to proteolysis and increases stability, activity and selectivity. For example, pleurocidin with all-D-amino acids resist degradation by trypsin, plasmin, and carboxypeptidase [279].
Terminal/ side chain modifications	N-/C-terminal modification is a simple but effective way to enhance peptide stability and effectiveness *in vivo*. Most frequently used modifications include amidation (C-terminal), acetylation (N-terminal), and methylation (N-terminal). For instance, Hy-al (N-terminal modification) increased antimicrobial activity against *E. coli*, *S. aureus*, *P.aeruginosa* and *B. subtilis* [280]. PEGylation and glycosylation of amino acid side chain groups are also used to improve biocompatibility and bioavailability of peptides [281].
Nano-particle formulation	Conjugation to NPs reduces the cytotoxicity, improves the antimicrobial potency and in vivo stability of AMPs compared to their free form. As Nano-Particles are excellent target-specific drug delivery systems [282] they can also be used in conjugation with AMPs for their controlled release to the target site. Moreover, the large surface area to volume ratio of NPs makes them ideal candidates for loading AMPs [283]. As Gold NPs are inert and non-toxic in nature and have high biocompatibility, they are

Optimization Strategy	Description
	considered ideal candidates for loading AMPs. For example, Esc (1-21) attached to AuNPs exhibited almost zero toxicity toward human keratinocytes and exhibited better resistance to proteolytic digestion, also its antibacterial activity increased 15-fold against *P. aeruginosa* [284].
Combination Therapy	AMPs that show synergistic interactions with conventional antibiotics can be used as a component in antibiotic combination therapy, instead of as the single active ingredient (Table 2.16). Combining AMPs along with other antibiotics extend the lifetime of current antibiotics and overcomes the inadequacies of antibiotic monotherapy.

Table 2.17: List of AMPs that show synergy with antibiotics

AMPs	Source	Synergistic Molecule	Target	Ref
PGLa	Frog skin	Magainin 2	E. coli and S. aureus	[285]
Ranalexin	Bullfrog *R. catesbeiana* and *Staphylococcus simulans*	Endopeptidase lysostaphin	S. aureus (MRSA)	[286]
Tridecaptin M	Mud bacterium	Rifampicin, vancomycin, and ceftazidime	Extremely drug-resistant *A. baumannii*	[287]
Dermaseptin	Amphibians skin	Dermaseptin	*E. coli, P. aeruginosa, S. aureus*	[288]
Bactenecin	Lactic acid bacteria	Bactenecin	*E. coli, P. aeruginosa, S. Typhimurium*	[289]
Lactoferricin	Mammalians	Ciprofloxacin, ceftazidime	*P. aeruginosa*	[290]
Nisin	*Lactococcus lactis*	Colistin	*Pseudomonas* biofilms	[291]

AMPs	Source	Synergistic Molecule	Target	Ref
P10	---	Ceftazidim/doripenem	MDR *A. baumannii* and colistin-resistant *P. aeruginosa*	[292]
Gad-1	Fish	Kanamycin, ciprofloxacin	*P. aeruginosa*	[293]

2.12 Bioinformatic Tools for AMP Prediction

In the current situation of antibiotic resistance and urgent surge of new and effective antimicrobial agents, AMPs show a ray of hope to overcome this scarcity. For the high-throughput design and screening of AMPs for clinical use and in-vitro testing against pathogens its essential to acquire in-depth knowledge of the chemical structure, target microbe, antimicrobial/hemolytic/cytotoxic activities, the physicochemical parameters responsible for antimicrobial activity, the interlink between peptide structure and antimicrobial activity. To serve that purpose, using a sequence-based computational tool to predict AMP is essential for identifying novel antimicrobial agents. Many AMP Prediction tools have been developed based on machine learning algorithms and widely used for the prediction and detailed information of AMPs such as CAMPR3 [9], ADAM [294], AMPA [11], AMP Scanner [10], DBAASP [13] etc. (Table 2.18)

Table 2.18: Different AMP prediction tools

Database	Description	Algorithms	Ref
CAMPR3	This database contains information of family-specific signatures for a large set of eukaryotic and prokaryotic AMPs. AMPs belonging to a particular AMP family can be effortlessly obtained using the family-based search.	Support Vector Machines (SVM), Random Forests (RF) Artificial Neural Network (ANN) and Discriminant Analysis (DA)	[9]

Database	Description	Algorithms	Ref
ADAM	This database provides easy access to AMP sequences, structures, and their relations. Two distinct features of ADAM are its size and sequence-structure analysis. This is the first comprehensive study to analyze various AMP structural folds.	SVM: uses amino acid composition as features 2.HMM: searches the AMP homologs	[294]
AMP Scanner	The server calculates the physicochemical properties and designed sequence motifs of AMPs and converts them to numerical vectors to classify it as an AMP or Non-AMP. Sequences classified as an AMP are further submitted to the 'Selectivity Model'.	Numerical matrix from deep neural network (DNN): to identify if a sequence is "AMP" or "Non-AMP"	[10]
AMPA	This database can automatically scan protein sequences to identify potential antimicrobial regions and classify proteins or domains thereof as either antimicrobial or non-antimicrobial. Also, AMPA-derived AI values can be used to compare different protein sequences and can serve as new templates for AMP design.	Antimicrobial propensity scale threshold: to generate antimicrobial profile through a sliding window system	[11]
DBAASP	It is a comprehensive resource for structure–activity studies and the de novo design of AMPs with desired biological functions. It provides users with complete information on the chemical and 3D structure of peptides. It's sequence-based prediction services allow end-users to perform de novo design of peptides with activity against particular microbial strains.	Hydrophobicity scale	[13]

Chapter 3. Metagenomics

3.1 Definition

Metagenome refers to the total genomes of bacteria and fungi in environmental samples. Metagenomics, a term first coined by Handelsman in 1998 [295], is the study of the total genetic material collected from a mixed community of organisms. It is a culture-independent method of analyzing unculturable microorganisms present in environmental samples [296]. Conventional microbial culture techniques used in the laboratory are only able to grow a few numbers of microorganisms (only 1%) [297] as most of the environmental parameters are almost impossible to mimic inside a lab. The atmosphere artificially created inside the laboratories for routine culture is not sufficient to promote the growth of most of the microbes and this insufficiency keeps a great number of novel genes, metabolic pathways and potentially valuable metabolites undiscovered. Therefore, Metagenomic provides a way of exploring the undiscovered diverse microbial communities. It also offers an effective way to study population structure, ecological roles, evolution and genetic diversity of microbial communities without the need of culturing them inside a laboratory. This technique also has wide applications in extracting novel enzymes of industrial values and bioactive compounds containing high therapeutic index from extreme environments.

3.2 Types

Metagenomic approaches can be broadly classified into two groups: Amplicon sequencing and metagenomic shotgun sequencing. Amplicon sequencing is a fast, accurate, and cost-effective way of metataxonomic detection that specifically amplifies highly conservative and variable regions of the genome like 16S, 18S, and ITS. This sequencing is widely used in phylogenetic and taxonomic studies of bacteria and archaea. Unlike amplicon-based sequencing, Metagenomic Shotgun Sequencing involves random shearing of DNA similar to the pattern of a ‘shotgun’ and sequences all DNA fragments present in the genome in a single run. As this approach includes sequencing the entire genome, reads generated from taxonomically informative regions such as 16S, 18S, and ITS1/2 helps to provide insight

into the biodiversity while reads generated from coding sequences provide insight into the biological functions encoded by the genome. A short comparison between the two metagenomic approaches are listen in the table below (Table 3.1)

Table 3.1: Comparison of Amplicon sequencing and Metagenomic Shotgun sequencing (adapted from [298])

Factor	Amplicon sequencing	Shotgun Metagenomic Sequencing
Principle	Oligonucleotide probes are designed to target and capture hypervariable regions of conserved genes or intergenic regions, followed by next-generation sequencing to provide genetic information of these amplified sequences.	It involves randomly shearing the DNA of the microbial genome into small fragments, addition of a universal primer at both ends of the fragments for PCR amplification and sequencing. Lastly, splicing the small fragments into a longer sequence is done via assembly.
Research Objectives	To analyze the phylogenetic relationship of species, the species composition, and the biodiversity of a microbial community.	To conduct in-depth research on genes and functions of a microbial community, such as pathway analysis, KEGG, GO, *etc.*
Taxonomic coverage	Bacteria and archaea	All taxa, including viruses
Taxonomic resolution	Genus-Species	Species-Strains
Bioinformatics requirements	Beginner to intermediate expertise	Intermediate to advanced expertise
Databases	Established, well-curated	Relatively new, still growing
Sensitivity to host DNA contamination	Low	High
Bias	Medium to high	Lower
Cost	Low	High

3.3 Workflow of Metagenomic Sequencing:

The basic steps of metagenomic sequencing include DNA extraction from environmental samples, library preparation, sequencing, assembly, gene prediction, functional annotation and statistical analysis. Each of these steps are briefly discussed below:

3.3.1 DNA Extraction

The first step of metagenomic workflow is extracting DNA from environmental samples in high concentration and large fragments. DNA isolation involves three basic steps: cell lysis to release the nucleic acids from the cells, DNA isolation and DNA quantification and quality assessment. Cell lysis can be performed by physical, enzymatic, chemical methods, or a combination thereof. DNA can then be isolated via filtration, precipitation, and centrifugation. Also, DNA extraction can be performed using several commercial Meta-Genomic DNA isolation kits available on the market such as the FastPrep DNA.

3.3.2 Library preparation

The first step of DNA library preparation is shearing the DNA via mechanical or enzymatic process to generate uniform sizes of DNA. In the next step, adapter ligation is done by a ligating adaptor to the 5' or 3' end of the DNA fragment in order to attach the DNA fragments to a flow cell or sequencing platform. The adapter can harbor barcodes to aid in sample identification. The resulting DNA fragments are isolated based on size by using gel electrophoresis or bead-based mechanisms. The last step is the quantification and quality assessment of the prepared library which is often done using a BioAnalyzer System or through qPCR. Instead, library preparation kits like Bioo Scientific NEXTflex PCR-Free DNA Sequencing Kit or Illumina TruSeq PCR-Free Library Preparation Kit can be used for metagenomics library construction.

3.3.3 Sequencing

Various high-throughput sequencing techniques like sanger sequencing, whole genome shotgun sequencing, NGS (Next generation sequencing) are available and used depending

on the purpose of sequencing and the amount of data to be sequenced. Over the past few years Next Generation Sequencing (NSG) techniques have become more popular and are replacing Sanger Sequencing because the Sanger method only sequences a single DNA fragment at a time whereas NGS is able to sequence millions of fragments simultaneously per run. Among the different NGS technologies, the 454/Roche and the Illumina/Solexa systems are extensively applied to metagenomic samples.

3.3.4 Assembly

Sequencing generates short DNA segments called reads that are insufficient to decode all the information about the organisms present in a sample. Longer stretches of sequence provides more information about genetic variations and repeats, and helps to recover the genomes of the microbes present in a sample. Assembly is the process of stitching together short reads to generate longer genomic contigs. There are two types of assembly: *de novo* assembly and reference-based assembly (co-assembly). Various software like AMOS, Newbler (Roche) and MIRA are used in reference-based assembly. On the other hand, Brujin assemblers, a software particularly designed to handle very large amounts of data, is used for *De novo* assembly.

3.3.5 ORF/Gene prediction

Gene prediction is a fundamental step for finding genes, determining their function and annotating them properly. There are two computational methods for prediction: homology-based prediction and *Ab initio* prediction. Homology based prediction tools use local alignment and global alignment to find sequence similarity between known genes, ESTs (expressed sequence tags) or proteins and the input sequence. Major drawback of this approach is that it is unable to predict novel genes that have no significant homology to known genes. On the other hand, *Ab initio* prediction uses gene structure as a template to detect unknown genes and is able to predict genes having no sequence similarity to other genes.

3.3.6 Functional Annotation

Functional annotation is very important to explore the abundance of information that is reserved in genes. Function annotation of genes can be performed on either nucleotide or translated sequences. Homology detection is the easiest and most frequently used method for functional annotation. In this method, algorithms like BLAST are used to find sequence homology against well-curated databases of proteins and conserved domains like NCBI-nr or Swiss-Prot. If protein function cannot be assigned by homology due to low sequence identity values, HMM searches are used instead. HMM uses databases like the Conserved Domain Database of NCBI, SEED or PFAM for interrogating protein functional domain profiles.

3.3.7 Binning/Taxonomic Analysis:

Binning refers to the process of taxonomical and phylogenetic sorting of DNA sequence into bins (FASTA files that contain specific contigs) that represent an individual genome or genomes from closely related organisms. Similarity-based binning and compositional binning are two approaches combinedly used for complete binning. Similarity-based binning gives information about the apparent taxonomic position of the source population. BLAST and hidden Markov models are examples of this binning. Compositional approach is applied for binning contigs with genes having no homology with the reference species. Tetranucleotide frequencies interpolated Markov models and Markov chain Monte Carlo models are examples of compositional binning.

3.4 Applications of Metagenomics:

Metagenomics has a wide range of applications in various fields including screening of industrially relevant enzymes and clinically important antimicrobials, detection of pathogenic microorganisms from different clinical samples and bioremediation of environmental pollution. The applications of metagenomics in different sectors are briefly discussed below:

3.4.1 Industrial Applications:

Metagenomics holds great potential in discovering novel industrially important enzymes as its high-throughput screening technologies provide the opportunity to study a wide range of uncultured microorganisms and obtain new genetic information on industrial enzymes. With the help of metagenomic approaches, many industrially important enzymes like proteases, amylases, lipases and nitrilases have been isolated by the activity-based and functional screening of bacteria (Table 3.2).

Table 3.2: List of enzymes isolated via screening different metagenomic samples (adapted from [299])

Enzyme	Method	Host	Metagenomic sample	Ref
α-amylase	Function-based screening	*E. coli*	Deep sea and acid soil	[300]
Esterase	Functional screening	*E. coli*	Rumen of dairy cow	[301]
β -agarase	Activity-based screening	*E. coli*	Soil	[302]
β-glucanase	Function-based screening	*E. coli*	Large bowel of mouse	[303]
Cellulase	Function-based screening	*E. coli*	Soil	[304]
Glycosyl hydrolase	Functional screening	*E. coli*	Cow rumen fluid	[305]
β -galactosidase	Function-based screening	*E. coli*	----------	[306]
Xylanase	Function-based screening	*E. coli*	China Holstein cow rumen	[307]
Alkaline serine protease	Activity-based screening	*E. coli*	Goat skin surface	[308]
Alcohol oxidoreductase	Function-based screening	*E. coli*	Soil	[309]

Enzyme	Method	Host	Metagenomic sample	Ref
Amidase	Function-based screening	*E. coli*	Activated sludge	[310]
Amylase	Function-based screening	*E. coli*	Soil	
lipase	Activity based screening	*E. coli*	Forest soil	[311]
Protease	Function-based screening	*E. coli*	Soil	[312]
Chitinase	Function-based screening	*E. coli*	Seawater	[313]
Dehydratase	Function-based screening	*E. coli*	Soil	[314]
Carboxylic ester hydrolase	Activity-based screening	*E. coli*	Forest soil	[315]
Nitrilase	Function-based screening	*E. coli*	Soil, water	[316]
Tannase	Activity-based screening	*E. coli*	Cotton field soil	[317]
Fibrinolytic metalloprotease	Activity-based screening	*E. coli*	Mud	[318]
Glycotransferase	Functional screening	*E. coli*	Elephant feces	[319]

3.4.2 Discovering Bioactive Compounds:

Various bioactive compounds with antimicrobial, immunosuppressive or anti-inflammatory properties are produced inside microorganisms as primary or secondary metabolites. With the help of metagenomic techniques these bioactive compounds can be obtained and used in food safety maintenance strategies or pharmaceutical industries (Table 3.3). For example, production of biotin for industrial purposes has been possible through screening of metagenomic libraries.

Table 3.3: List of bioactive genes and pathways discovered through metagenomic screening (adapted from [299])

Bioactive/gene/pathway	Method	Host	Metagenomic sample	Ref
Pederin	Targeted sequencing		*Paederus beetles*	[320]
Biotin	Selection based screening	*E. coli*	Horse excrement	[321]
vibrioferrin	Function-based screening	*E. coli*	Tidal-flat sediment	[322]
Polyketide synthase gene	Targeted sequencing		*Discodermia dissoluta*	[323]
Novel serine protease inhibitor gene	Sequence based screening	*E. coli*	Uncultured marine organism	[324]
Borregomycin A and B	Homology based screening		Desert soil	[325]
Novel prebiotic degradation pathways	Hydrolytic activity-based selective screening	*E. coli*	Human ilium mucosa and fecal microbiota	[326]
Novel salt tolerant genes	Function-based screening	*E. coli*	Human gut microbiota and fecal sample	[327]
15 Acid resistant genes	Function-based screening	*E. coli*	Planktonic and rhizosphere microbiota	[328]

3.4.3 Discovery of Novel Antibiotics

Antimicrobials are produced in microbes as a defense mechanism to hinder competitor microorganism's growth. With the help of metagenomics, it is possible to identify novel antimicrobial molecules that can be identified by screening microbial populations for antimicrobial activity against indicator or clinically relevant microorganisms. These novel antimicrobials can be isolated with the help of high-throughput metagenomic tools and

clinically used against pathogens. Analyzing soil metagenomic samples has led to the discovery of several novel antibiotics like turbomycin A and B which is effective against both gram-positive and gram-negative bacteria (Table 3.4)

Table 3.4: List of Novel antimicrobials discovered from soil metagenomic samples (adapted from [329])

Antibiotics	Habitat	Library type	Ref
Beta-lactamases	Soil	Fosmid	[330]
Fasamycin A and B	Soil	Cosmid	[331]
Indirubin	Soil	Fosmid	[332]
Terragine	Soil	Cosmid	[333]
Turbomycin A and B	Soil	BAC	[334]
Violacein	Soil	Cosmid	[335]

3.4.4 Detection of Antibiotic Resistance Genes

In order to stop bacteria from becoming multidrug resistant it is very important to identify the genes causing bacterial resistance, understand bacterial mechanisms of developing resistance and tracing how resistance is passed to the progeny bacteria. Analyzing different metagenomic samples has made it possible to identify antibiotic resistance genes (Table 3.5) which might help scientists to deal with resistant bacterial strains.

Table 3.5: List of antibiotic resistance genes detected via metagenomic screening (adapted from [299])

Antibiotic resistance determinant	Screening Method	Host	Metagenomic sample	Ref
Novel florfenicol and chloramphenicol resistance gene	Functional screening	*E. coli*	Island soil	[336]
Novel kanamycin and ceftazidime resistant gene	Activity-based screening of	*E. coli*	Soil from apple orchard	[337]

Antibiotic resistance determinant	Screening Method	Host	Metagenomic sample	Ref
Chloramphenicol, ampicillin and kanamycin resistance gene	Functional screening	*E. coli*	Activated sludge	[338]
Novel chloramphenicol hydrolase (resistant to chloramphenicol and florfenicol)	Activity-based screening	*E. coli*	Alluvial soil	[339]
Novel carboxylesterase	Activity-based screening	*E. coli*	Wetland soil	[340]
Ampicillin, amoxicillin, tetracycline, penicillin and class A and C β-lactamase resistant genes	Activity-based screening	*E. coli*	Fecal sample of Herring gull	[341]
Kanamycin, gentamicin, chloramphenicol, rifampin, trimethoprim and tetracycline resistant genes	Activity-based screening	*E. coli*	Urban soil	[342]
β-lactam, aminoglycoside, amphenicol, sulfonamide and tetracycline resistant genes	Activity-based screening	*E. coli*	soil	[343]
Tetracycline, minocycline, aminoglycoside, streptomycin, gentamicin, kanamycin, amikacin, chloramphenicol and rifampicin resistant genes	Activity-based screening	*E. coli*	agricultural soil sample	[344]
Fluoroquinolone, cephalosporin and trimethoprim resistant genes	Activity-based screening	*E. coli*	Retail spinach	[345]
Ampicillin and kanamycin resistant genes	Activity-based screening	*E. coli*	Mozzarella Cheese	[346]

3.4.5 Diagnosis of Infectious Disease:

Metagenomic next-generation sequencing (mNGS) makes it possible to identify multiple pathogens simultaneously with a single assay making the diagnosis process fast and easy. Characterization of microorganisms from various clinical samples (Table 3.6) like saliva, blood, cough, urine and tissue fluid helps in the analysis of human host response (transcriptomic) data to predict causes of infection and evaluate disease risk. Most recently, mNGS has also been applied for the rapid identification of SARS-CoV-2b [347].

Table 3.6: Clinical diagnosis of pathogenic microbes through metagenomic sequencing (adapted from [348])

Sample	Metagenomic Analysis	Diagnosis	Ref
Stool	Shotgun pyrosequencing (Roche 454)	*Clostridium jejuni*	[349]
CSF	Shotgun sequencing (Illumina).	*Leptospira*	[350]
CSF	Shotgun sequencing (Illumina)	*Brucella*	[351]
Brain biopsy	Shotgun sequencing (Illumina)	*M. tuberculosis*	[352]
CSF	Shotgun sequencing (Ion Torrent)	*L. monocytogenes*	[353]
Articular fluid	Shotgun sequencing (Illumina)	*Mycoplasma salivarium*	[354]
Cardiac valve	Shotgun sequencing (Illumina)	*Abiotrophia defective*	[355]
Plasma	Shotgun sequencing (Illumina)	*Capnocytophaga canimorsus*	[356]
Blood	Shotgun sequencing (Ion)	*Propionibacterium acnes*	[357]
BAL	Shotgun sequencing (Ion)	*Pseudomonas aeruginosa; Staphylococcus aureus*	[358]

Bronchoalveolar lavage (BAL), Cerebrospinal fluid (CSF)

3.4.6 Bioremediation of Pollutants:

Metagenomics is considered as one of the potent tools to remove contaminants from the environment [359]. Metagenomic technologies can be used to obtain information about the

characteristics of bacterial communities present in various contaminated sites, screen potential microbial degraders for bioremediation and find out the corresponding genes for the degradation and detoxification. Recently, multiple studies conducting metagenomic approaches in bioremediation have been reported (Table 3.7).

Table 3.7: Metagenomic screening of bioremediating bacteria from polluted environment

Purpose	Metagenomic approach	Result	Ref
To evaluate the petroleum degradation abilities of metagenomic bacteria derived from petroleum reservoirs	Construction of metagenomic library	*Dietzia maris* CBMAI 705 and *Micrococcus sp.* CBMAI 636 could biodegrade up to 99% of phenanthrene and methyl phenanthrenes	[360]
Bioremediation of chromium-contaminated water	metagenomic analysis	*Enterobactercloacae*, *Flavobacterium sp.* and *Ralstonia sp.* reduced metal contamination up to 100%	[361]
To identify hydrocarbon degrading microbes and their corresponding genes active in cold temperature	Metagenomic sequencing	Pseudomonas and Rhodococcus species actively expressed hydrocarbon degradation genes at cold temperature	[362]
To analyze microbial community composition and diversity at a diesel-contaminated site	pyrosequencing	Proteobacteria, Firmicutes, Actinobacteria, Acidobacteria, and Chloroflexi found in abundance	[363]

Chapter 4. Meta-transcriptomics

4.1 Definition

Meta-transcriptomics, first introduced in the early 2000s deals with the study and analysis of the transcriptome (mRNA) of a metagenomic sample. The transcriptome consists of the total mRNA sequences extracted from an environmental sample. Metagenomic approaches only identify the genetic content of bacterial populations which is not enough to differentiate the active microbes from inactive members of a microbiome. On the other hand, meta-transcriptomics can reveal details about populations that are transcriptionally active by generating information on the real-time gene expression of that microbiome. From the transcriptome analysis of a given environmental sample, we can learn about the gene expression and functional roles of the complex microbial communities present in the environmental samples.

4.2 Applications

Meta-transcriptomics has vast application in characterizing active microbes in a community [345], discovering novel microbial interactions [346], detection of regulatory antisense RNA [347], tracking gene expressions and determining the relationship between viruses and their host [348]. Metatranscriptomics has been applied to various fields from the study of microbiomes found in humans or animals or plants, within soils, and in aquatic environments. For example, Marchetti et al. [364] and Mason et al. [365] sequenced the transcriptomes of ocean microbes to identify active members and their functional responses after environmental changes. Maurice et al. [366] conducted metatranscriptome profiling, 16S rRNA gene sequencing, and flow cytometry to identify dominant bacterial species in the human gut microbiota as well as the physiology and gene expression responses of bacteria to xenobiotics. In this article, we will specifically discuss the applications of meta transcriptomics in the marine environment.

4.3 Meta-transcriptomic analysis of Marine Sediment Microbiota

Marine ecosystem contains about 103–1010 microbe/cm3 sediment, consisting of about

0.18 to 3.6% of Earth's total living biomass [367]. It indeed is very rich in biodiversity and might be the biggest reservoir of metabolites containing huge pharmaceutical and industrial importance. However, most of the marine sediment microbiota remains undiscovered till date due to the inability to mimic their natural growth conditions inside laboratories, their slow growth rates, poor number of colonies, dependance on the metabolites generated by other microbes, and dormancy. As routine culture procedure is futile for studying these microbes, meta-transcriptomic analysis can be the best option for understanding their structural diversity, functional profiles and gene expression patterns. For instance, Metatranscriptomics has already been applied to characterize the dynamics of cyanobacterial blooms in the Baltic Sea [368], to detect small RNAs in the open ocean [369], and analyze viral-host relationships of marine eukaryotes [370]. Already, a good number of antimicrobials have been discovered from the marine sediment microbiota (listed in table 4.1). Hence, with the aim to discover more novel AMPs from the marine sediment microbiota, we collected some meta-transcriptomic data from a previous study based on the enrichment culture of marine sediment [371] and tried to predict AMP from those data via multiple AMP prediction tools. We further predicted the three-dimensional structures and protein parameters of the transcripts that resulted in AMP positive.

Table 4.1: List of bioactive compounds/drugs isolated from various marine sediment sample (adapted from [372])

Antimicrobial compounds	**Origin**	**Sediment Sample**	**Biological activity**	**Ref**
Lynamicins, spiroindimicins	*Streptomyces sp.*	Deep sea marine sediment	Antibacterial	[373]
Saccharomonopyrones	*Saccharomonospora sp.*	Deep sea sediment	Weak antioxidant	[374]
Strepchazolins	*Streptomyces chartreusis*	Marine sediments	Antibacterial	[375]
Bonnevillamides	*Streptomyces sp.*	Sediment	Modulate heart	[376]

Antimicrobial compounds	Origin	Sediment Sample	Biological activity	Ref
			growth and cardiac function	
Paulomycin G	*Micromonospora matsumotoense*	Deep sea marine sediment	Antitumor properties	[377]
Rifamycin B	*Salinispora sp.*	Sediment	Antibacterial	[378]
Ananstreps	*Streptomyces anandii*	Sea sediment at a mangrove	Antibacterial, Cytotoxic	[379]
Actinonin	*Streptomyces sp.*	Sediment	Antibacterial	[380]
Xiamenmycin	*Streptomyces xiamenensis*	Sediment	Anti-fibrotic	[381]
2-Methyl butyl propyl phthalate	*Streptomyces cheonan-ensis*	Sediment of mangrove ecosystem	Antibacterial, antifungal, cytotoxic	[382]
N-(4-minocyclooctyl)-3,5-dinitrobenzamide	*Pseudonocardia endophytica*	Sediment of mangrove ecosystem	Antibacterial, cytotoxic	[383]
Dinactin	*Streptomyces sp.*	marine sediments	Antiproliferative and antimalarial	[384]
N-acetyl-N-demethyl-mayamycin and Strep-toanthraquinone A	*Streptomyces sp.*	Marine sediments	Antiglioma and antibacterial	[385]
Violapyrone B	*Streptomyces somaliensis*	Deep sea marine sediment	Antibacterial	[386]
Akaeolide	*Streptomyces sp.*	Marine sediment	Antimicrobial	[386]
Marangucycline A	*Streptomyces sp.*	Deep sea marine sediment	Antibacterial and Cytotoxic	[387]
Isoikarugamycin	*Streptomyces*	Marine sediment	Antibacterial and	[388]

Antimicrobial compounds	Origin	Sediment Sample	Biological activity	Ref
	zhaozho-uensis		antifungal	
Bohemamine	*Streptomyces spinover-rucosus*	Marine sediment from mangrove	Antibacterial	[389]
Hormaomycins	*Streptomyces sp.*	Marine sediment	Antibacterial and antifungal	[390]
Lobosamides	*Micromonospora sp.*	Sediment	Antitrypanosome	[391]
Salinipostins	*Salinispora sp.*	Sediment	Antimalarial	[392]

Chapter 5. A Case Study of AMP Identification from Metatranscriptome

5.1 Overview of the Case Study

In this study, we designed a workflow for the *in silico* identification of AMPs from the metatranscriptome of uncultured marine sediment microbiota. We used a combination of machine-learning based AMP prediction tools that are designed to identify AMPs from a given proteome of an organism. The tools effectively identified the potential AMP candidates from the given data and we proceeded with their functional characterization and structure prediction. We also determined the physicochemical features of the potential AMP candidates, knowledge of which might be useful in case of their purification in laboratories. Through this study we wanted to demonstrate the potential use of available metatranscriptomic data to identify novel proteins/antimicrobials/bioactive. We hope to add value to the quest of discovering novel AMPs from different sources for designing drugs that are more resilient to antibiotic resistance and able to rescue the world from the ongoing crisis of multidrug resistance.

5.2 Rationale of the Case Study

We know that AMPs have a wide range of bactericidal, antifungal, antiviral and antiparasitic activities and they are able to act on even the severely drug-resistant bacterial strains. A wide array of research and trials are ongoing to identify novel AMPs from different organisms and produce AMP-based drugs. Among all the different microbial communities of the world, marine microbial communities are one of the richest producers of diverse bioactive compounds, yet they are among the least studied ones due to their unculturable nature in the artificial lab environment. Fortunately, the recent advances in machine-learning techniques and metagenomic sequencing have paved the way for discovering AMPs produced inside the unculturable organisms too. Till now, little research has been conducted on marine sediment microbiota, but the results of these studies have led to the discovery of a range of novel antimicrobials and bioactive (Table 4.1). Utilizing metagenomic/metatranscriptomic data for discovering AMP is a comparatively new approach, and using the metatranscriptome of marine microbiota only adds to the

uniqueness of this study. Thus, we chose metatranscriptome marine microbiota samples as our focus and proposed an *in silico* workflow for predicting AMP from the predicted proteins. We opted for a *in silico* approach for discovering AMP because it is faster, cost effective and more efficient than the *in vivo* or *in vitro* procedures of AMP identification. Our workflow can lead to the discovery of novel AMPs which can further be purified in the lab and tested on different pathogenic strains to design effective antimicrobial drugs.

5.3 Objective of the Case Study

The objectives of this study are as follows:

1. Prediction of AMPs from metatranscriptome data
2. Elimination of bias via filtering the potential AMP candidates through multiple tools
3. Prediction of the secondary and the tertiary structures of the AMP candidates
4. Functional characterization of the AMP candidates
5. Prediction of the physicochemical parameters of the AMP candidates
6. Establishment of a standard pipeline for AMP prediction

Chapter 6. Methodology

Firstly, metatranscriptomic data for our experiment was collected from MGnify database of EMBL-EBI [393]. Then the collected data was assessed on four AMP prediction tools namely CAMPR3, AMP scanner, ADAM, and AMPA for identifying reads that encode for potential AMPs. After that, the sequences that were predicted as AMP by at least three different predictors were filtered out from the rest to eliminate false positive results. The filtered AMP candidates were put forth for their functional characterization using InterProScan followed by the secondary and the tertiary structure prediction via PSIPRED and PEPFOLD3 tools respectively. Lastly, the important physiological parameters of the AMPs were predicted using the ProtParam tool. The basic workflow of our experiment is shown in the figure below:

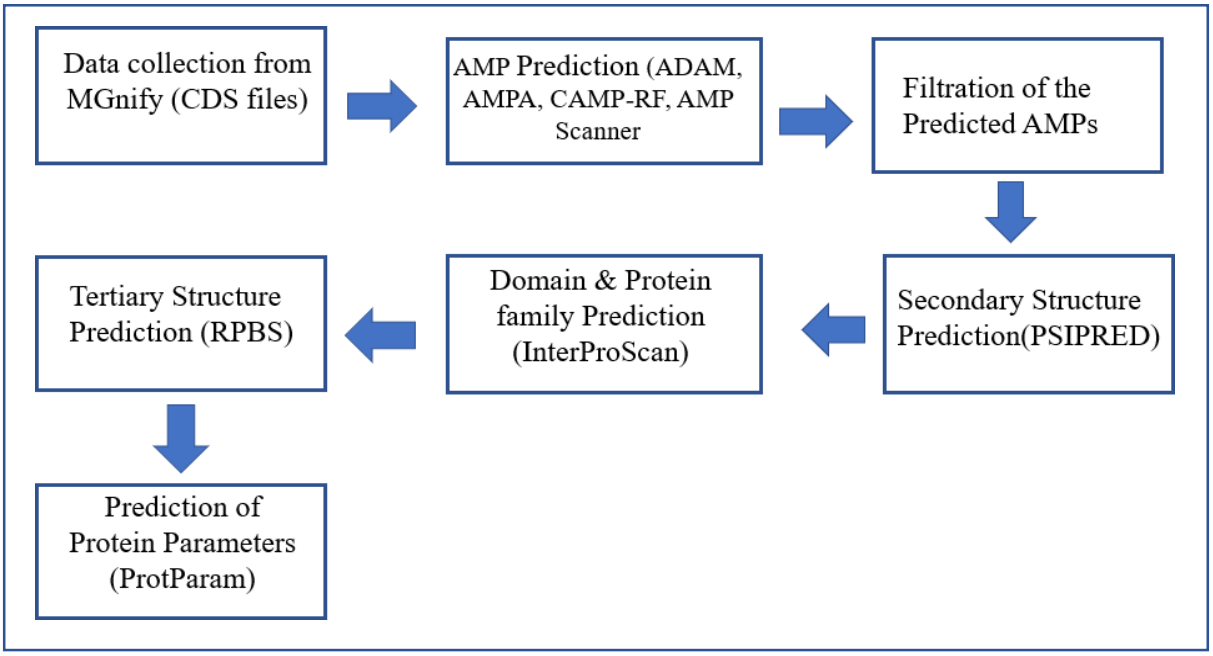


Figure 6.1: Overview of the entire workflow

6.1 Collection of Meta-transcriptomic Data

We collected metatranscriptomic data from the study named "Metatranscriptomic and comparative genomic insights into resuscitation mechanisms during enrichment culturing" [371] conducted by Mu *et al.* in 2018 and accessed the study data from MGnify [393], a

free website that provides an automated pipeline for analyzing and archiving microbiome data from metagenomic and meta-transcriptomic samples. The study contained three different marine sediment samples collected from the intertidal zone of Weihai, China and named as HGD, S and XSD. As the purpose of our experiment was to predict antimicrobial potency and characterize protein, we collected the CDS files of the samples present in the website. CDS (Coding Sequence) refers to the region of DNA that translates into a protein and the sequence of CDS determines the sequence of amino acids in a protein. In the website, each sample had five CDS files named HGD1, HGD2, HGD3 etc. We downloaded 15 "Predicted CDS without annotation" files in total from which we obtained 417 protein sequences (Table 6.1) in FASTA format and proceeded for the next step.

Table 6.1: Collection of metatranscriptomic data

Sl.	File name	Sample Name	Number of sequences
1	HGD0_cDNA_Predicted CDS without annotation	SRS2998334	17
2	HGD1_cDNA_Predicted CDS without annotation	SRS2998333	56
3	HGD2_cDNA_Predicted CDS without annotation	SRS2998336	28
4	HGD3_cDNA_Predicted CDS without annotation	SRS2998335	19
5	HGD4_cDNA_Predicted CDS without annotation	SRS2998338	15
6	S0_cDNA_Predicted CDS without annotation	SRS2998337	47
7	S1_cDNA_Predicted CDS without annotation	SRS2998340	41
8	S2_cDNA_Predicted CDS without annotation	SRS2998339	31
9	S3_cDNA_Predicted CDS without annotation	SRS2998331	28
10	S4_cDNA_Predicted CDS without annotation	SRS2998332	9
11	XSD0_cDNA_Predicted CDS without annotation	SRS2998341	24
12	XSD1_cDNA_Predicted CDS without annotation	SRS2998342	28
13	XSD2_cDNA_Predicted CDS without annotation	SRS2998344	20
14	XSD3_cDNA_Predicted CDS without annotation	SRS2998343	29
15	XSD4_cDNA_Predicted CDS without annotation	SRS2998345	25
Total number of sequences			417

6.2 Prediction of AMP

We ran all the 15 files in four different AMP prediction tools namely CAMPR3 [9], AMP Scanner [9], AMPA [11], and ADAM [294]. CAMPR3 had different tools e.g. SVM, ANN, RF and Discriminant Analysis, among which we chose CAMPR3(RF), since it was previously determined to be the one of the best-performing AMP prediction methods. The parameters used and interpretation of results are described in the Table below (Table 6.2). After getting the results for all files, we organized them in separate excel sheets.

Table 6.2: Search parameters and result interpretation of different AMP predictors

Tool	Search parameters	Input	Result interpretation
CAMPR3	Random Forest	FASTA sequence	A probability score (0 to 1) is given for the prediction. Higher the probability, greater the possibility of the peptide being antimicrobial
AMP Scanner (vr 2)	N/A	FASTA sequences (length should be between 10-200AA)	Sequence having a Probability > 0.5 is predicted as AMP
AMPA	Window size = '7' Threshold value = 0.225.	FASTA	The number and location of the antimicrobial stretches, the mean antimicrobial index of both the protein and the predicted stretches as well as the probability that the predicted stretch is found by chance in a non-antimicrobial protein are provided.
ADAM	HMM	FASTA	Generates result based on a probabilistic model where, the already known AMPs are used to find the most likely homologue for the query peptide. Sequences with > 50% probability are considered to be classified as AMP.

6.3 Filtration of the Predicted AMPs

Different predictors use different parameters and algorithms to predict AMP, for example ADAM and CAMPR3 are able to identify any variety of AMPs, whereas other tools can only predict specific subclasses of AMPs. This is why one particular protein sequence that is detected as AMP in one tool can be predicted as non-AMP by some other tool. To find out the AMP-coding sequences with 100% accuracy, we constructed five different filters by putting the four AMP predictors in different combinations (Table 6.3) and selected only those sequences that were predicted as AMP-positive by at least three of the four different prediction tools.

Table 6.3: Filters constructed for FASTA sequences

Sl.	Description
Filter 1	Positive result in CAMPR3 (RF), ADAM, AMPA and AMP Scanner
Filter 2	Positive result in CAMPR3 (RF), AMP Scanner and ADAM
Filter 3	Positive result in ADAM, AMPA and AMP Scanner
Filter 4	Positive result in CAMPR3 (RF), AMP Scanner and AMPA
Filter 5	Positive result in CAMPR3 (RF), ADAM and AMPA

6.4 Protein Characterization

Firstly, we submitted the FASTA sequence of the proteins positive in the above-mentioned filters to PSIPRED [394, 395] to predict the secondary structure. Later, we searched for the protein domains and family memberships of our AMPs with InterProScan [396]. InterProScan provides functional analysis of proteins by classifying them into families and predicting domains and important sites. To classify proteins in this way, InterProScan uses predictive models, known as signatures, provided by several different databases (referred to as member databases). It scans query protein sequences against the protein signatures of the InterPro member databases to find out if the given protein belongs to any of the protein families in the database and if yes then which protein family they belong to. It is very important to know the protein-family because the proteins within a particular family tend to interact with certain molecules in similar ways and perform similar functions within the cell. So, if we could trace the family of a protein, we can get a clear idea about its functions

and the way it interacts with other compounds. After that, the tertiary structure visualization was done using PEP-FOLD3 online tool [397-399]. Fourthly, we submitted our peptides' amino acid sequences as one-letter codes in the ProtParam software [400] for the calculation of various physical and chemical parameters including the molecular weight, theoretical pI, amino acid composition, atomic composition, estimated half-life, instability index and grand average of hydropathicity (GRAVY).

Chapter 7. Results

We have divided our results into three sections. Section one contains the results obtained from different AMP prediction tools, section two contains the filtered AMPs, and the third section contains the results of the protein characterization of the AMPs.

7.1 Result of AMP Prediction

AMP prediction results obtained from the four AMP prediction tools for each of the three-sediment samples (HGD, S, XSD) are organized in separate Tables and kept as supplementary material. Some of the FASTA sequences contained unusual amino acids in them, thus CAMPR3 could neither detect nor predict AMP from those particular sequences. Sequences showing positive results in AMP prediction tools are labeled as AMP and highlighted in yellow, sequences showing negative results in the prediction tools are labeled as NAMP (Non-AMP) and sequences showing no results in the prediction tools are kept blank in the Tables (Supplementary material). As different predictors are based on different machine learning algorithms, a sequence which is predicted as AMP by one predictor gets predicted as non-AMP by another predictor. Thus, we found different numbers of AMP-coding regions for the same sequences in different predictors (Table 7.16).

Table 7.1: AMPs prediction result of the AMP predictors

Sample Name	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
HGD0	3	0	2	1
HGD1	5	9	10	22
HGD2	3	2	5	2
HGD3	3	3	4	0
HGD4	3	3	2	1
S0	9	0	5	10
S1	9	3	6	11
S2	10	8	5	10
S3	7	5	5	6

Sample Name	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
S4	2	4	1	0
XSD0	6	2	2	12
XSD1	5	4	3	10
XSD2	2	4	5	6
XSD3	9	3	5	1
XSD3	9	6	4	2

7.2 Result of AMP Filtration:

The sequences that showed positive results in any of the filters mentioned below are named according to their sequence number in the source CDS file, for example HGD1-P50 means that the 50th sequence of the HGD1 FASTA file was predicted AMP by the predictor software. Only HGD1-P50, S1-P8 and S2-P15 have been predicted as AMP by all the prediction tools (Table 7.17; Figure 7.1), rest of the sequences placed in the Table below are predicted as AMP by at least three of the four different predictors used. By applying five different filters, we got 22 AMP-coding sequences from our experiment in total where three sequences are obtained from filter-1, three sequences from filter-2, seven sequences from filter-3 and nine sequences from filter-4 but no sequences from filter-5. (Table 7.17). Only the sequences that were successful in generating positive results in any of the set of predictors are considered as true AMPs and analyzed further.

Table 7.2: Filtration result of the predicted AMP-coding sequences

Sample Name	Filter 1	Filter 2	Filter 3	Filter 4	Filter 5
HGD1	1(P50)	1(P38)	1(P33)	0	0
HGD3	0	0	0	1(P6)	0
HGD4	0	0	0	2(P5, P8)	0
S1	1(P8)	0	1(P34)	0	0
S2	1(P15)	0	1(P2)	2 (P26, P31)	0
S3	0	0	1(P8)	1(P15)	0

Sample Name	Filter 1	Filter 2	Filter 3	Filter 4	Filter 5
XSD1	0	0	1(P11)	1(P10)	0
XSD2	0	0	2(P9, P10)	1(P19)	0
XSD3	0	0	0	1(P21)	0
XSD4	0	2(P5, P20)	0	0	0
Total	3	3	7	9	0

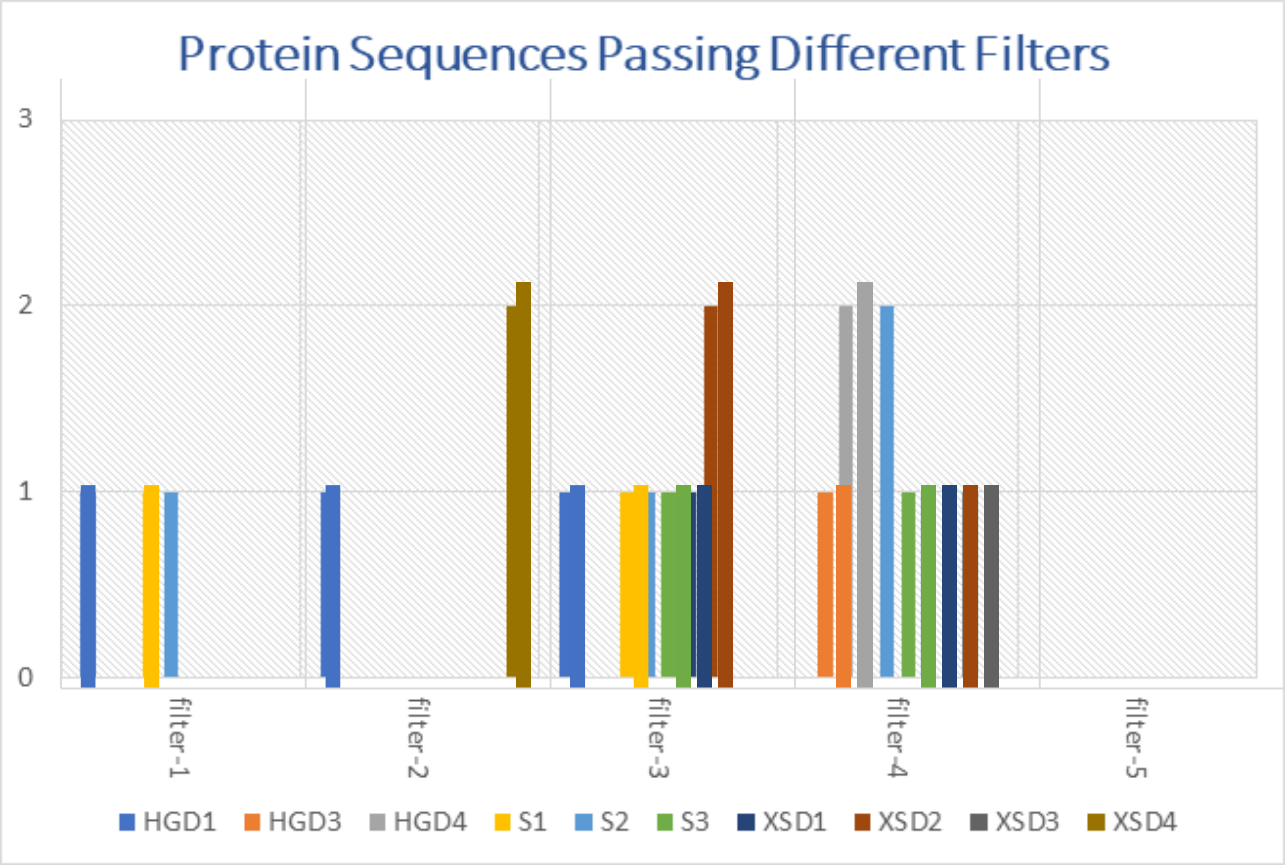


Figure 7.1: Number of Protein sequence that successfully passed the filters

7.3 Result of Protein Characterization

A total of 22 AMPs successfully passed the filters, and these were further analyzed for structure prediction. The secondary structure predicted by PSIPRED shows the number

and position of β-strands (highlighted in yellow), α-helix (highlighted in pink) and coils (highlighted in gray) in a given protein. The secondary structure of the proteins containing lesser than thirty amino acids were unable to be predicted by this software. InterProScan could not generate any information about the protein domain or protein family membership of any of the submitted sequence as all of them were very small proteins. The tertiary structures of the proteins were visualized via PEP-FOLD3. When we compared both the secondary and tertiary structures of the peptides where we found that most of the proteins had differences in the secondary and tertiary structures predicted by the two software. For example, the protein HGD1-P33 had one β-strand and one α-helix according to the secondary structure predicted by PSIPRED, however, the tertiary structure of the same protein predicted by PEP-FOLD3 did not have any β-strand but contained two α-helixes instead of one (Figure 5.2). In the cases of dissimilarities between the two results, we considered the three-dimensional structure predicted by PEP-FOLD3 to be more accurate. ProtParam identified the different parameters of our query protein sequences like the amino acid length, molecular weight, theoretical pI, estimated half-life in Mammalian reticulocytes (*in vitro*), instability index, and grand average of hydropathicity (GRAVY).

7.3.1 Structure Prediction of the AMPs

HGD1-P33: We found that the secondary structure of HGD1-P33 contains one β-strand and one α-helix. On the contrary, its tertiary structure contains two α-helixes only (Fig 7.2).

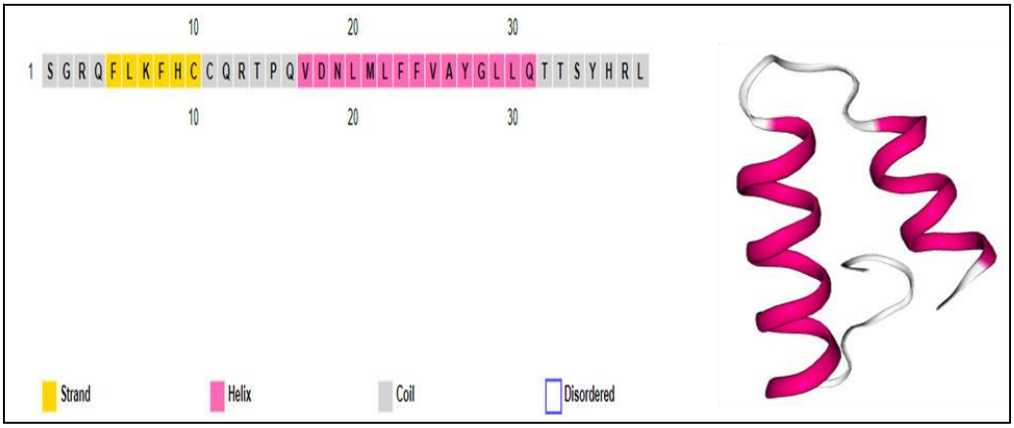


Figure 7.2: The secondary and tertiary structure of HGD1-P33

HGD1-P38: We saw that, both the secondary and the tertiary structure of HGD1-P33 contains two β-strands only (Figure 7.3).

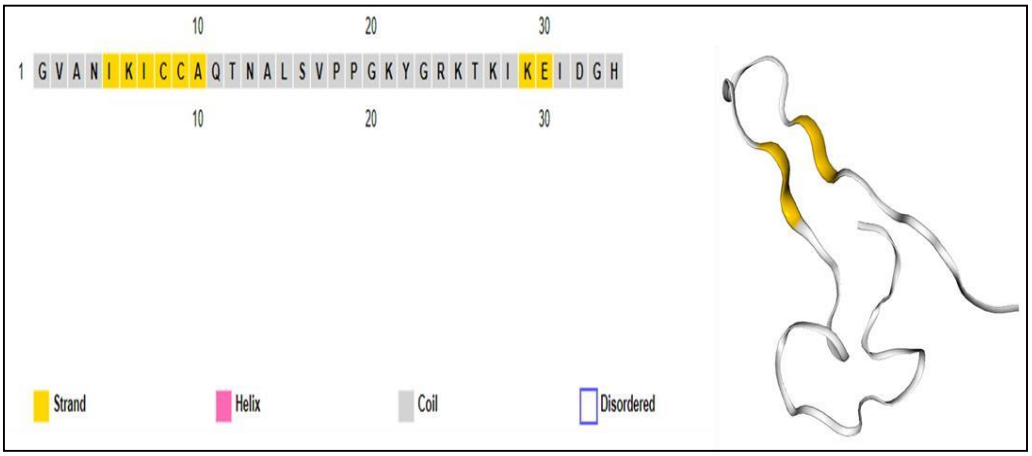


Figure 7.3: The secondary and tertiary structure of HGD1-P38

HGD1-P50: We found that the secondary structure of HGD1-P50 comprises two β-strands and one α-helix, but its tertiary structure contains only one α-helix (Figure 7.4).

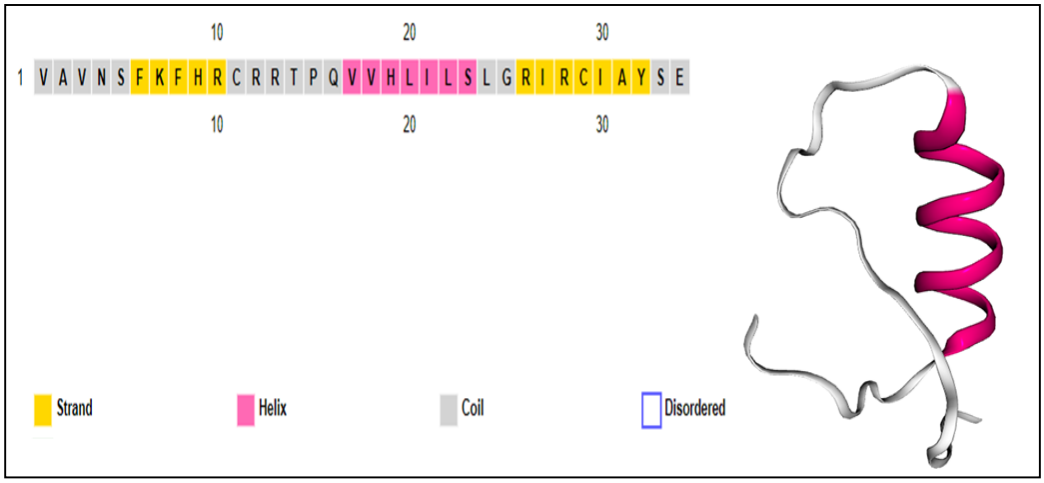


Figure 7.4: The secondary and tertiary structure of HGD1-P50

HGD3-P6: We found that the secondary structure of HGD3-P6 contains two β-strands and two α-helixes, but its tertiary structure consists of two β-strands and one α-helix (Figure 7.5).

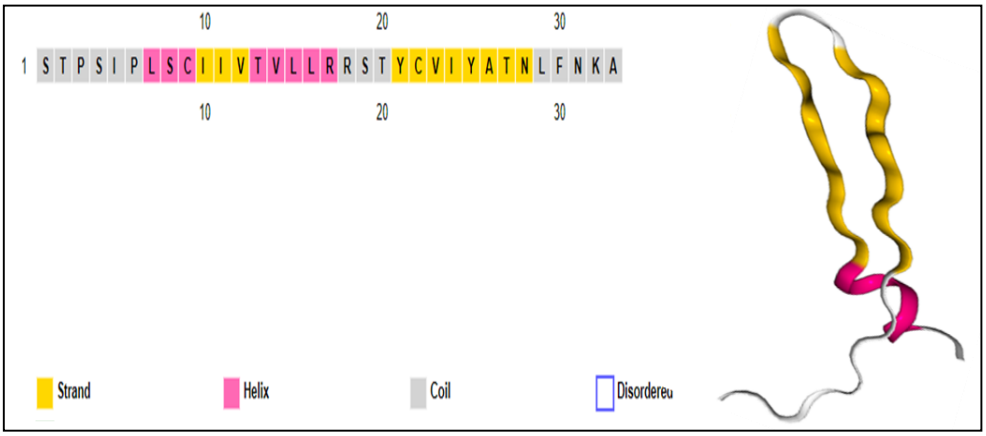


Figure 7.5: The secondary and tertiary structure of HGD3-P6

HGD4-P5: We found that the secondary structure of HGD4-P5 contains two β-strands and one α-helix, but its tertiary structure consists of two α-helixes (Figure 7.6).

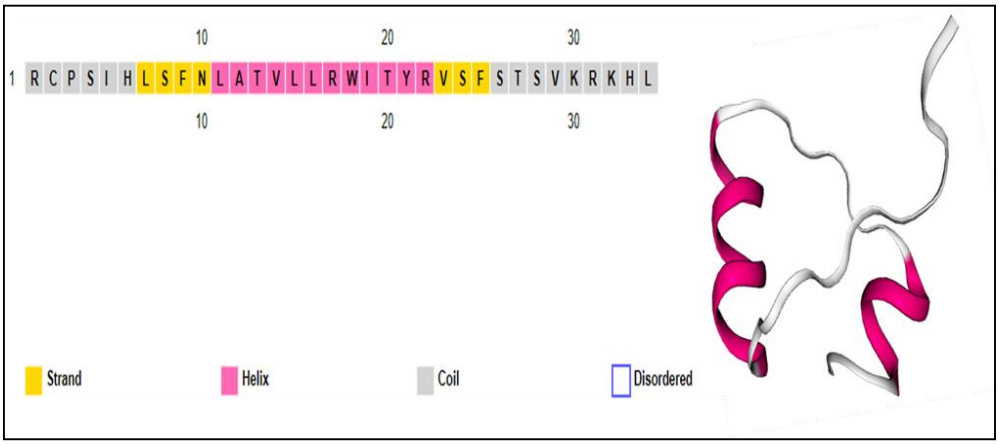


Figure 7.6: The secondary and tertiary structure of HGD4-P5

HGD4-P8: We found that, both the secondary and the tertiary structure of HGD4-P8 contains two β-strands and one α-helix (Figure 7.7).

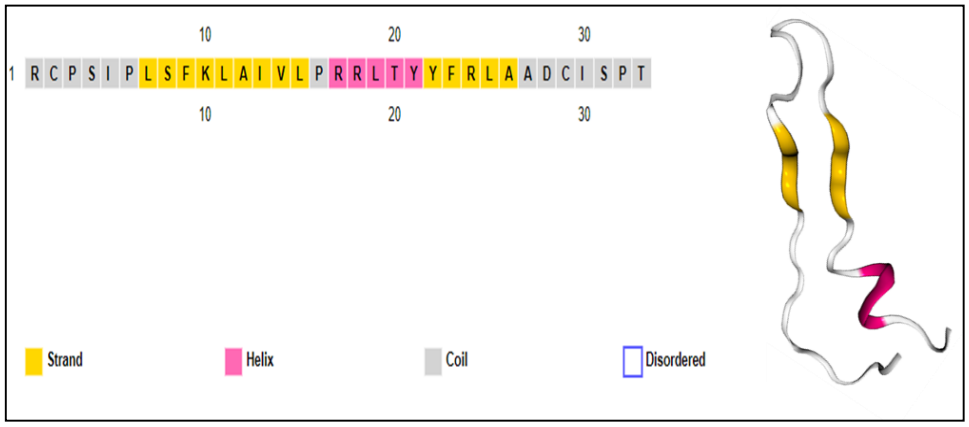


Figure 7.7: The secondary and tertiary structure of HGD4-P8

S1-P8: We found that the secondary structure of S1-P8 contains two β-strands and one α-helix but its tertiary structure consists of two β-strands only (Figure 7.8).

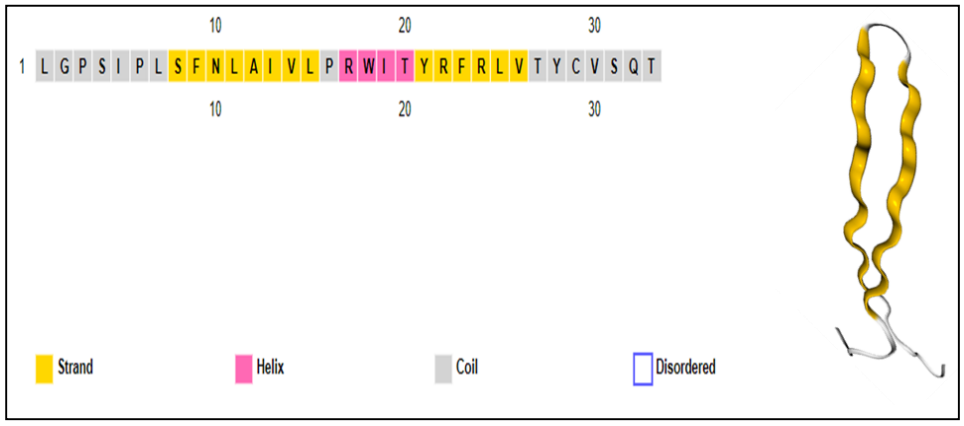


Figure 7.8: The secondary and tertiary structure of S1-P8

S1-P34: We found that the secondary structure of S1-P34 contains one β-strand and one α-helix, but its tertiary structure contains two α-helixes (Figure 7.9).

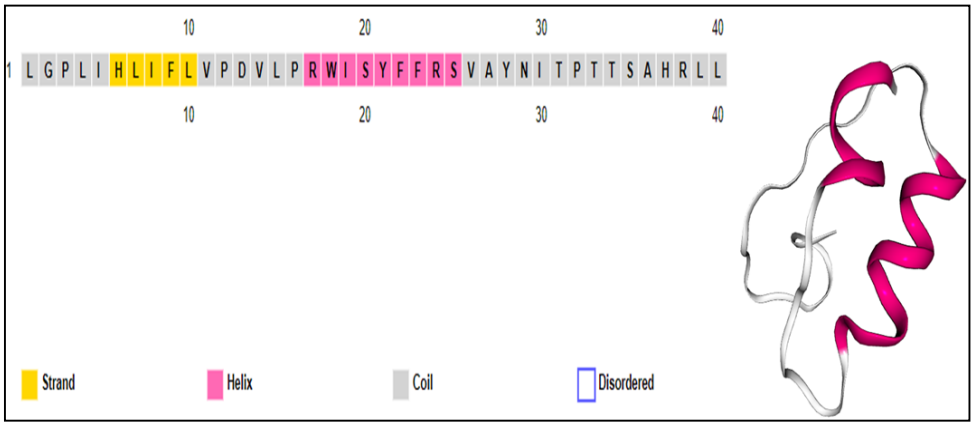


Figure 7.9: The secondary and tertiary structure of S1-P34

S2-P2: We found that the secondary structure of S2-P2 contains two β-strands and one α-helix but its tertiary structure consists of two α-helixes (Figure 7.10).

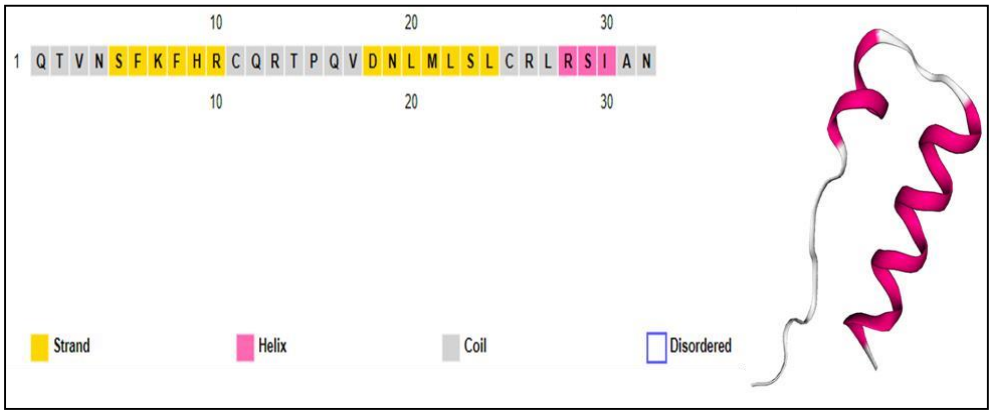


Figure 7.10: The secondary and tertiary structure of S2-P2

S2-P15: We found that, both the secondary and the tertiary structures of S2-P15 contain two β-strands and one α-helix (Figure 7.11).

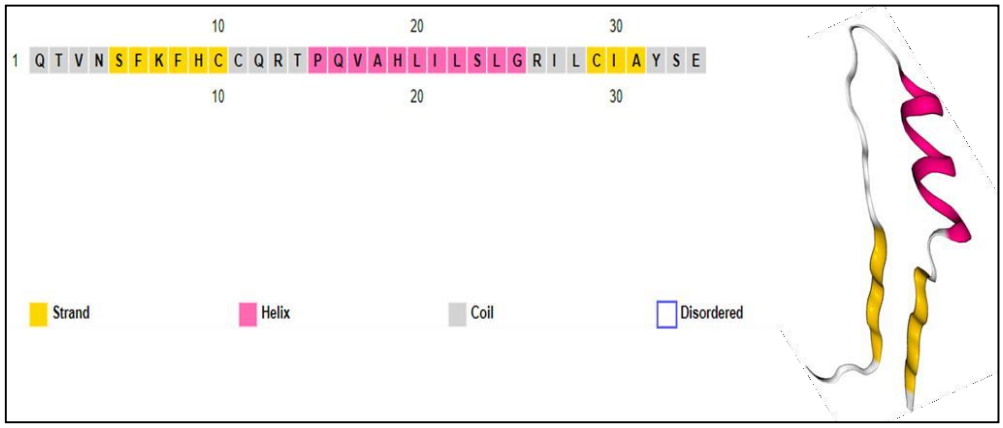


Figure 7.11: The secondary and tertiary structure of S2-P15

S2-P26: We found that the secondary structure of S2-P26 could not be predicted as its amino acid length was below 30. However, its tertiary structure consists of two α-helixes (Figure 7.12).

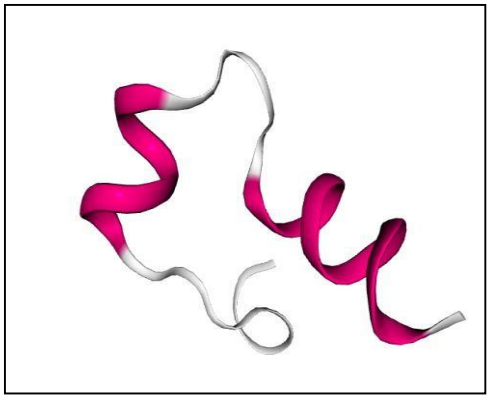


Figure 7.12: The tertiary structure of S2-P26

S2-P31: We found that the secondary structure of S2-P31 contains one β-strand and two α-helixes, but its tertiary structure consists of two α-helixes (Figure 7.13).

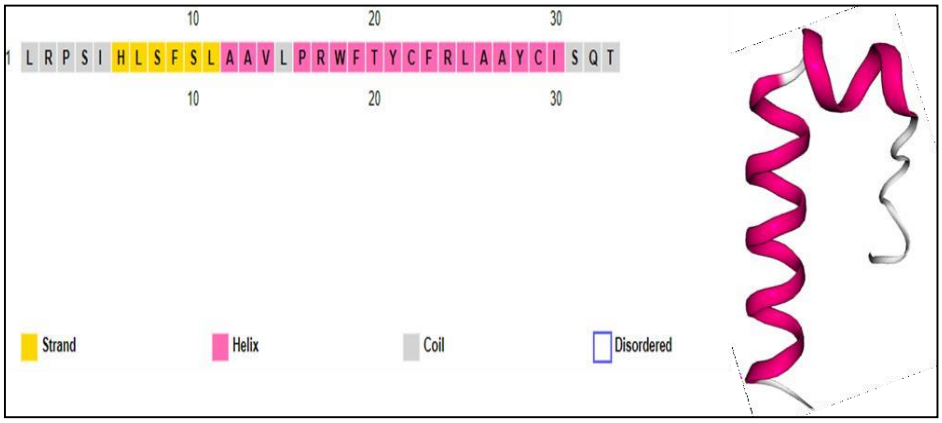


Figure 7.13: The secondary and tertiary structure of S2-P31

S3-P8: We found that the secondary structure of S2-P31 contains two β-strands and two α-helixes but its tertiary structure consists of two β-strands (Figure 7.14).

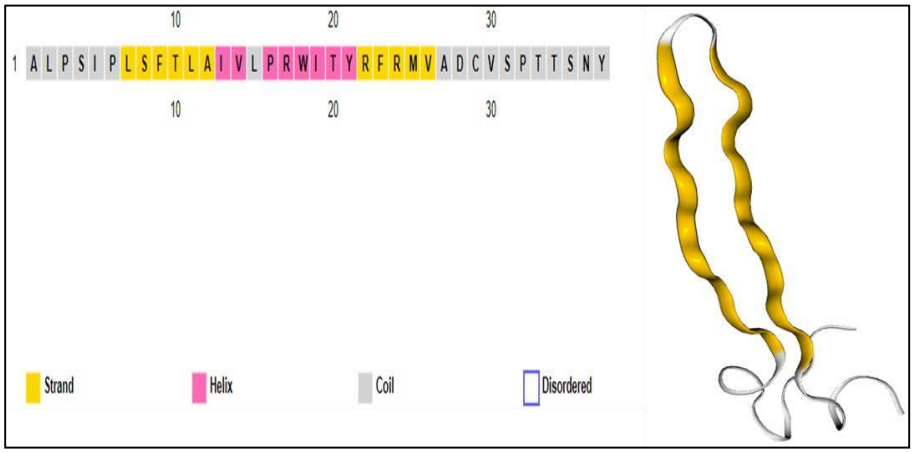


Figure 7.14: The secondary and tertiary structure of S3-P8

S3-P15: We found that, both the secondary and the tertiary structure of S3-P15 contains two β-strands. (Figure 7.15).

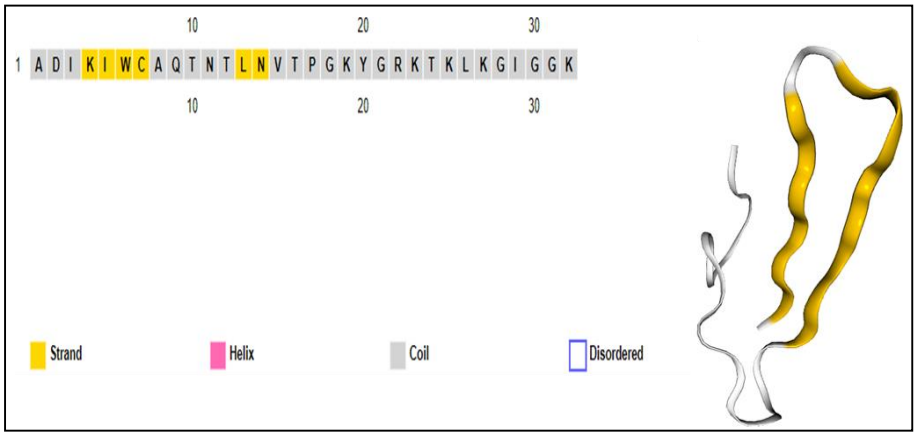


Figure 7.15: The secondary and tertiary structure of S3-P15

XSD1-P10: We found that the secondary structure of XSD1-P10 contains only coiled-coil structure (highlighted in gray). However, its tertiary structure contains one α-helix (Figure 7.16).

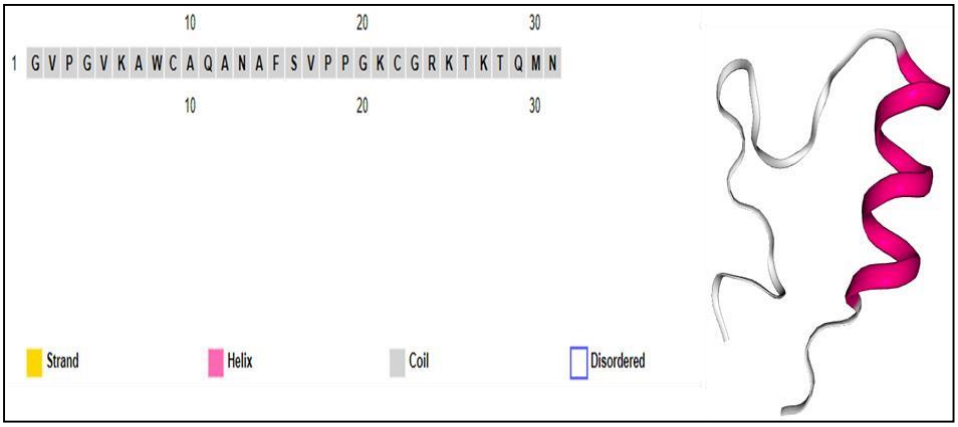


Figure 7.16: The secondary and tertiary structure of XSD1-P10

XSD1-P11: We found that the secondary structure of XSD1-P11 contains two β-strands and one α-helix but its tertiary structure shows only two α-helixes (Figure 7.17).

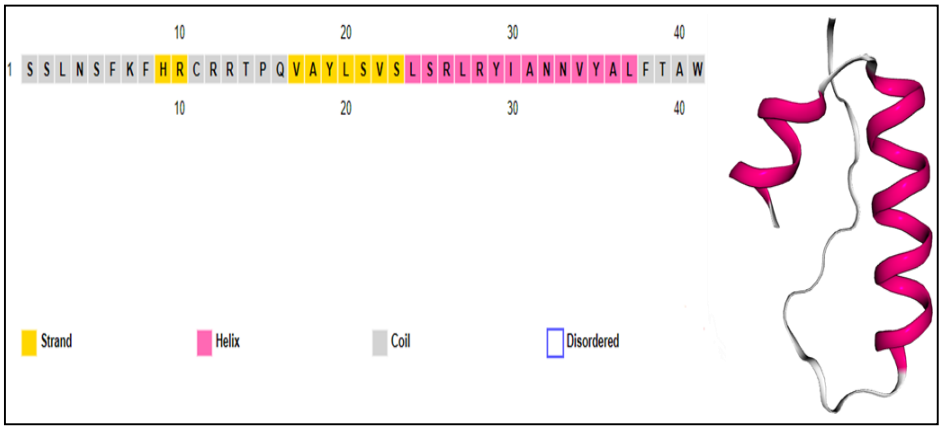


Figure 7.17: The secondary and tertiary structure of XSD1-P11

XSD2-P9: We found that the secondary structure of XSD2-P9 contains two β-strands and one α-helix, but its tertiary structure consists of one α-helix (Figure 7.18).

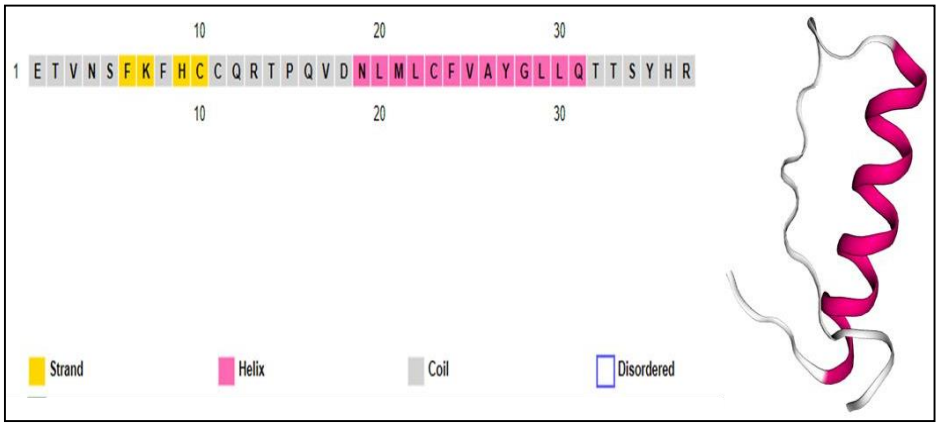


Figure 7.18: The secondary and tertiary structure of XSD2-P9

XSD2-P10: We found that the secondary structure of XSD2-P10 contains three β-strands and three α-helices but its tertiary structure contains only one α-helix (Figure 7.19).

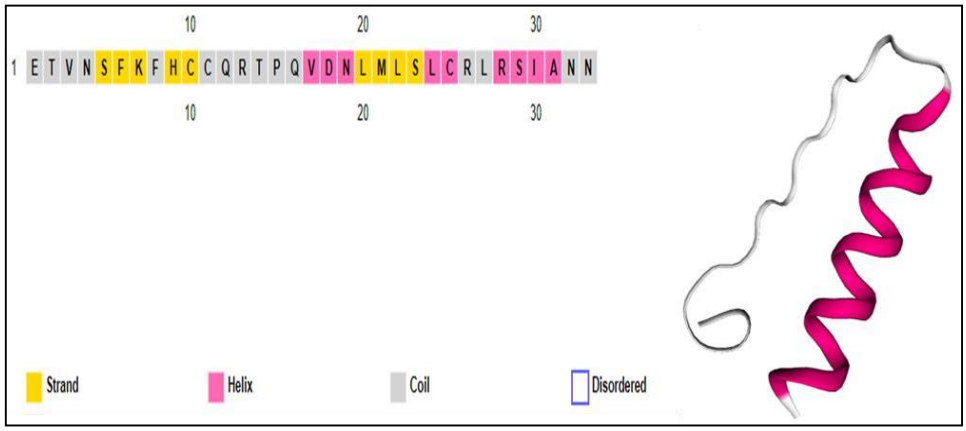


Figure 7.19: The secondary and tertiary structure of XSD2-P10

XSD2-P19: We found that the secondary structure of XSD2-P19 contains one β-strand but its tertiary structure consists of two α-helixes (Figure 7.20).

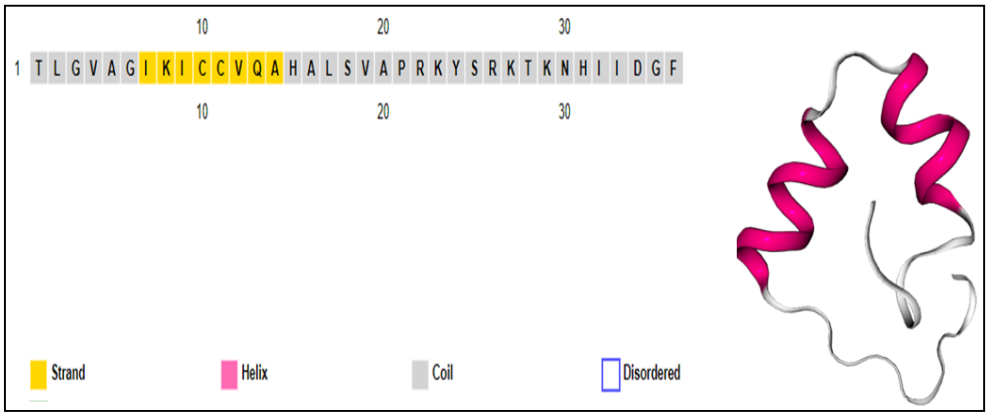


Figure 7.20: The secondary and tertiary structure of XSD2-P19

XSD3-P21: We found that the secondary structure of XSD3-P21 could not be predicted as its amino acid length was below thirty. However, its tertiary structure consists of one α-helix and one β-strand (Figure 7.21)

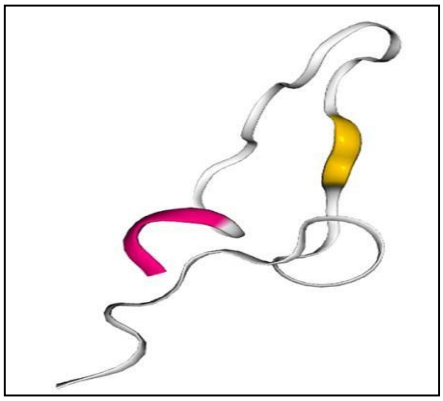


Figure 7.21: The tertiary structure of XSD3-P21

XSD4-P5: We found that the secondary structure of XSD4-P5 contains two β-strands and one α-helix but its tertiary structure consists of two α-helixes and only one β-strand (Figure 7.22).

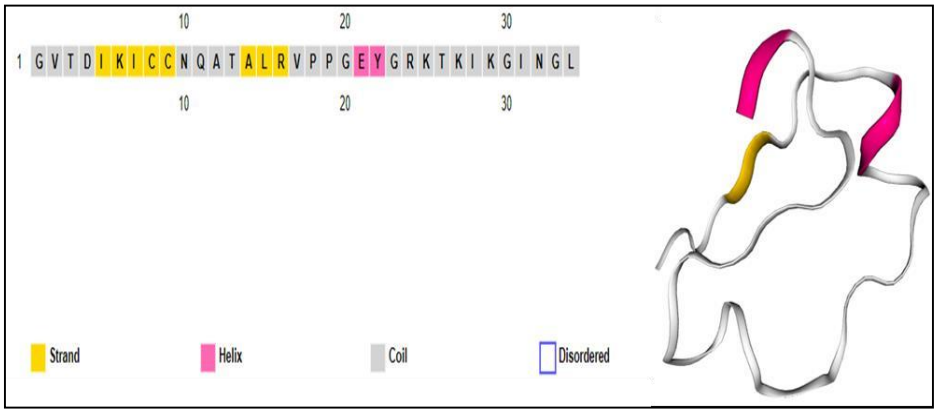


Figure 7.22: The secondary and tertiary structure of XSD4-P5

XSD4-P20: We found that the secondary structure of XSD4-P20 contains three β-strands but its tertiary structure consists of two β-strands and one α-helix (Figure 7.23).

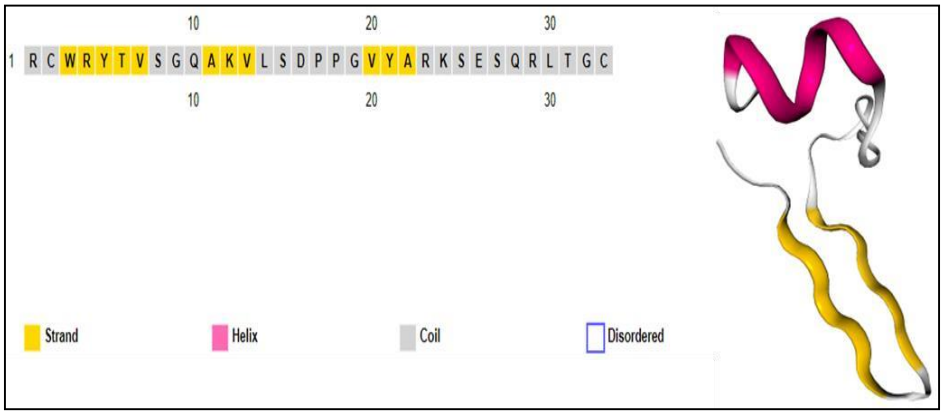


Figure 7.23: The secondary and tertiary structure of XSD4-P5

7.3 ProtParam Results

From ProtParam website, we were able to find out the amino acid length, molecular weight, isoelectric point, number of positively and negatively charged residues, half-life and GRAVY (average hydrophobicity) (Table 7.18). The results show that the length of the AMPs ranged from 28 to 41 amino-acid residues, with an average length of 34 residues. The molecular masses of the AMPs ranged from 3000kD up to 4800kD. Their theoretical pI (isoelectric point) values ranged from 8.04 to 11.57 in the AMPs. We observed that the number of positively charged residues ranged from 3 to 7 whereas the number of negatively charged residues ranged from 0 to 2. On average the AMPs contained 1 negatively charged residue and 5 positively charged residue meaning that all of them contained overall positive charge and were cationic in nature. The estimated half-lives of the proteins ranged from 0.8 hours up to 30 hours. Almost half of the AMPs had positive GRAVY value (The grand average of the hydropathicity) and the other half had negative GRAVY score. Negative value indicates that the peptide is non-polar and less hydrophobic, while positive GRAVY indicates that the AMP is polar and more hydrophobic. Higher positive value indicates a greater hydrophobicity, among the proteins described in Table 5.18, HGD3-P6 was the most hydrophobic AMP with a GRAVY value of 0.773. Lastly, 16 AMPs were stable with a stability index below 40 whereases 6 AMPs were unstable with stability indexes above 40.

Table 7.3: Result of protein characterization with ProtParam

Sequence name	Number of amino acids	Mol. weight	Theoretical pI	Total number of negatively charged residues (Asp + Glu)	Total number of positively charged residues (Arg + Lys)	Estimated half-life: (Mammalian reticulocytes, *in vitro*)	GRAVY	Instability Index (S: Stable; U: Unstable)
HGD1-33	38	4520.27	9.38	1	4	1.9 hours	-0.063	17.32 (S)
HGD1-38	34	3611.24	9.51	2	6	30 hours	-0.365	20.22 (S)
HGD1-50	34	3970.72	10.81	1	6	100 hours	0.182	36.91 (S)
HGD3-6	33	3658.38	9.38	0	3	1.9 hours	0.773	63.42 (U)
HGD4-5	34	4030.8	11.57	0	6	1 hours	0.144	50.94 (U)
HGD4-8	33	3782.57	9.84	1	5	1 hours	0.439	46.19 (U)
S1-8	33	3825.57	10.05	0	3	5.5 hours	0.645	28.85 (U)
S1-34	40	4638.52	9.98	1	3	5.5 hours	0.688	52.86 (U)

Sequence name	**Number of amino acids**	**Mol. weight**	**Theoretical pI**	**Total number of negatively charged residues (Asp + Glu)**	**Total number of positively charged residues (Arg + Lys)**	**Estimated half-life: (Mammalian reticulocytes, *in vitro*)**	**GRAVY**	**Instability Index (S: Stable; U: Unstable)**
S2-2	32	3777.4	10.72	1	5	0.8 hours	-0.394	21.35 (U)
S2-15	34	3877.56	8.68	1	3	0.8 hours	0.291	21.01 (S)
S2-26	28	3002.59	10.21	0	6	30 hours	-0.25	41.49 (U)
S2-31	33	3832.54	9.5	0	3	5.5 hours	0.645	57.69 (U)
S3-8	37	4200.97	9.31	1	3	4.4 hours	0.549	38.69 (S)
S3-15	32	3461.09	10.14	1	7	4.4 hours	-0.597	-22.99 (S)
XSD1-10	31	3232.79	10.07	0	5	30 hours	-0.455	22.22 (S)

Sequence name	**Number of amino acids**	**Mol. weight**	**Theoretical pI**	**Total number of negatively charged residues (Asp + Glu)**	**Total number of positively charged residues (Arg + Lys)**	**Estimated half-life: (Mammalian reticulocytes, *in vitro*)**	**GRAVY**	**Instability Index (S: Stable; U: Unstable)**
XSD1-11	41	4837.59	10.95	0	6	1.9 hours	-0.049	39.11 (S)
XSD2-9	37	4352	8.04	2	3	1 hours	-0.146	12.59 (S)
XSD2-10	33	3839.44	8.72	2	4	1 hours	-0.276	21.01 (S)
XSD2-19	36	3896.63	9.9	1	6	7.2 hours	0.128	19.01 (S)
XSD3-21	29	3203.88	10.2	1	7	5.5 hours	-0.434	5.25 (S)
XSD4-5	34	3615.27	9.59	2	6	30 hours	-0.238	-2.13 (S)
XSD4-20	33	3700.21	9.69	2	6	1 hours	-0.712	35.11 (S)

Chapter 8. Discussion & Suggestions

In recent years, many studies have been conducted to find novel drugs to tackle new variants of disease-causing pathogens and also to get potential alternatives for antibiotics to effectively treat MDR bacteria. Years of research and numerous experiments indicate that natural antimicrobial peptides synthesized in practically all living beings to fight disease have all the capabilities of becoming clinically applicable drugs to treat disease as well as being considered as high potential replacements for conventional antibiotics to tackle multidrug resistant bacterial infections. In addition to their antimicrobial activities, AMPs have antifungal, antiparasitic, antiviral antitumoral and immunomodulatory activities as we discussed in the review. Along with this broad spectrum of activity, AMPs offer many possibilities for chemical modification, and can be used in conjugation with other drugs which might lead to designing more effective and safe therapeutic drugs [401]. Consequently, an extensive array of these drugs are currently in different stages of development [402]. However, it is labor-intensive and time-consuming to design experimental methods to discover natural AMPs. We were thus inspired to design a database-assisted workflow for predicting AMP from metatranscriptomic data available online and performing comprehensive analysis of functional and physicochemical parameters of the predicted AMPs.

Already more than a few AMP-based drugs have been synthesized from different species including mammals [403], insects [404], amphibians [405], plants [406] and bacteria. We wanted to search for antimicrobial peptides from a source which has not been studied much till now. Marine sediment microbiota was a perfect choice for that as it is less explored compared to the terrestrial microbes. Just like 99% of the microbes in the world, marine sediment microbiota is difficult to culture in a laboratory. So, we took the help of metagenomic and meta-transcriptomic studies for getting the genomic data of marine sediment microbiota. Meta-transcriptomic data was more suitable for AMP prediction, as it specifically deals with the actively expressed genes and functional gene products.

In order to predict potential AMPs from the marine sedimentary microbiome, we retrieved 15 CDS files containing meta-transcriptome of marine sediment microbiota obtained from

the intertidal zone of Weihai, China and named as HGD, S and XSD. The 15 files contained 417 protein sequences in total on which we employed four machine learning tools namely CAMPR3, AMP scanner, AMPA and ADAM for sequence-based prediction of AMPs. As different predictors are based on different machine learning algorithms the result generated in each tool varied from the others (Table 7.16). For example, 20.38% of the protein sequences code for AMPs according to CAMPR3, 13.43% according to AMP scanner, 15.35% according to AMPA and 22.54% according to ADAM. To find out the AMP-coding sequences with 100% accuracy, we constructed five different filters by putting the four AMP predictors in different combinations (Table 7.3) and selected only those sequences that were predicted as AMP-positive by at least three of the four different prediction tools (Table 7.17). By applying five different filters, we got 22 AMPs in total from our experiment.

Further we visualized the secondary structures of the predicted AMPs with PSIPRED and tertiary structures with the help of PEP-FOLD3. We observed that in many cases, the predicted secondary structure of the AMP differed from the tertiary structure. Between the two structure prediction results, we consider the tertiary structure prediction to be more accurate as secondary structure is predicted merely on the knowledge of amino acid sequence whereas the tertiary structure is predicted by calculating the spatial disposition of each atom and the different types of side chain interactions, such as hydrogen bonding, disulfide bonds, and hydrophobic interactions etc. From the structure analysis of our predicted AMPs, we saw that among the 22 AMPs 8 of them comprised of two α-helices, another 8 of them comprised of either β-strands or α-helices and 6 of them contained both β-strands and α-helices. All of their structural information makes them fall into the category of small peptides as none of the AMPs contained more than 3 different structural patterns.

After that, we wanted to perform functional analysis of the AMPs with InterProScan which classifies them into families and predicts their domains and important sites. However, we were unsuccessful in this because InterProScan could neither retrieve information on the protein domains nor the protein-family memberships of the AMPs as we stated earlier that all of them are very small proteins. Lastly, we interpreted different important physicochemical parameters of the AMPs like molecular mass, amino acid length,

isoelectric points, number of positive and negatively charged residues, half-life and GRAVY with ProtParam as having prior knowledge of all these aspects are crucial for performing laboratory research/experiment on these proteins. For instance, knowing the isoelectric point is essential for the isolation and purification of a protein as the protein's solubility is minimal at this pH, mobility in an electro-focusing system is zero and we can collect the protein from a solution at this point. It is also important to know about the half-life of the AMP for making drugs based on it, as this value determines the time that AMP needs to get excreted from the body. After one half-life has passed, 50% of the initial AMP amount is removed from the body. Also, for determining the hydrophobicity of our AMPs, we predicted the GRAVY, the grand average of hydropathy, which is calculated by the sum of hydropathy values of all amino acids divided by the protein length. Proteins with GRAVY score below 0 are considered as hydrophilic globular proteins while proteins that have GRAVY scores above 0 are more likely to be hydrophobic membranous proteins.

Nevertheless, as all the 22 AMPs are obtained by following a workflow of *in silico* prediction of AMPs from amino acid sequences of transcriptomic data, functionality of these AMPs is yet to be discovered and their antimicrobial potency and spectrum of effectiveness on different microorganisms is yet to be validated. Extensive laboratory experiments and *in vivo* testing is required for the fruitfulness of this study. These AMP candidates need further experimental evaluation and *in vitro* testing against a broad panel of pathogenic bacterial, fungal and yeast tester strains to ensure their antimicrobial potency and effectiveness against disease. The most promising lead candidate can further be tested for potency, toxicity and efficacy in animal models. If the AMPs are proven effective against pathogens, further studies can be conducted for the large-scale synthesis of these AMPs for clinical applications.

Chapter 9. Concluding Remarks

AMPs are promising alternatives of conventional antibiotics to tackle MDR bacterial infections. Also, they have applications in cancer treatment, wound healing and battling other non-bacterial infections. In the past, scientists have always suffered from the high cost and low success rate associated with the identification of novel and effective AMPs from natural sources. This struggle has been reduced considerably with the advent of new technological advances and computational methods. Using a combination of metatranscriptomics and computational approaches, we were able to show that the marine sediment microbiome is a promising resource for bio-prospecting novel AMPs as we could identify twenty-two AMPs (Table 5.17) from 15 marine sediment metatranscriptomic dataset. Here we proposed a computational workflow for predicting AMPs from metatranscriptomic data, and we were successful in finding a significant number of potential AMPs by following this workflow. Through this experiment we wanted to show that metatranscriptomic datasets are great resources for finding novel proteins and antimicrobial peptides if explored with proper computational tools. Thus, we encourage more studies that include functional characterization of other metatranscriptomic data available on different databases which will lead to the discovery of valuable biological compounds.

Given the technological advances, improvements in genomic methods and computational analytic approaches as well as the growing abundance of omics data, the approach developed and presented here, alongside other rational design and deep learning approaches will facilitate the discovery of novel AMPs and other bioactive from environments where conventional isolation and cultivation of microorganisms has been challenging. The 22 AMPs identified from 15 marine sediment metatranscriptome dataset shows that the microbial community within the marine sediment serve as an invaluable resource for urgently needed alternatives to currently available antibiotics. Furthermore, the workflow performed here emphasizes the usefulness of *in silico* techniques for the rapid identification of new AMP candidates from metatranscriptomic datasets. We hope that we were successful in demonstrating an easy method of obtaining AMPs from metagenomic

samples by using a number of computational tools will encourage the idea of discovering novel proteins or bioactive compounds from available metatranscriptomic datasets. We believe that this experiment will encourage others to utilize available data to decipher new information.

Bibliography

1. Luong, T., A.-C. Salabarria, and D.R. Roach, *Phage Therapy in the Resistance Era: Where Do We Stand and Where Are We Going?* Clinical Therapeutics, 2020. **42**(9): p. 1659-1680.
2. Haney, E.F., S.K. Straus, and R.E.J.F.i.c. Hancock, *Reassessing the host defense peptide landscape*. 2019. **7**: p. 43.
3. Bhattacharjya, S. and S.K.J.I.J.o.M.S. Straus, *Design, engineering and discovery of novel α-helical and β-boomerang antimicrobial peptides against drug resistant bacteria.* 2020. **21**(16): p. 5773.
4. Mahlapuu, M., C. Björn, and J.J.C.r.i.b. Ekblom, *Antimicrobial peptides as therapeutic agents: Opportunities and challenges.* 2020. **40**(7): p. 978-992.
5. Luong, H.X., T.T. Thanh, and T.H.J.L.s. Tran, *Antimicrobial peptides–Advances in development of therapeutic applications.* 2020. **260**: p. 118407.
6. Liu, S., et al., *Computational resources and tools for antimicrobial peptides.* 2017. **23**(1): p. 4-12.
7. B Hadley, E. and R.J.C.t.i.m.c. EW Hancock, *Strategies for the discovery and advancement of novel cationic antimicrobial peptides.* 2010. **10**(18): p. 1872-1881.
8. Clara Pestana-Calsa, M., et al., *Bioinformatics-coupled molecular approaches for unravelling potential antimicrobial peptides coding genes in Brazilian native and crop plant species.* 2010. **11**(3): p. 199-209.
9. Waghu, F.H., et al., *CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides.* Nucleic Acids Research, 2015. **44**(D1): p. D1094-D1097.
10. Veltri, D., U. Kamath, and A. Shehu, *Deep learning improves antimicrobial peptide recognition.* Bioinformatics, 2018. **34**(16): p. 2740-2747.
11. Torrent, M., et al., *AMPA: An automated web server for prediction of protein antimicrobial regions.* Bioinformatics (Oxford, England), 2011. **28**: p. 130-1.
12. Lata, S., N.K. Mishra, and G.P.S. Raghava, *AntiBP2: improved version of antibacterial peptide prediction.* BMC Bioinformatics, 2010. **11**(1): p. S19.
13. Pirtskhalava, M., et al., *DBAASP v3: database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics.* Nucleic Acids Res, 2021. **49**(D1): p. D288-d297.
14. Jhong, J.H., et al., *dbAMP 2.0: updated resource for antimicrobial peptides with an enhanced scanning method for genomic and proteomic data.* Nucleic Acids Res, 2022. **50**(D1): p. D460-d470.
15. Huang, K.-Y., et al., *Identification of natural antimicrobial peptides from bacteria through metagenomic and metatranscriptomic analysis of high-throughput transcriptome data of Taiwanese oolong teas.* BMC Systems Biology, 2017. **11**(7): p. 131.
16. Lin, Y., et al., *An advanced approach to identify antimicrobial peptides and their function types for penaeus through machine learning strategies.* BMC Bioinformatics, 2019. **20**(Suppl 8): p. 291.
17. Qutb, A.M., F. Wei, and W. Dong, *Prediction and Characterization of Cationic Arginine-Rich Plant Antimicrobial Peptide SM-985 From Teosinte (Zea mays ssp. mexicana).* 2020. **11**.
18. Zhang, C. and S.-K.J.M.d. Kim, *Research and application of marine microbial enzymes: status and prospects.* 2010. **8**(6): p. 1920-1934.
19. Fulzele, R., et al., *Characterization of novel extracellular protease produced by marine bacterial isolate from the Indian Ocean.* 2011. **42**: p. 1364-1373.
20. El-Hassayeb, H. and S.J.I.J.C.M.A.S. Abdel Aziz, *Screening, production and industrial application of protease enzyme from marine bacteria.* 2016. **5**.
21. Pettit, G.R., et al., *Isolation and structure of bryostatin 1.* 1982. **104**(24): p. 6846-6848.
22. Sigel, M., et al. *Anticellular and antitumor activity of extracts from tropical marine invertebrates*. in *Food-drugs from the sea proceedings*. 1969. Marine Technology Society Washington DC.
23. Kashfi, R., et al., *Metabolomic Diversity and Identification of Antibacterial Activities of Bacteria Isolated From Marine Sediments in Hawai'i and Puerto Rico.* 2020. **7**.
24. Zasloff, M.J.n., *Antimicrobial peptides of multicellular organisms.* 2002. **415**(6870): p. 389-395.
25. Hancock, R.E. and G.J.T.i.m. Diamond, *The role of cationic antimicrobial peptides in innate host defences.* 2000. **8**(9): p. 402-410.
26. Hancock, R.E., D.S.J.A.a. Chapple, and chemotherapy, *Peptide antibiotics.* 1999. **43**(6): p. 1317-1323.

27. Hancock, R.E.J.E.o.o.i.d., *Cationic antimicrobial peptides: towards clinical applications.* 2000. **9**(8): p. 1723-1729.
28. Wang, Z. and G. Wang, *APD: the Antimicrobial Peptide Database.* Nucleic Acids Res, 2004. **32**(Database issue): p. D590-2.
29. Wang, G., X. Li, and Z. Wang, *APD2: the updated antimicrobial peptide database and its application in peptide design.* Nucleic Acids Res, 2009. **37**(Database issue): p. D933-7.
30. Seshadri Sundararajan, V., et al., *DAMPD: a manually curated antimicrobial peptide database.* 2012. **40**(D1): p. D1108-D1112.
31. Thomas, S., et al., *CAMP: a useful resource for research on antimicrobial peptides.* 2010. **38**(suppl_1): p. D774-D780.
32. Piotto, S.P., et al., *YADAMP: yet another database of antimicrobial peptides.* 2012. **39**(4): p. 346-351.
33. Wang, G., X. Li, and Z. Wang, *APD3: the antimicrobial peptide database as a tool for research and education.* Nucleic Acids Research, 2015. **44**(D1): p. D1087-D1093.
34. Kang, X., et al., *DRAMP 2.0, an updated data repository of antimicrobial peptides.* 2019. **6**(1): p. 1-10.
35. Boparai, J.K., P.K.J.P. Sharma, and P. Letters, *Mini review on antimicrobial peptides, sources, mechanism and recent applications.* 2020. **27**(1): p. 4-16.
36. Vogel, H., et al., *RNA-sequencing analysis reveals abundant developmental stage-specific and immunity-related genes in the pollen beetle M eligethes aeneus.* 2014. **23**(1): p. 98-112.
37. Abry, M.F., et al., *Comparative genomics identifies male accessory gland proteins in five Glossina species.* 2017. **2**.
38. Farouk, A.-E., et al., *Inducible antimicrobial compounds (Halal) production in honey bee larvae (Apis mellifera) from Rumaida, Taif by injecting of various dead microorganisms extracts.* 2017. **5**(2): p. 023-029.
39. Lee, J., D.G.J.J.o.m. Lee, and biotechnology, *Antimicrobial peptides (AMPs) with dual mechanisms: membrane disruption and apoptosis.* 2015. **25**(6): p. 759-764.
40. Hansen, I.A., et al., *Small mosquitoes, large implications: crowding and starvation affects gene expression and nutrient accumulation in Aedes aegypti*. 2015.
41. Allocca, M., S. Zola, and P.J.D.m. Bellosta, *The Fruit Fly, Drosophila melanogaster: modeling of human diseases (Part II).* 2018: p. 131-156.
42. Thiyonila, B., et al., *Dung beetle gut microbes: diversity, metabolic and immunity related roles in host system.* 2018. **1**.
43. Manabe, T. and K.J.S.r. Kawasaki, *D-form KLKLLLLLKLK-NH2 peptide exerts higher antimicrobial properties than its L-form counterpart via an association with bacterial cell wall components.* 2017. **7**(1): p. 1-10.
44. Yang, Y.T., et al., *Tenebrio molitor Gram-negative-binding protein 3 (TmGNBP3) is essential for inducing downstream antifungal Tenecin 1 gene expression against infection with Beauveria bassiana JEF-007.* 2018. **25**(6): p. 969-977.
45. Duwadi, D., et al., *Identification and screening of potent antimicrobial peptides in arthropod genomes.* 2018. **103**: p. 26-30.
46. Sheehan, G., et al., *The human cathelicidin antimicrobial peptide LL-37 promotes the growth of the pulmonary pathogen Aspergillus fumigatus.* 2018. **86**(7): p. e00097-18.
47. Schaal, J.B., et al., *Macrocyclic θ-defensins suppress tumor necrosis factor-α (TNF-α) shedding by inhibition of TNF-α–converting enzyme.* 2018. **293**(8): p. 2725-2734.
48. Khurshid, Z., et al., *Histatin peptides: Pharmacological functions and their applications in dentistry.* 2017. **25**(1): p. 25-31.
49. Baxter, A.A., et al., *Tumor cell membrane-targeting cationic antimicrobial peptides: novel insights into mechanisms of action and therapeutic prospects.* 2017. **74**(20): p. 3809-3825.
50. V Panteleev, P., et al., *A therapeutic potential of animal β-hairpin antimicrobial peptides.* 2017. **24**(17): p. 1724-1746.
51. Young-Speirs, M., et al., *Host defense cathelicidins in cattle: types, production, bioactive functions and potential therapeutic and diagnostic applications.* 2018. **51**(6): p. 813-821.
52. Savelyeva, A., et al., *Anticancer Genes.* 2014, Springer.
53. Sun, T., B. Zhan, and Y.J.G. Gao, *A novel cathelicidin from Bufo bufo gargarizans Cantor showed specific activity to its habitat bacteria.* 2015. **571**(2): p. 172-177.

54. Upadhyay, R.K.J.I.J.o.G.P., *Spider venom toxins, its purification, solubilization, and antimicrobial activity*. 2018. **12**(01).
55. Belmadani, A., A. Semlali, and M.J.J.o.a.m. Rouabhia, *Dermaseptin-S1 decreases Candida albicans growth, biofilm formation and the expression of hyphal wall protein 1 and aspartic protease genes.* 2018. **125**(1): p. 72-83.
56. Tahir, H.M., et al., *Antibacterial potential of venom extracted from wolf spider, Lycosa terrestris (Araneae: Lycosiade).* 2018. **52**(2): p. 286-290.
57. Kuzmin, D., et al., *Effect of N-and C-terminal modifications on cytotoxic properties of antimicrobial peptide tachyplesin I.* 2017. **162**(6): p. 754-757.
58. Coulen, S.C., et al., *Valorisation of proteins from rubber tree.* 2017. **8**(4): p. 1027-1041.
59. Lan, N.T.N., H.T. Thao, and C.H. Mau, *Overexpression of VrPDF1 gene confers resistance to weevils in transgenic mung bean plants*. 2017, PeerJ Preprints.
60. Mills, S., et al., *A multibacteriocin cheese starter system, comprising nisin and lacticin 3147 in Lactococcus lactis, in combination with plantaricin from Lactobacillus plantarum.* 2017. **83**(14): p. e00799-17.
61. Su, Z., et al., *EIS and PM-IRRAS studies of alamethicin ion channels in a tethered lipid bilayer.* 2018. **812**: p. 213-220.
62. Braïek, O.B., et al., *Biotechnological potential, probiotic and safety properties of newly isolated enterocin-producing Enterococcus lactis strains.* 2018. **92**: p. 361-370.
63. Ebrahimipour, G.H., et al., *Isolation, Partial Purification and Characterization of an Antimicrobial Compound, Produced by Bacillus atrophaeus.* 2014. **7**(9).
64. Sharma, G., et al., *Antibacterial activity, cytotoxicity, and the mechanism of action of bacteriocin from Bacillus subtilis GAS101*. 2018. **27**(2): p. 186-192.
65. Jiang, H., et al., *Plantaricin NC8 from Lactobacillus plantarum causes cell membrane disruption to Micrococcus luteus without targeting lipid II.* 2018. **102**(17): p. 7465-7473.
66. Hammi, I., et al., *Maltaricin CPN, a new class II a bacteriocin produced by Carnobacterium maltaromaticum CPN isolated from mould-ripened cheese.* 2016. **121**(5): p. 1268-1274.
67. Chen, Y.-s., et al., *Leucocin C-607, a novel bacteriocin from the multiple-bacteriocin-producing Leuconostoc pseudomesenteroides 607 isolated from persimmon.* 2018. **10**(2): p. 148-156.
68. Singh, R., et al., *Identification of unstructured model for subtilin production through Bacillus subtilis using hybrid genetic algorithm.* 2017. **60**: p. 1-12.
69. Guzmán-Rodríguez, J.J., et al., *Plant antimicrobial peptides as potential anticancer agents.* 2015. **2015**.
70. Chan, J.C. and E.C. Li-Chan, *Production of lactoferricin and other cationic peptides from food grade bovine lactoferrin with various iron saturation levels.* J Agric Food Chem, 2007. **55**(2): p. 493-501.
71. Conlon, J. and Á. Sonnevend, *Antimicrobial Peptides in Frog Skin Secretions.* Methods in molecular biology (Clifton, N.J.), 2010. **618**: p. 3-14.
72. Conlon, J.M. and M.J.P. Mechkarska, *Host-defense peptides with therapeutic potential from skin secretions of frogs from the family Pipidae.* 2014. **7**(1): p. 58-77.
73. Lu, Y., et al., *The first antimicrobial peptide from sea amphibian.* 2008. **45**(3): p. 678-681.
74. Dutta, P., et al., *Beneficial role of insect-derived bioactive components against inflammation and its associated complications (colitis and arthritis) and cancer.* 2019. **313**: p. 108824.
75. Zahedifard, F., et al., *Comparative study of different forms of Jellein antimicrobial peptide on Leishmania parasite*. 2020. **209**: p. 107823.
76. Ghodhbane, H., et al., *Bacteriocins active against multi-resistant gram negative bacteria implicated in nosocomial infections.* 2015. **15**(1): p. 2-12.
77. Meade, E., M.A. Slattery, and M. Garvey, *Bacteriocins, Potent Antimicrobial Peptides and the Fight against Multi Drug Resistant Species: Resistance Is Futile?* Antibiotics (Basel), 2020. **9**(1).
78. Wang, S., et al., *Antimicrobial peptides as potential alternatives to antibiotics in food animal industry.* 2016. **17**(5): p. 603.
79. Campos, M.L., et al., *The role of antimicrobial peptides in plant immunity.* 2018. **69**(21): p. 4997-5011.
80. Stotz, H.U., J.G. Thomson, and Y. Wang, *Plant defensins: defense, development and application.* Plant Signal Behav, 2009. **4**(11): p. 1010-2.
81. Faulkner, D.J.J.N.p.r., *Marine natural products.* 2001. **18**(1): p. 1R-49R.

82. Kennedy, J., et al., *Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems.* 2011. **111**(4): p. 787-799.
83. Oh, R., et al., *Myticusin-beta, antimicrobial peptide from the marine bivalve, Mytilus coruscus.* 2020. **99**: p. 342-352.
84. Rathinakumar, R. and W.C.J.T.F.j. Wimley, *High-throughput discovery of broad-spectrum peptide antibiotics.* 2010. **24**(9): p. 3232-3238.
85. Yeaman, M.R. and N.Y. Yount, *Mechanisms of Antimicrobial Peptide Action and Resistance.* 2003. **55**(1): p. 27-55.
86. Takahashi, D., et al., *Structural determinants of host defense peptides for antimicrobial activity and target cell selectivity.* 2010. **92**(9): p. 1236-1241.
87. Lei, J., et al., *The antimicrobial peptides and their potential clinical applications.* 2019. **11**(7): p. 3919.
88. Yeaman, M.R. and N.Y.J.P.r. Yount, *Mechanisms of antimicrobial peptide action and resistance.* 2003. **55**(1): p. 27-55.
89. Powers, J.-P.S. and R.E.J.P. Hancock, *The relationship between peptide structure and antibacterial activity.* 2003. **24**(11): p. 1681-1691.
90. Nguyen, L.T., E.F. Haney, and H.J.J.T.i.b. Vogel, *The expanding scope of antimicrobial peptide structures and their modes of action.* 2011. **29**(9): p. 464-472.
91. Kumar, P., J.N. Kizhakkedathu, and S.K. Straus, *Antimicrobial Peptides: Diversity, Mechanism of Action and Strategies to Improve the Activity and Biocompatibility In vivo.* Biomolecules, 2018. **8**(1).
92. Ehrenstein, G. and H.J.Q.r.o.b. Lecar, *Electrically gated ionic channels in lipid bilayers.* 1977. **10**(1): p. 1-34.
93. Bauer, R. and L.J.I.j.o.f.m. Dicks, *Mode of action of lipid II-targeting lantibiotics.* 2005. **101**(2): p. 201-216.
94. Yang, L., et al., *Barrel-stave model or toroidal model? A case study on melittin pores.* 2001. **81**(3): p. 1475-1485.
95. Li, Y., et al., *Overview on the recent study of antimicrobial peptides: origins, functions, relative mechanisms and application.* 2012. **37**(2): p. 207-215.
96. Brogden, K.A.J.N.r.m., *Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?* 2005. **3**(3): p. 238-250.
97. Münch, D. and H.-G.J.B.e.B.A.-B. Sahl, *Structural variations of the cell wall precursor lipid II in Gram-positive bacteria—Impact on binding and efficacy of antimicrobial peptides.* 2015. **1848**(11): p. 3062-3071.
98. de Leeuw, E., et al., *Functional interaction of human neutrophil peptide-1 with the cell wall precursor lipid II.* 2010. **584**(8): p. 1543-1548.
99. Park, C.B., et al., *Mechanism of action of the antimicrobial peptide buforin II: buforin II kills microorganisms by penetrating the cell membrane and inhibiting cellular functions.* 1998. **244**(1): p. 253-257.
100. Subbalakshmi, C. and N.J.F.m.l. Sitaram, *Mechanism of antimicrobial action of indolicidin.* 1998. **160**(1): p. 91-96.
101. Lehrer, R., et al., *Interaction of human defensins with Escherichia coli. Mechanism of bactericidal activity.* 1989. **84**(2): p. 553-561.
102. Sharma, H. and R.J.P.O. Nagaraj, *Human β-defensin 4 with non-native disulfide bridges exhibit antimicrobial activity.* 2015. **10**(3): p. e0119525.
103. Meng, S., et al., *Research advances of antimicrobial peptides and applications in food industry and agriculture.* 2010. **11**(4): p. 264-273.
104. Carrera, M., et al., *Characterization of foodborne strains of Staphylococcus aureus by shotgun proteomics: Functional networks, virulence factors and species-specific peptide biomarkers.* 2017. **8**: p. 2458.
105. Nagarajan, K., et al., *Peptide therapeutics versus superbugs: highlight on current research and advancements.* 2018. **24**(1): p. 19-33.
106. Le, C.-F., et al., *Intracellular targeting mechanisms by antimicrobial peptides.* 2017. **61**(4): p. e02340-16.

107. Gordon, Y.J., E.G. Romanowski, and A.M.J.C.e.r. McDermott, *A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs.* 2005. **30**(7): p. 505-515.
108. Mirski, T., et al., *Utilisation of peptides against microbial infections–a review.* 2018. **25**(2).
109. Wuerth, K., *Combating Pseudomonas aeruginosa lung infections using synthetic host defense peptides.* 2017, UNIVERSITY OF BRITISH COLUMBIA (Vancouver.
110. Barlow, P.G., et al., *The human cationic host defense peptide LL-37 mediates contrasting effects on apoptotic pathways in different primary cells of the innate immune system.* J Leukoc Biol, 2006. **80**(3): p. 509-20.
111. De, Y., et al., *LL-37, the neutrophil granule- and epithelial cell-derived cathelicidin, utilizes formyl peptide receptor-like 1 (FPRL1) as a receptor to chemoattract human peripheral blood neutrophils, monocytes, and T cells.* J Exp Med, 2000. **192**(7): p. 1069-74.
112. Kosikowska, P. and A.J.E.o.o.t.p. Lesner, *Antimicrobial peptides (AMPs) as drug candidates: a patent review (2003–2015).* 2016. **26**(6): p. 689-702.
113. Taniguchi, M., et al., *Wound healing activity and mechanism of action of antimicrobial and lipopolysaccharide-neutralizing peptides from enzymatic hydrolysates of rice bran proteins.* 2019. **128**(2): p. 142-148.
114. van der Does, A.M., P.S. Hiemstra, and N.J.A.P. Mookherjee, *Antimicrobial host defence peptides: immunomodulatory functions and translational prospects.* 2019: p. 149-171.
115. Barabas, N., et al., *Beta-defensins activate macrophages and synergize in pro-inflammatory cytokine expression induced by TLR ligands.* Immunobiology, 2013. **218**(7): p. 1005-1011.
116. Niyonsaba, F., et al., *A cathelicidin family of human antibacterial peptide LL-37 induces mast cell chemotaxis.* 2002. **106**(1): p. 20-26.
117. García, J.-R., et al., *Identification of a novel, multifunctional β-defensin (human β-defensin 3) with specific antimicrobial activity.* 2001. **306**(2): p. 257-264.
118. Liu, Y.-J.J.C., *Dendritic cell subsets and lineages, and their functions in innate and adaptive immunity.* 2001. **106**(3): p. 259-262.
119. Biragyn, A., et al., *Mediators of innate immunity that target immature, but not mature, dendritic cells induce antitumor immunity when genetically fused with nonimmunogenic tumor antigens.* J Immunol, 2001. **167**(11): p. 6644-53.
120. Boniotto, M., et al., *Human beta-defensin 2 induces a vigorous cytokine response in peripheral blood mononuclear cells.* Antimicrob Agents Chemother, 2006. **50**(4): p. 1433-41.
121. Kim, J., et al., *Expression of Beta-Defensin 131 Promotes an Innate Immune Response in Human Prostate Epithelial Cells.* PloS one, 2015. **10**: p. e0144776.
122. Funderburg, N., et al., *Human -defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2.* Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18631-5.
123. Chaly, Y., et al., *Neutrophil alpha-defensin human neutrophil peptide modulates cytokine production in human monocytes and adhesion molecule expression in endothelial cells.* European cytokine network, 2000. **11**: p. 257-66.
124. Grigat, J., et al., *Chemoattraction of macrophages, T lymphocytes, and mast cells is evolutionarily conserved within the human alpha-defensin family.* J Immunol, 2007. **179**(6): p. 3958-65.
125. Yang, D., et al., *Multiple roles of antimicrobial defensins, cathelicidins, and eosinophil-derived neurotoxin in host defense.* Annu Rev Immunol, 2004. **22**: p. 181-215.
126. Tavano, R., et al., *The honeybee antimicrobial peptide apidaecin differentially immunomodulates human macrophages, monocytes and dendritic cells.* J Innate Immun, 2011. **3**(6): p. 614-22.
127. Tang, J., et al., *A small peptide with potential ability to promote wound healing.* PLoS One, 2014. **9**(3): p. e92082.
128. Pena, O.M., et al., *Synthetic cationic peptide IDR-1018 modulates human macrophage differentiation.* PLoS One, 2013. **8**(1): p. e52449.
129. Nijnik, A., et al., *Synthetic cationic peptide IDR-1002 provides protection against bacterial infections through chemokine induction and enhanced leukocyte recruitment.* 2010. **184**(5): p. 2539-2550.
130. Scott, M.G., et al., *An anti-infective peptide that selectively modulates the innate immune response.* 2007. **25**(4): p. 465-472.
131. Mattsby-Baltzer, I., et al., *Lactoferrin or a fragment thereof inhibits the endotoxin-induced interleukin-6 response in human monocytic cells.* 1996. **40**(2): p. 257-262.

132. Nijnik, A., et al., *Synthetic cationic peptide IDR-1002 provides protection against bacterial infections through chemokine induction and enhanced leukocyte recruitment.* J Immunol, 2010. **184**(5): p. 2539-50.
133. Bahar, A.A. and D.J.P. Ren, *Antimicrobial peptides.* 2013. **6**(12): p. 1543-1575.
134. Lau, Q.Y., et al., *Elucidating the bactericidal mechanism of action of the linear antimicrobial tetrapeptide BRBR-NH2.* 2018. **1860**(8): p. 1517-1527.
135. Rathinakumar, R., W.F. Walkenhorst, and W.C.J.J.o.t.A.C.S. Wimley, *Broad-spectrum antimicrobial peptides by rational combinatorial design and high-throughput screening: the importance of interfacial activity.* 2009. **131**(22): p. 7609-7617.
136. Galván Márquez, I.J., et al., *Mode of action of nisin on Escherichia coli.* 2020. **66**(2): p. 161-168.
137. Sultana, A., H. Luo, and S. Ramakrishna, *Antimicrobial Peptides and Their Applications in Biomedical Sector.* Antibiotics (Basel), 2021. **10**(9).
138. Kerenga, B.K., et al., *Salt-tolerant antifungal and antibacterial activities of the corn defensin ZmD32.* 2019. **10**: p. 795.
139. Shurko, J.F., et al., *Evaluation of LL-37 antimicrobial peptide derivatives alone and in combination with vancomycin against S. aureus.* 2018. **71**(11): p. 971-974.
140. Yasir, M., D. Dutta, and M.D.J.S.r. Willcox, *Comparative mode of action of the antimicrobial peptide melimine and its derivative Mel4 against Pseudomonas aeruginosa.* 2019. **9**(1): p. 1-12.
141. Fenner, A.J.N.R.U., *Antimicrobial peptide derived from moths can eradicate UPEC biofilms and could offer a novel therapeutic option.* 2020. **17**(4): p. 191-191.
142. Dong, M., et al., *BING, a novel antimicrobial peptide isolated from Japanese medaka plasma, targets bacterial envelope stress response by suppressing cpxR expression.* 2021. **11**(1): p. 1-17.
143. Zaet, A., et al., *D-Cateslytin, a new antimicrobial peptide with therapeutic potential.* 2017. **7**(1): p. 1-12.
144. Porto, W.F., et al., *In silico optimization of a guava antimicrobial peptide enables combinatorial exploration for peptide design.* 2018. **9**(1): p. 1-12.
145. Ma, B., et al., *The antimicrobial peptide thanatin disrupts the bacterial outer membrane and inactivates the NDM-1 metallo-β-lactamase.* 2019. **10**(1): p. 1-11.
146. Piras, A.M., et al., *Chitosan nanoparticles loaded with the antimicrobial peptide temporin B exert a long-term antibacterial activity in vitro against clinical isolates of Staphylococcus epidermidis.* 2015. **6**: p. 372.
147. Knappe, D., et al., *Oncocin (VDKPPYLPRPRPPRRIYNR-NH2): a novel antibacterial peptide optimized against gram-negative human pathogens.* 2010. **53**(14): p. 5240-5247.
148. Costerton, J.W., P.S. Stewart, and E.P.J.S. Greenberg, *Bacterial biofilms: a common cause of persistent infections.* 1999. **284**(5418): p. 1318-1322.
149. Donlan, R.M.J.C.i.d., *Biofilm formation: a clinically relevant microbiological process.* 2001. **33**(8): p. 1387-1392.
150. Hall-Stoodley, L., J.W. Costerton, and P.J.N.r.m. Stoodley, *Bacterial biofilms: from the natural environment to infectious diseases.* 2004. **2**(2): p. 95-108.
151. Reffuveille, F., et al., *A broad-spectrum antibiofilm peptide enhances antibiotic action against bacterial biofilms.* 2014. **58**(9): p. 5363-5371.
152. Batoni, G., G. Maisetta, and S.J.B.e.B.A.-B. Esin, *Antimicrobial peptides and their interaction with biofilms of medically relevant bacteria.* 2016. **1858**(5): p. 1044-1060.
153. Yasir, M., M.D.P. Willcox, and D. Dutta, *Action of Antimicrobial Peptides against Bacterial Biofilms.* Materials (Basel), 2018. **11**(12).
154. Overhage, J., et al., *Human host defense peptide LL-37 prevents bacterial biofilm formation.* 2008. **76**(9): p. 4176-4182.
155. De La Fuente-Núñez, C., et al., *D-enantiomeric peptides that eradicate wild-type and multidrug-resistant biofilms and protect against lethal Pseudomonas aeruginosa infections.* 2015. **22**(2): p. 196-205.
156. Gordya, N., et al., *Natural antimicrobial peptide complexes in the fighting of antibiotic resistant biofilms: Calliphora vicina medicinal maggots.* 2017. **12**(3): p. e0173559.
157. Zhu, C., et al., *Human β-defensin 3 inhibits antibiotic-resistant Staphylococcus biofilm formation.* 2013. **183**(1): p. 204-213.

158. Wang, H.-y., et al., *Molecular pathways underlying inhibitory effect of antimicrobial peptide Nal-P-113 on bacteria biofilms formation of Porphyromonas gingivalis W83 by DNA microarray*. 2017. **17**(1): p. 1-7.
159. Okuda, K.-i., et al., *Effects of bacteriocins on methicillin-resistant Staphylococcus aureus biofilm.* 2013. **57**(11): p. 5572-5579.
160. Libardo, M.D.J., et al., *Nuclease activity gives an edge to host-defense peptide piscidin 3 over piscidin 1, rendering it more effective against persisters and biofilms.* 2017. **284**(21): p. 3662-3683.
161. Yuan, Y., et al., *A novel membrane-disruptive antimicrobial peptide from frog skin secretion against cystic fibrosis isolates and evaluation of anti-MRSA effect using Galleria mellonella model*. 2019. **1863**(5): p. 849-856.
162. Khozani, R.S., et al., *Kinetics study of antimicrobial peptide, melittin, in simultaneous biofilm degradation and eradication of potent biofilm producing MDR Pseudomonas aeruginosa isolates.* 2019. **25**(1): p. 329-338.
163. Mohammadi, M., et al., *Identification and characterization of novel antimicrobial peptide from hippocampus comes by In silico and experimental studies.* 2018. **20**(6): p. 718-728.
164. Von Borowski, R.G., et al., *First-in-class matrix anti-assembly peptide prevents staphylococcal biofilm in vitro and in vivo.* 2020.
165. Wanderley, J.L., et al., *Mimicry of apoptotic cells by exposing phosphatidylserine participates in the establishment of amastigotes of Leishmania (L) amazonensis in mammalian hosts.* 2006. **176**(3): p. 1834-1839.
166. Epand, R.M. and R.F.J.J.o.P.S. Epand, *Bacterial membrane lipids in the action of antimicrobial agents.* 2011. **17**(5): p. 298-305.
167. Pretzel, J., et al., *Antiparasitic peptides.* Adv Biochem Eng Biotechnol, 2013. **135**: p. 157-92.
168. Rivas, L., J.R. Luque-Ortega, and D. Andreu, *Amphibian antimicrobial peptides and Protozoa: lessons from parasites.* Biochim Biophys Acta, 2009. **1788**(8): p. 1570-81.
169. Gwadz, R.W., et al., *Effects of magainins and cecropins on the sporogonic development of malaria parasites in mosquitoes.* 1989. **57**(9): p. 2628-2633.
170. Kokoza, V., et al., *Blocking of Plasmodium transmission by cooperative action of Cecropin A and Defensin A in transgenic Aedes aegypti mosquitoes.* 2010. **107**(18): p. 8111-8116.
171. Boman, H., et al., *Antibacterial and antimalarial properties of peptides that are cecropin-melittin hybrids.* 1989. **259**(1): p. 103-106.
172. Soravia, E., G. Martini, and M.J.F.l. Zasloff, *Antimicrobial properties of peptides from Xenopus granular gland secretions.* 1988. **228**(2): p. 337-340.
173. Zasloff, M., B. Martin, and H.-C.J.P.o.t.n.a.o.s. Chen, *Antimicrobial activity of synthetic magainin peptides and several analogues.* 1988. **85**(3): p. 910-913.
174. Giacometti, A., et al., *Activity of buforin II alone and in combination with azithromycin and minocycline against Cryptosporidium parvum in cell culture.* Journal of Antimicrobial Chemotherapy, 2001. **47**(1): p. 97-99.
175. Brand, G.D., et al., *Dermaseptins from Phyllomedusa oreades andPhyllomedusa distincta: ANTI-TRYPANOSOMA CRUZI ACTIVITY WITHOUT CYTOTOXICITY TO MAMMALIAN CELLS **. Journal of Biological Chemistry, 2002. **277**(51): p. 49332-49340.
176. Brand, G.D., et al., *Novel dermaseptins from Phyllomedusa hypochondrialis (Amphibia).* Biochemical and Biophysical Research Communications, 2006. **347**(3): p. 739-746.
177. Mangoni, M.L., et al., *Effect of natural L- to D-amino acid conversion on the organization, membrane binding, and biological function of the antimicrobial peptides bombinins H*. Biochemistry, 2006. **45**(13): p. 4266-76.
178. Mangoni, M.L., et al., *Temporins, small antimicrobial peptides with leishmanicidal activity.* J Biol Chem, 2005. **280**(2): p. 984-90.
179. Leite, J.R., et al., *Phylloseptins: a novel class of anti-bacterial and anti-protozoan peptides from the Phyllomedusa genus.* Peptides, 2005. **26**(4): p. 565-73.
180. Waugh, C.D., *Caspofungin*, in *xPharm: The Comprehensive Pharmacology Reference*, S.J. Enna and D.B. Bylund, Editors. 2007, Elsevier: New York. p. 1-4.
181. Li, T., et al., *Activity and Mechanism of Action of Antifungal Peptides from Microorganisms: A Review.* Molecules, 2021. **26**(11).

182. Zhao, X., et al., *Isolation and identification of antifungal peptides from Bacillus BH072, a novel bacterium isolated from honey.* Microbiol Res, 2013. **168**(9): p. 598-606.
183. Gao, W., et al., *[Antifungal mechanism of Bacillus marinus B-9987].* Wei Sheng Wu Xue Bao, 2009. **49**(11): p. 1494-501.
184. Lei, S., et al., *Capability of iturin from Bacillus subtilis to inhibit Candida albicans in vitro and in vivo.* Appl Microbiol Biotechnol, 2019. **103**(11): p. 4377-4392.
185. Wang, J., et al., *[Antagonism against Beauveria bassiana by lipopeptide metabolites produced by entophyte Bacillus amyloliquefaciens strain SWB16].* Wei Sheng Wu Xue Bao, 2014. **54**(7): p. 778-85.
186. Chang, W.T., C.S. Chen, and S.L. Wang, *An antifungal chitinase produced by Bacillus cereus with shrimp and crab shell powder as a carbon source.* Curr Microbiol, 2003. **47**(2): p. 102-8.
187. Mehmood, M.A., et al., *Heterologous expression of the antifungal β-chitin binding protein CBP24 from bacillus thuringiensis and its synergistic action with bacterial chitinases.* Protein Pept Lett, 2015. **22**(1): p. 39-44.
188. Ji, Z.L., et al., *Identification and characterization of a serine protease from Bacillus licheniformis W10: A potential antifungal agent.* Int J Biol Macromol, 2020. **145**: p. 594-603.
189. Yan, H., et al., *Two novel cationic antifungal peptides isolated from Bacillus pumilus HN-10 and their inhibitory activity against Trichothecium roseum.* World J Microbiol Biotechnol, 2018. **34**(2): p. 21.
190. Kajimura, Y. and M. Kaneda, *Fusaricidin A, a new depsipeptide antibiotic produced by Bacillus polymyxa KT-8. Taxonomy, fermentation, isolation, structure elucidation and biological activity.* J Antibiot (Tokyo), 1996. **49**(2): p. 129-35.
191. Sorensen, K.N., K.H. Kim, and J.Y. Takemoto, *In vitro antifungal and fungicidal activities and erythrocyte toxicities of cyclic lipodepsinonapeptides produced by Pseudomonas syringae pv. syringae.* Antimicrob Agents Chemother, 1996. **40**(12): p. 2710-3.
192. Ribeiro, P.D. and E. Medina-Acosta, *Prevention of lethal murine candidiasis using HP (2-20), an antimicrobial peptide derived from the N-terminus of Helicobacter pylori ribosomal protein L1.* Peptides, 2003. **24**(11): p. 1807-14.
193. Graham, C.E., et al., *Enterococcus faecalis bacteriocin EntV inhibits hyphal morphogenesis, biofilm formation, and virulence of Candida albicans.* Proc Natl Acad Sci U S A, 2017. **114**(17): p. 4507-4512.
194. Nyfeler, R. and W. Keller-Schierlein, *[Metabolites of microorganisms. 143. Echinocandin B, a novel polypeptide-antibiotic from Aspergillus nidulans var. echinulatus: isolation and structural components].* Helv Chim Acta, 1974. **57**(8): p. 2459-77.
195. Skouri-Gargouri, H. and A. Gargouri, *First isolation of a novel thermostable antifungal peptide secreted by Aspergillus clavatus.* Peptides, 2008. **29**(11): p. 1871-7.
196. Wen, C., W. Guo, and X. Chen, *Purification and identification of a novel antifungal protein secreted by Penicillium citrinum from the Southwest Indian Ocean.* J Microbiol Biotechnol, 2014. **24**(10): p. 1337-45.
197. Takesako, K., et al., *Biological properties of aureobasidin A, a cyclic depsipeptide antifungal antibiotic.* J Antibiot (Tokyo), 1993. **46**(9): p. 1414-20.
198. Nakamura, I., et al., *ASP2397: a novel antifungal agent produced by Acremonium persicinum MF-347833.* J Antibiot (Tokyo), 2017. **70**(1): p. 45-51.
199. Han, Y., et al., *Characterization of antifungal chitinase from marine Streptomyces sp. DA11 associated with South China Sea sponge Craniella australiensis.* Mar Biotechnol (NY), 2009. **11**(1): p. 132-40.
200. Maskey, R.P., et al., *Chandrananimycins A approximately C: production of novel anticancer antibiotics from a marine Actinomadura sp. isolate M048 by variation of medium composition and growth conditions.* J Antibiot (Tokyo), 2003. **56**(7): p. 622-9.
201. Becker, J.M., et al., *Polyoxin D inhibits growth of zoopathogenic fungi.* Antimicrob Agents Chemother, 1983. **23**(6): p. 926-9.
202. Hector, R.F., B.L. Zimmer, and D. Pappagianis, *Evaluation of nikkomycins X and Z in murine models of coccidioidomycosis, histoplasmosis, and blastomycosis.* Antimicrob Agents Chemother, 1990. **34**(4): p. 587-93.

203. Ganesan, L.T., et al., *In-vitro activity of nikkomycin Z alone and in combination with polyenes, triazoles or echinocandins against Aspergillus fumigatus.* Clin Microbiol Infect, 2004. **10**(11): p. 961-6.
204. Galdiero, S., et al., *Peptide inhibitors against herpes simplex virus infections.* 2013. **19**(3): p. 148-158.
205. Belaid, A., et al., *In vitro antiviral activity of dermaseptins against herpes simplex virus type 1.* 2002. **66**(2): p. 229-234.
206. Keogan, S., S. Passic, and F.C. Krebs, *Infection by CXCR4-Tropic Human Immunodeficiency Virus Type 1 Is Inhibited by the Cationic Cell-Penetrating Peptide Derived from HIV-1 Tat.* International Journal of Peptides, 2012. **2012**: p. 349427.
207. Sitaram, N. and R.J.B.e.B.A.-B. Nagaraj, *Interaction of antimicrobial peptides with biological and model membranes: structural and charge requirements for activity.* 1999. **1462**(1-2): p. 29-54.
208. Pärn, K., E. Eriste, and Ü.J.C.-p.p. Langel, *The antimicrobial and antiviral applications of cell-penetrating peptides.* 2015: p. 223-245.
209. Mulder, K., et al., *Current scenario of peptide-based drugs: the key roles of cationic antitumor and antiviral peptides.* 2013. **4**.
210. Agarwal, G. and R. Gabrani, *Antiviral Peptides: Identification and Validation.* Int J Pept Res Ther, 2021. **27**(1): p. 149-168.
211. Gao, Y., et al., *Synthesis and disulfide bond connectivity–activity studies of a kalata B1-inspired cyclopeptide against dengue NS2B–NS3 protease.* 2010. **18**(3): p. 1331-1336.
212. Chernysh, S., et al., *Antiviral and antitumor peptides from insects.* 2002. **99**(20): p. 12628-12632.
213. Wachinger, M., et al., *Antimicrobial peptides melittin and cecropin inhibit replication of human immunodeficiency virus 1 by suppressing viral gene expression.* 1998. **79**(4): p. 731-740.
214. Hultmark, D., et al., *Insect immunity. Purification and properties of three inducible bactericidal proteins from hemolymph of immunized pupae of Hyalophora cecropia.* 1980. **106**(1): p. 7-16.
215. Matanic, V.C.A. and V.J.I.j.o.a.a. Castilla, *Antiviral activity of antimicrobial cationic peptides against Junin virus and herpes simplex virus.* 2004. **23**(4): p. 382-389.
216. Holthausen, D.J., et al., *An amphibian host defense peptide is virucidal for human H1 hemagglutinin-bearing influenza viruses.* 2017. **46**(4): p. 587-595.
217. Wu, Z., et al., *Human neutrophil α-defensin 4 inhibits HIV-1 infection in vitro.* 2005. **579**(1): p. 162-166.
218. Meyer-Hoffert, U., et al., *Expression of human beta-defensin-2 and-3 in verrucae vulgares and condylomata acuminata.* 2008. **22**(9): p. 1050-1054.
219. Quiñones-Mateu, M.E., et al., *Human epithelial β-defensins 2 and 3 inhibit HIV-1 replication.* 2003. **17**(16): p. F39-F48.
220. Sørensen, O.E., et al., *Human cathelicidin, hCAP-18, is processed to the antimicrobial peptide LL-37 by extracellular cleavage with proteinase 3.* 2001. **97**(12): p. 3951-3959.
221. Gordon, Y.J., et al., *Human cathelicidin (LL-37), a multifunctional peptide, is expressed by ocular surface epithelia and has potent antibacterial and antiviral activity.* 2005. **30**(5): p. 385-394.
222. Robinson Jr, W.E., et al., *Anti-HIV-1 activity of indolicidin, an antimicrobial peptide from neutrophils.* 1998. **63**(1): p. 94-100.
223. Van der Strate, B., et al., *Antiviral activities of lactoferrin.* 2001. **52**(3): p. 225-239.
224. Rothan, H.A., et al., *Protegrin-1 inhibits dengue NS2B-NS3 serine protease and viral replication in MK2 cells.* 2012. **2012**.
225. Mahendran, A.S.K., et al., *The potential of antiviral peptides as COVID-19 therapeutics.* 2020. **11**: p. 575444.
226. Li, Q., et al., *Virucidal activity of a scorpion venom peptide variant mucroporin-M1 against measles, SARS-CoV and influenza H5N1 viruses.* 2011. **32**(7): p. 1518-1525.
227. Xia, S., et al., *Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein.* 2020. **17**(7): p. 765-767.
228. Xia, S., et al., *A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike.* 2019. **5**(4): p. eaav4580.
229. Channappanavar, R., et al., *Protective effect of intranasal regimens containing peptidic Middle East respiratory syndrome coronavirus fusion inhibitor against MERS-CoV infection.* 2015. **212**(12): p. 1894-1903.

230. Zhao, H., et al., *A novel peptide with potent and broad-spectrum antiviral activities against multiple respiratory viruses.* 2016. **6**(1): p. 1-13.
231. Wang, C., et al., *Lectin-like intestinal defensin inhibits 2019-nCoV spike binding to ACE2.* 2020.
232. Wohlford-Lenane, C.L., et al., *Rhesus theta-defensin prevents death in a mouse model of severe acute respiratory syndrome coronavirus pulmonary disease.* 2009. **83**(21): p. 11385-11390.
233. Sinthuvanich, C., et al., *Anticancer β-Hairpin Peptides: Membrane-Induced Folding Triggers Activity.* Journal of the American Chemical Society, 2012. **134**(14): p. 6210-6217.
234. van Zoggel, H., et al., *Antitumor and Angiostatic Activities of the Antimicrobial Peptide Dermaseptin B2.* PLOS ONE, 2012. **7**(9): p. e44351.
235. Wang, Y.-s., et al., *Intratumoral Expression of Mature Human Neutrophil Peptide-1 Mediates Antitumor Immunity in MiceHNP1 Mediates Tumor Immunity In situ.* 2009. **15**(22): p. 6901-6911.
236. Leuschner, C. and W. Hansel, *Targeting Breast and Prostate Cancers Through Their Hormone Receptors1.* Biology of Reproduction, 2005. **73**(5): p. 860-865.
237. Ourth, D.D., *Antitumor cell activity in vitro by myristoylated-peptide.* Biomedicine & Pharmacotherapy, 2011. **65**(4): p. 271-274.
238. Koskimaki, J.E., et al., *Peptides derived from type IV collagen, CXC chemokines, and thrombospondin-1 domain-containing proteins inhibit neovascularization and suppress tumor growth in MDA-MB-231 breast cancer xenografts.* Neoplasia, 2009. **11**(12): p. 1285-91.
239. Gaspar, D., A.S. Veiga, and M.A.R.B. Castanho, *From antimicrobial to anticancer peptides. A review.* 2013. **4**.
240. Aghazadeh, H., et al., *The activity and action mechanism of novel short selective LL-37-derived anticancer peptides against clinical isolates of Escherichia coli.* 2019. **93**(1): p. 75-83.
241. Fruitwala, S., D.W. El-Naccache, and T.L. Chang. *Multifaceted immune functions of human defensins and underlying mechanisms*. in *Seminars in cell & developmental biology*. 2019. Elsevier.
242. Liu, S., et al., *Linear analogues of human β-defensin 3: concepts for design of antimicrobial peptides with reduced cytotoxicity to mammalian cells.* 2008. **9**(6): p. 964-973.
243. Zweytick, D.J.C.S., *LTX-315–a promising novel antitumor peptide and immunotherapeutic agent.* 2019. **3**(11): p. 328.
244. Jeyamogan, S., et al., *Sera/organ lysates of selected animals living in polluted environments exhibit cytotoxicity against cancer cell lines.* 2019. **19**(18): p. 2251-2268.
245. Siddiqua, A., et al., *Venom proteins; prospects for anticancer therapy.* 2019. **52**(2): p. 68-78.
246. Brady, D., et al., *Insect cecropins, antimicrobial peptides with potential therapeutic applications.* 2019. **20**(23): p. 5862.
247. Pinto, I.B., et al., *Utilization of antimicrobial peptides, analogues and mimics in creating antimicrobial surfaces and bio-materials.* 2019. **150**: p. 107237.
248. Zahedifard, F., et al., *Anti-leishmanial activity of Brevinin 2R and its Lauric acid conjugate type against L. major: In vitro mechanism of actions and in vivo treatment potentials.* 2019. **13**(2): p. e0007217.
249. Li, B., et al., *LFB: a novel antimicrobial brevinin-like peptide from the skin secretion of the Fujian large headed frog, Limnonectes fujianensi.* 2019. **9**(6): p. 242.
250. Liu, Y., O. Tavana, and W.J.J.o.m.c.b. Gu, *p53 modifications: exquisite decorations of the powerful guardian.* 2019. **11**(7): p. 564-577.
251. Chen, X., et al., *A novel antimicrobial peptide, Ranatuerin-2PLx, showing therapeutic potential in inhibiting proliferation of cancer cells.* 2018. **38**(6).
252. Tornesello, A.L., et al., *Antimicrobial peptides as anticancer agents: functional properties and biological activities.* 2020. **25**(12): p. 2850.
253. Hansen, I.K., et al., *Isolation and characterization of antimicrobial peptides with unusual disulfide connectivity from the colonial ascidian Synoicum turgens.* 2020. **18**(1): p. 51.
254. Tripathi, A.K., et al., *Identification of GXXXXG motif in Chrysophsin-1 and its implication in the design of analogs with cell-selective antimicrobial and anti-endotoxin activities.* 2017. **7**(1): p. 1-16.
255. Matsuzaki, K.J.A.P., *Membrane permeabilization mechanisms.* 2019: p. 9-16.
256. Kumar, P., J.N. Kizhakkedathu, and S.K.J.B. Straus, *Antimicrobial peptides: diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo.* 2018. **8**(1): p. 4.

257. Pires, J., et al., *In vitro activity of the novel antimicrobial peptide dendrimer G3KL against multidrug-resistant Acinetobacter baumannii and Pseudomonas aeruginosa*. 2015. **59**(12): p. 7915-7918.
258. Rajamuthiah, R., et al., *A defensin from the model beetle Tribolium castaneum acts synergistically with telavancin and daptomycin against multidrug resistant Staphylococcus aureus.* 2015. **10**(6): p. e0128576.
259. Feng, X., et al., *The human antimicrobial peptide LL-37 and its fragments possess both antimicrobial and antibiofilm activities against multidrug-resistant Acinetobacter baumannii.* 2013. **49**: p. 131-137.
260. Luz, C., et al., *Antimicrobial packaging based on ε-polylysine bioactive film for the control of mycotoxigenic fungi in vitro and in bread.* 2018. **42**(1): p. e13370.
261. Elsser-Gravesen, D., A.J.B.o.f. Elsser-Gravesen, and f. additives, *Biopreservatives.* 2013: p. 29-49.
262. Miltz, J., et al., *Potency evaluation of a dermaseptin S4 derivative for antimicrobial food packaging applications.* 2006. **19**(6): p. 345-354.
263. Józefiak, D., et al., *Dietary nisin modulates the gastrointestinal microbial ecology and enhances growth performance of the broiler chickens.* 2013. **8**(12): p. e85347.
264. Cutler, S.A., et al., *Dietary inclusion of colicin e1 is effective in preventing postweaning diarrhea caused by F18-positive Escherichia coli in pigs.* 2007. **51**(11): p. 3830-3835.
265. Tang, Z., et al., *Effects of dietary supplementation with an expressed fusion peptide bovine lactoferricin–lactoferrampin on performance, immune function and intestinal mucosal morphology in piglets weaned at age 21 d.* 2008. **101**(7): p. 998-1005.
266. Wu, S., et al., *Effects of the antimicrobial peptide cecropin AD on performance and intestinal health in weaned piglets challenged with Escherichia coli.* 2012. **35**(2): p. 225-230.
267. Zhang, J., et al., *Mammary gland expression of antibacterial peptide genes to inhibit bacterial pathogens causing mastitis.* 2007. **90**(11): p. 5218-5225.
268. Yin, Z.-X., et al., *Cloning, expression and antimicrobial activity of an antimicrobial peptide, epinecidin-1, from the orange-spotted grouper, Epinephelus coioides.* 2006. **253**(1-4): p. 204-211.
269. Chiou, P.P., et al., *Effect of cecropin B and a synthetic analogue on propagation of fish viruses in vitro.* 2002. **4**(3): p. 294-302.
270. Shan, Z., et al., *NKL-24: a novel antimicrobial peptide derived from zebrafish NK-lysin that inhibits bacterial growth and enhances resistance against Vibrio parahaemolyticus infection in yesso scallop, Patinopecten yessoensis.* 2020. **106**: p. 431-440.
271. Hu, B., et al., *EmPis-1L, an effective antimicrobial peptide against the antibiotic-resistant VBNC state cells of pathogenic bacteria.* 2019. **11**(2): p. 667-675.
272. Ferre, R., et al., *Inhibition of plant-pathogenic bacteria by short synthetic cecropin A-melittin hybrid peptides.* 2006. **72**(5): p. 3302-3308.
273. Schaefer, S.C., et al., *Enhanced resistance to early blight in transgenic tomato lines expressing heterologous plant defense genes.* 2005. **222**(5): p. 858-866.
274. Li, X., et al., *Effects of the peptide H-OOWW-NH2 and its derived lipopeptide C12-OOWW-NH2 on controlling of citrus postharvest green mold.* 2019. **158**: p. 110979.
275. Shi, G., et al., *DRAMP 3.0: an enhanced comprehensive data repository of antimicrobial peptides.* 2022. **50**(D1): p. D488-D496.
276. Zhang, Z., et al., *A small peptide with therapeutic potential for inflammatory acne vulgaris.* 2013. **8**(8): p. e72923.
277. Mwangi, J., et al., *The antimicrobial peptide ZY4 combats multidrug-resistant Pseudomonas aeruginosa and Acinetobacter baumannii infection.* 2019. **116**(52): p. 26516-26522.
278. Torcato, I.M., et al., *Design and characterization of novel antimicrobial peptides, R-BP100 and RW-BP100, with activity against Gram-negative and Gram-positive bacteria.* 2013. **1828**(3): p. 944-955.
279. Lee, J. and D. Lee, *Structure-antimicrobial activity relationship between pleurocidin and its enantiomer.* Experimental & molecular medicine, 2008. **40**: p. 370-6.
280. Crusca Jr, E., et al., *Influence of N-terminus modifications on the biological activity, membrane interaction, and secondary structure of the antimicrobial peptide hylin-a1.* 2011. **96**(1): p. 41-48.
281. Veronese, F.M.J.B., *Peptide and protein PEGylation: a review of problems and solutions.* 2001. **22**(5): p. 405-417.

282. Rai, A., et al., *One-step synthesis of high-density peptide-conjugated gold nanoparticles with antimicrobial efficacy in a systemic infection model.* 2016. **85**: p. 99-110.
283. Jiang, Z., et al., *Cell surface-based sensing with metallic nanoparticles.* 2015. **44**(13): p. 4264-4274.
284. Casciaro, B., et al., *Gold-Nanoparticles coated with the antimicrobial peptide Esculentin-1a(1-21)NH2 as a reliable strategy for antipseudomonal drugs.* Acta Biomaterialia, 2016. **47**.
285. Bechinger, B., et al., *Revealing the mechanisms of synergistic action of two magainin antimicrobial peptides.* 2020. **2**: p. 615494.
286. Desbois, A.P., C.G. Gemmell, and P.J.J.I.j.o.a.a. Coote, *In vivo efficacy of the antimicrobial peptide ranalexin in combination with the endopeptidase lysostaphin against wound and systemic meticillin-resistant Staphylococcus aureus (MRSA) infections.* 2010. **35**(6): p. 559-565.
287. Jangra, M., V. Raka, and H.J.M. Nandanwar, *In vitro evaluation of antimicrobial peptide Tridecaptin M in combination with other antibiotics against multidrug resistant Acinetobacter baumannii.* 2020. **25**(14): p. 3255.
288. Zaïri, A., et al., *In vitro activities of dermaseptins K4S4 and K4K20S4 against Escherichia coli, Staphylococcus aureus, and Pseudomonas aeruginosa planktonic growth and biofilm formation.* 2014. **58**(4): p. 2221-2228.
289. Yin, X., et al., *Bacteriocin biosynthesis contributes to the anti-inflammatory capacities of probiotic Lactobacillus plantarum.* 2018. **9**(2): p. 333-344.
290. Oo, T.Z., et al., *Evaluation of synergistic activity of bovine lactoferricin with antibiotics in corneal infection.* 2010. **65**(6): p. 1243-1251.
291. Field, D., et al., *Synergistic Nisin-Polymyxin Combinations for the Control of Pseudomonas Biofilm Formation.* Front Microbiol, 2016. **7**: p. 1713.
292. Jahangiri, A., et al., *Synergistic effect of two antimicrobial peptides, Nisin and P10 with conventional antibiotics against extensively drug-resistant Acinetobacter baumannii and colistin-resistant Pseudomonas aeruginosa isolates.* Microb Pathog, 2021. **150**: p. 104700.
293. Portelinha, J. and A.M. Angeles-Boza, *The Antimicrobial Peptide Gad-1 Clears Pseudomonas aeruginosa Biofilms under Cystic Fibrosis Conditions.* Chembiochem, 2021. **22**(9): p. 1646-1655.
294. Lee, H.T., et al., *A large-scale structural classification of antimicrobial peptides.* Biomed Res Int, 2015. **2015**: p. 475062.
295. Handelsman, J., et al., *Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products.* 1998. **5**(10): p. R245-R249.
296. Daniel, R.J.N.R.M., *The metagenomics of soil.* 2005. **3**(6): p. 470-478.
297. Torsvik, V., et al., *High diversity in DNA of soil bacteria.* 1990. **56**(3): p. 782-787.
298. Sekse, C., et al., *High throughput sequencing for detection of foodborne pathogens.* 2017. **8**: p. 2029.
299. Coughlan, L.M., et al., *Biotechnological applications of functional metagenomics in the food and pharmaceutical industries.* Front Microbiol, 2015. **6**: p. 672.
300. Richardson, T.H., et al., *A novel, high performance enzyme for starch liquefaction: discovery and optimization of a low pH, thermostable α-amylase.* 2002. **277**(29): p. 26501-26507.
301. Ferrer, M., et al., *Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora.* 2005. **7**(12): p. 1996-2010.
302. Voget, S., et al., *Prospecting for novel biocatalysts in a soil metagenome.* 2003. **69**(10): p. 6235-6242.
303. Walter, J., et al., *Construction, analysis, and β-glucanase screening of a bacterial artificial chromosome library from the large-bowel microbiota of mice.* 2005. **71**(5): p. 2347-2354.
304. Jiang, C., et al., *Characterization of a novel β-glucosidase-like activity from a soil metagenome.* 2009. **47**(5): p. 542-548.
305. Palackal, N., et al., *A multifunctional hybrid glycosyl hydrolase discovered in an uncultured microbial consortium from ruminant gut.* 2007. **74**(1): p. 113-124.
306. Wang, K., et al., *Enzymatic synthesis of Galacto-oligosaccharides in an organic–aqueous biphasic system by a novel β-Galactosidase from a metagenomic library.* 2012. **60**(15): p. 3940-3946.
307. Cheng, F., et al., *Novel xylanase from a holstein cattle rumen metagenomic library and its application in xylooligosaccharide and ferulic acid production from wheat straw.* 2012. **60**(51): p. 12516-12524.
308. Pushpam, P.L., T. Rajesh, and P.J.A.e. Gunasekaran, *Identification and characterization of alkaline serine protease from goat skin surface metagenome.* 2011. **1**(1): p. 1-10.

309. Knietsch, A., et al., *Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on Escherichia coli.* 2003. **69**(3): p. 1408-1416.
310. Uchiyama, T., K.J.A. Miyazaki, and e. microbiology, *Product-induced gene expression, a product-responsive reporter assay used to screen metagenomic libraries for enzyme-encoding genes.* 2010. **76**(21): p. 7029-7035.
311. Faoro, H., et al., *Characterization of a new Acidobacteria-derived moderately thermostable lipase from a Brazilian Atlantic Forest soil metagenome.* 2012. **81**(2): p. 386-394.
312. Gupta, R., et al., *Bacterial alkaline proteases: molecular approaches and industrial applications.* 2002. **59**(1): p. 15-32.
313. Cottrell, M.T., et al., *Chitinases from uncultured marine microorganisms.* 1999. **65**(6): p. 2553-2557.
314. Knietsch, A., et al., *Identification and characterization of coenzyme B12-dependent glycerol dehydratase-and diol dehydratase-encoding genes from metagenomic DNA libraries derived from enrichment cultures.* 2003. **69**(6): p. 3048-3060.
315. Biver, S., M.J.J.o.I.M. Vandenbol, and Biotechnology, *Characterization of three new carboxylic ester hydrolases isolated by functional screening of a forest soil metagenomic library.* 2013. **40**(2): p. 191-200.
316. Robertson, D.E., et al., *Exploring nitrilase sequence space for enantioselective catalysis.* 2004. **70**(4): p. 2429-2436.
317. Yao, J., et al., *Isolation and characterization of a novel tannase from a metagenomic library.* 2011. **59**(8): p. 3812-3818.
318. Lee, D.-G., et al., *Screening and characterization of a novel fibrinolytic metalloprotease from a metagenomic library.* 2007. **29**(3): p. 465-472.
319. Rabausch, U., et al., *Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes.* 2013. **79**(15): p. 4551-4563.
320. Piel, J.J.P.o.t.N.A.o.S., *A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles.* 2002. **99**(22): p. 14002-14007.
321. Entcheva, P., et al., *Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia.* 2001. **67**(1): p. 89-99.
322. Fujita, M.J., et al., *Cloning and heterologous expression of the vibrioferrin biosynthetic gene cluster from a marine metagenomic library.* 2011. **75**(12): p. 2283-2287.
323. Schirmer, A., et al., *Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge Discodermia dissoluta.* 2005. **71**(8): p. 4840-4849.
324. Jiang, C.-J., et al., *Characterization of a novel serine protease inhibitor gene from a marine metagenome.* 2011. **9**(9): p. 1487-1501.
325. Chang, F.-Y. and S.F.J.P.o.t.N.A.o.S. Brady, *Discovery of indolotryptoline antiproliferative agents by homology-guided metagenomic screening.* 2013. **110**(7): p. 2478-2483.
326. Lakhdari, O., et al., *Functional metagenomics: a high throughput screening method to decipher microbiota-driven NF-κB modulation in the human gut.* 2010. **5**(9): p. e13092.
327. Culligan, E.P., et al., *Functional environmental screening of a metagenomic library identifies stlA; a unique salt tolerance locus from the human gut microbiome.* 2013. **8**(12): p. e82985.
328. Guazzaroni, M.E., et al., *Novel acid resistance genes from the metagenome of the T into R iver, an extremely acidic environment.* 2013. **15**(4): p. 1088-1102.
329. Bashir, Y., S. Pradeep Singh, and B. Kumar Konwar, *Metagenomics: An Application Based Perspective*. Chinese Journal of Biology, 2014. **2014**: p. 146030.
330. Allen, H.K., et al., *Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil.* 2009. **3**(2): p. 243-251.
331. Feng, Z., et al., *Environmental DNA-encoded antibiotics fasamycins A and B inhibit FabF in type II fatty acid biosynthesis.* 2012. **134**(6): p. 2981-2987.
332. Lim, H.K., et al., *Characterization of a forest soil metagenome clone that confers indirubin and indigo production on Escherichia coli.* 2005. **71**(12): p. 7768-7777.
333. Wang, G.-Y.-S., et al., *Novel Natural Products from Soil DNA Libraries in a Streptomycete Host.* Organic Letters, 2000. **2**(16): p. 2401-2404.

334. Gillespie, D.E., et al., *Isolation of Antibiotics Turbomycin A and B from a Metagenomic Library of Soil Microbial DNA.* 2002. **68**(9): p. 4301-4306.
335. Feng, Z., et al., *Environmental DNA-Encoded Antibiotics Fasamycins A and B Inhibit FabF in Type II Fatty Acid Biosynthesis.* Journal of the American Chemical Society, 2012. **134**(6): p. 2981-2987.
336. Lang, K.S., et al., *Novel florfenicol and chloramphenicol resistance gene discovered in Alaskan soil by using functional metagenomics.* Appl Environ Microbiol, 2010. **76**(15): p. 5321-6.
337. Donato, J.J., et al., *Metagenomic analysis of apple orchard soil reveals antibiotic resistance genes encoding predicted bifunctional proteins.* Appl Environ Microbiol, 2010. **76**(13): p. 4396-401.
338. Parsley, L.C., et al., *Identification of diverse antimicrobial resistance determinants carried on bacterial, plasmid, or viral metagenomes from an activated sludge microbial assemblage.* Appl Environ Microbiol, 2010. **76**(11): p. 3753-7.
339. Tao, W., et al., *Inactivation of chloramphenicol and florfenicol by a novel chloramphenicol hydrolase.* Appl Environ Microbiol, 2012. **78**(17): p. 6295-301.
340. Jeon, J.H., et al., *Novel metagenome-derived carboxylesterase that hydrolyzes β-lactam antibiotics.* Appl Environ Microbiol, 2011. **77**(21): p. 7830-6.
341. Martiny, A.C., et al., *Functional metagenomics reveals previously unrecognized diversity of antibiotic resistance genes in gulls.* Front Microbiol, 2011. **2**: p. 238.
342. McGarvey, K.M., K. Queitsch, and S. Fields, *Wide variation in antibiotic resistance proteins identified by functional metagenomic screening of a soil DNA library*. Appl Environ Microbiol, 2012. **78**(6): p. 1708-14.
343. Forsberg, K.J., et al., *The shared antibiotic resistome of soil bacteria and human pathogens.* Science, 2012. **337**(6098): p. 1107-11.
344. Su, J.Q., et al., *Functional metagenomic characterization of antibiotic resistance genes in agricultural soils from China.* Environ Int, 2014. **65**: p. 9-15.
345. Berman, H.F. and L.W. Riley, *Identification of novel antimicrobial resistance genes from microbiota on retail spinach.* BMC Microbiol, 2013. **13**: p. 272.
346. Devirgiliis, C., et al., *Functional screening of antibiotic resistance genes from a representative metagenomic library of food fermenting microbiota.* Biomed Res Int, 2014. **2014**: p. 290967.
347. Castañeda-Mogollón, D., et al., *A metagenomics workflow for SARS-CoV-2 identification, co-pathogen detection, and overall diversity.* J Clin Virol, 2021. **145**: p. 105025.
348. Amrane, S. and J.-C. Lagier, *Metagenomic and clinical microbiology.* Human Microbiome Journal, 2018. **9**: p. 1-6.
349. Nakamura, S., et al., *Metagenomic diagnosis of bacterial infections.* Emerg Infect Dis, 2008. **14**(11): p. 1784-6.
350. Wilson, M.R., et al., *Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing.* 2014. **370**(25): p. 2408-2417.
351. Mongkolrattanothai, K., et al., *Neurobrucellosis: Unexpected Answer From Metagenomic Next-Generation Sequencing.* Journal of the Pediatric Infectious Diseases Society, 2017. **6**(4): p. 393-398.
352. Salzberg, S.L., et al., *Next-generation sequencing in neuropathologic diagnosis of infections of the nervous system.* 2016. **3**(4): p. e251.
353. Yao, M., et al., *Detection of Listeria monocytogenes in CSF from Three Patients with Meningoencephalitis by Next-Generation Sequencing.* jcn, 2016. **12**(4): p. 446-451.
354. Thoendel, M., et al., *A Novel Prosthetic Joint Infection Pathogen, Mycoplasma salivarium, Identified by Metagenomic Shotgun Sequencing.* Clinical Infectious Diseases, 2017. **65**(2): p. 332-335.
355. Fukui, Y., et al., *Metagenomic analysis for detecting pathogens in culture-negative infective endocarditis.* Journal of Infection and Chemotherapy, 2015. **21**(12): p. 882-884.
356. Abril, M.K., et al., *Diagnosis of Capnocytophaga canimorsus Sepsis by Whole-Genome Next-Generation Sequencing.* Open Forum Infectious Diseases, 2016. **3**(3).
357. Ye, M., et al., *Rapid diagnosis of Propionibacterium acnes infection in patient with hyperpyrexia after hematopoietic stem cell transplantation by next-generation sequencing: a case report.* BMC Infectious Diseases, 2016. **16**(1): p. 5.
358. Pendleton, K.M., et al., *Rapid Pathogen Identification in Bacterial Pneumonia Using Real-Time Metagenomics.* Am J Respir Crit Care Med, 2017. **196**(12): p. 1610-1612.

359. Röling, W.F., *Maths on microbes: adding microbial ecophysiology to metagenomics.* Microb Biotechnol, 2015. **8**(1): p. 21-2.
360. Dellagnezze, B.M., et al., *Bioremediation potential of microorganisms derived from petroleum reservoirs.* Marine Pollution Bulletin, 2014. **89**(1): p. 191-200.
361. Williams, P.J., et al., *Effective bioreduction of hexavalent chromium-contaminated water in fixed-film bioreactors %J Water SA.* 2014. **40**: p. 549-554.
362. Yergeau, E., et al., *Metagenomic Analysis of the Bioremediation of Diesel-Contaminated Canadian High Arctic Soils.* PLOS ONE, 2012. **7**(1): p. e30058.
363. Sutton, N.B., et al., *Impact of Long-Term Diesel Contamination on Soil Microbial Community Structure.* 2013. **79**(2): p. 619-630.
364. Marchetti, A., et al., *Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability.* 2012. **109**(6): p. E317-E325.
365. Mason, O.U., et al., *Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill.* 2012. **6**(9): p. 1715-1727.
366. Maurice, C.F., H.J. Haiser, and P.J.J.C. Turnbaugh, *Xenobiotics shape the physiology and gene expression of the active human gut microbiome.* 2013. **152**(1-2): p. 39-50.
367. Kallmeyer, J., et al., *Global distribution of microbial abundance and biomass in subseafloor sediment.* 2012. **109**(40): p. 16213-16216.
368. Berg, C., et al., *Dissection of Microbial Community Functions during a Cyanobacterial Bloom in the Baltic Sea via Metatranscriptomics.* 2018. **5**.
369. Shi, Y., G.W. Tyson, and E.F. DeLong, *Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column.* Nature, 2009. **459**(7244): p. 266-269.
370. Moniruzzaman, M., et al., *Virus-host relationships of marine single-celled eukaryotes resolved from metatranscriptomics.* Nature Communications, 2017. **8**(1): p. 16054.
371. Mu, D.-S., et al., *Metatranscriptomic and comparative genomic insights into resuscitation mechanisms during enrichment culturing.* Microbiome, 2018. **6**(1): p. 230.
372. Mahapatra, G.P., et al., *Metagenomics Approaches in Discovery and Development of New Bioactive Compounds from Marine Actinomycetes.* Current Microbiology, 2020. **77**(4): p. 645-656.
373. Zhang, W., et al., *Marine Streptomyces sp. derived antimycin analogues suppress HeLa cells via depletion HPV E6/E7 mediated by ROS-dependent ubiquitin–proteasome system.* 2017. **7**(1): p. 1-14.
374. Xie, C.-L., et al., *Anti-allergic compounds from the deep-sea-derived actinomycete Nesterenkonia flava MCCC 1K00610.* 2017. **15**(3): p. 71.
375. Wu, G., et al., *Bonnevillamides, Linear Heptapeptides Isolated from a Great Salt Lake-Derived Streptomyces sp.* 2017. **15**(7): p. 195.
376. Li, H., et al., *Discovery of antimycin-type depsipeptides from a wbl gene mutant strain of deepsea-derived Streptomyces somaliensis SCSIO ZH66 and their effects on pro-inflammatory cytokine production.* 2017. **8**: p. 678.
377. Kim, J., et al., *Borrelidins C–E: New Antibacterial Macrolides from a Saltern-Derived Halophilic Nocardiopsis sp.* 2017. **15**(6): p. 166.
378. Xu, M.-J., et al., *Deciphering the streamlined genome of Streptomyces xiamenensis 318 as the producer of the anti-fibrotic drug candidate xiamenmycin.* 2016. **6**(1): p. 1-11.
379. Waters, A.L., et al., *An analysis of the sponge Acanthostrongylophora igens' microbiome yields an actinomycete that produces the natural product manzamine A.* 2014. **1**: p. 54.
380. Mangamuri, U., et al., *Bioactive metabolites produced by Streptomyces Cheonanensis VUK-A from Coringa mangrove sediments: isolation, structure elucidation and bioactivity.* 2016. **6**(1): p. 1-8.
381. Chen, Z., et al., *New α-glucosidase inhibitors from marine algae-derived Streptomyces sp. OUCMDZ-3434.* 2016. **6**(1): p. 1-9.
382. Zhou, T., et al., *Biosynthesis of akaeolide and lorneic acids and annotation of type I polyketide synthase gene clusters in the genome of Streptomyces sp. NPS554.* 2015. **13**(1): p. 581-596.
383. Zhen, X., et al., *A new analogue of echinomycin and a new cyclic dipeptide from a marine-derived Streptomyces sp. LS298.* 2015. **13**(11): p. 6947-6961.
384. Zhang, Y., et al., *Activation and enhancement of Fredericamycin A production in deepsea-derived Streptomyces somaliensis SCSIO ZH66 by using ribosome engineering and response surface methodology.* 2015. **14**(1): p. 1-11.

385. Vicente, J., et al., *Monacyclinones, new angucyclinone metabolites isolated from Streptomyces sp. M7_15 associated with the Puerto rican sponge Scopalina ruetzleri.* 2015. **13**(8): p. 4682-4700.
386. Huang, H., et al., *Activation of a plasmid-situated type III PKS gene cluster by deletion of a wbl gene in deepsea-derived Streptomyces somaliensis SCSIO ZH66.* 2016. **15**(1): p. 1-9.
387. Bekiesch, P., P. Basitta, and A.K.J.A.d.P. Apel, *Challenges in the heterologous production of antibiotics in Streptomyces.* 2016. **349**(8): p. 594-601.
388. Kimura, N., *Novel biological resources screened from uncultured bacteria by a metagenomic method*, in *Metagenomics*. 2018, Elsevier. p. 273-288.
389. Martínez-García, E., et al., *Pseudomonas 2.0: genetic upgrading of P. putida KT2440 as an enhanced host for heterologous gene expression.* 2014. **13**(1): p. 1-15.
390. Liu, N., et al., *Mycemycins A–E, new dibenzoxazepinones isolated from two different Streptomycetes.* 2015. **13**(10): p. 6247-6258.
391. Schulze, C.J., et al., *Genome-directed lead discovery: biosynthesis, structure elucidation, and biological evaluation of two families of polyene macrolactams against Trypanosoma brucei.* 2015. **10**(10): p. 2373-2381.
392. Schulze, C.J., et al., *Salinipostins A–K, long-chain bicyclic phosphotriesters as a potent and selective antimalarial chemotype.* 2015. **80**(3): p. 1312-1320.
393. Mitchell, A.L., et al., *MGnify: the microbiome analysis resource in 2020.* Nucleic Acids Research, 2019. **48**(D1): p. D570-D578.
394. Jones, D.T.J.J.o.m.b., *Protein secondary structure prediction based on position-specific scoring matrices.* 1999. **292**(2): p. 195-202.
395. Buchan, D.W. and D.T.J.N.a.r. Jones, *The PSIPRED protein analysis workbench: 20 years on.* 2019. **47**(W1): p. W402-W407.
396. Blum, M., et al., *The InterPro protein families and domains database: 20 years on.* 2021. **49**(D1): p. D344-D354.
397. Lamiable, A., et al., *PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex.* 2016. **44**(W1): p. W449-W454.
398. Shen, Y., et al., *Improved PEP-FOLD approach for peptide and miniprotein structure prediction.* 2014. **10**(10): p. 4745-4758.
399. Thevenet, P., et al., *PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides.* 2012. **40**(W1): p. W288-W293.
400. Gasteiger, E., et al., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor. 2005, Humana Press: Totowa, NJ. p. 571-607.
401. Abdel Monaim, S.A.H., et al., *Converting Teixobactin into a Cationic Antimicrobial Peptide (AMP).* J Med Chem, 2017. **60**(17): p. 7476-7482.
402. E Greber, K. and M.J.C.t.i.m.c. Dawgul, *Antimicrobial peptides under clinical trials.* 2017. **17**(5): p. 620-628.
403. Dutta, P. and S. Das, *Mammalian Antimicrobial Peptides: Promising Therapeutic Targets Against Infection and Chronic Inflammation.* Curr Top Med Chem, 2016. **16**(1): p. 99-129.
404. Wu, Q., J. Patočka, and K. Kuča, *Insect Antimicrobial Peptides, a Mini Review.* Toxins (Basel), 2018. **10**(11).
405. Wang, G., *Bioinformatic Analysis of 1000 Amphibian Antimicrobial Peptides Uncovers Multiple Length-Dependent Correlations for Peptide Design and Prediction.* Antibiotics (Basel), 2020. **9**(8).
406. Thomma, B.P., B.P. Cammue, and K.J.P. Thevissen, *Plant defensins.* 2002. **216**(2): p. 193-202.

Supplementary Table 1: AMP prediction result of HGD0_cDNA

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
1	AMP	NAMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	NAMP
3		NAMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	NAMP
5		NAMP	NAMP	NAMP
6	NAMP	NAMP	AMP	NAMP
7	NAMP	NAMP	AMP	NAMP
8	NAMP	NAMP	NAMP	AMP
9	NAMP	NAMP	NAMP	NAMP
10		NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	NAMP
14	AMP	NAMP	NAMP	NAMP
15		NAMP	NAMP	NAMP
16	AMP	NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	NAMP
Total number of AMPs	3	0	2	1

Supplementary Table 2: AMP prediction result of HGD1_cDNA

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	AMP	AMP
2	NAMP	AMP	NAMP	NAMP
3	NAMP	NAMP	NAMP	NAMP
4		NAMP	NAMP	AMP
5	NAMP	NAMP	NAMP	NAMP
6	NAMP	NAMP	NAMP	AMP
7	NAMP	AMP	NAMP	AMP
8	NAMP	NAMP	NAMP	NAMP
9	NAMP	NAMP	AMP	NAMP
10	NAMP	NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	AMP
13	NAMP	NAMP	NAMP	AMP
14	NAMP	NAMP	NAMP	NAMP
15	NAMP	NAMP	NAMP	NAMP

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
16		NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	NAMP
18	NAMP	NAMP	NAMP	AMP
19	NAMP	NAMP	NAMP	NAMP
20		NAMP	NAMP	NAMP
21	NAMP	AMP	AMP	NAMP
22	NAMP	NAMP	AMP	NAMP
23	NAMP	NAMP	NAMP	NAMP
24	NAMP	NAMP	NAMP	AMP
25	NAMP	NAMP	NAMP	NAMP
26	NAMP	AMP	NAMP	NAMP
27	NAMP	AMP	NAMP	NAMP
28	NAMP	NAMP	AMP	NAMP
29	NAMP	NAMP	NAMP	AMP
30	NAMP	NAMP	NAMP	NAMP
31		NAMP	AMP	AMP
32	NAMP	NAMP	NAMP	NAMP
33	NAMP	AMP	AMP	AMP
34	NAMP	NAMP	NAMP	AMP
35	AMP	NAMP	NAMP	NAMP
36	AMP	NAMP	AMP	NAMP
37	NAMP	NAMP	NAMP	NAMP
38	AMP	AMP	NAMP	AMP
39	NAMP	NAMP	NAMP	NAMP
40	NAMP	NAMP	NAMP	NAMP
41	NAMP	AMP	NAMP	AMP
42	NAMP	NAMP	NAMP	NAMP
43	NAMP	NAMP	NAMP	AMP
44	NAMP	NAMP	NAMP	AMP
45	NAMP	NAMP	NAMP	AMP
46	NAMP	NAMP	NAMP	NAMP
47	NAMP	NAMP	AMP	AMP
48	AMP	NAMP	NAMP	NAMP
49		NAMP	NAMP	NAMP
50	AMP	AMP	AMP	AMP
51	NAMP	NAMP		AMP
52		NAMP		NAMP
53	NAMP	NAMP		NAMP

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
54	NAMP	NAMP		AMP
55		NAMP		NAMP
56	NAMP	NAMP		AMP
Total number of AMPs	5	9	10	22

Supplementary Table 3: AMP prediction result of HGD2_cDNA

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
1		NAMP	NAMP	NAMP
2	NAMP	AMP	AMP	NAMP
3	NAMP	NAMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	NAMP
5	NAMP	NAMP	NAMP	NAMP
6	NAMP	NAMP	NAMP	NAMP
7	NAMP	NAMP	NAMP	NAMP
8	NAMP	NAMP	NAMP	NAMP
9	NAMP	NAMP	NAMP	NAMP
10	NAMP	AMP	AMP	NAMP
11	AMP	NAMP	NAMP	AMP
12		NAMP	AMP	NAMP
13	NAMP	NAMP	NAMP	NAMP
14		NAMP	NAMP	NAMP
15	NAMP	NAMP	NAMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	AMP	NAMP	NAMP	NAMP
18		NAMP	NAMP	NAMP
19	NAMP	NAMP	NAMP	NAMP
20	NAMP	NAMP	NAMP	NAMP
21	NAMP	NAMP	NAMP	NAMP
22	NAMP	NAMP	NAMP	AMP
23	NAMP	NAMP	NAMP	NAMP
24	NAMP	NAMP	AMP	NAMP
25	NAMP	NAMP	NAMP	NAMP
26	AMP	NAMP	NAMP	NAMP
27	NAMP	NAMP	AMP	NAMP
28	NAMP	NAMP	NAMP	NAMP
Total number of AMPs	3	2	5	2

Supplementary Table 4: AMP prediction result of HGD3_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	NAMP	AMP	NAMP	NAMP
2	NAMP	AMP	NAMP	NAMP
3	NAMP	NAMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	NAMP
5	NAMP	NAMP	AMP	NAMP
6	AMP	AMP	AMP	NAMP
7	AMP	NAMP	NAMP	NAMP
8	NAMP	NAMP	NAMP	NAMP
9	NAMP	NAMP	NAMP	NAMP
10	NAMP	NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	NAMP
14	NAMP	NAMP	NAMP	NAMP
15	NAMP	NAMP	NAMP	NAMP
16	AMP	NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	NAMP
18	NAMP	NAMP	AMP	NAMP
19	NAMP	NAMP	AMP	NAMP
Total number of AMPs	3	3	4	0

Supplementary Table 5: AMP prediction result of HGD4_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	NAMP
3	NAMP	NAMP	NAMP	NAMP
4		NAMP	NAMP	AMP
5	AMP	AMP	AMP	NAMP
6		NAMP	NAMP	NAMP
7	NAMP	NAMP	NAMP	NAMP
8	AMP	AMP	AMP	NAMP
9	NAMP	NAMP	NAMP	NAMP
10	NAMP	NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
12	NAMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	NAMP
14		NAMP	NAMP	NAMP
15	AMP	AMP	NAMP	NAMP
Total number of AMPs	3	3	2	1

Supplementary Table 6: AMP prediction result of S0_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	NAMP	AMP
2	NAMP	NAMP	NAMP	AMP
3	NAMP	NAMP	NAMP	AMP
4	NAMP	NAMP	NAMP	NAMP
5	AMP	NAMP	NAMP	NAMP
6	AMP	NAMP	NAMP	NAMP
7	NAMP	NAMP	AMP	NAMP
8	NAMP	NAMP	AMP	NAMP
9		NAMP	NAMP	NAMP
10		NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	AMP
13	NAMP	NAMP	NAMP	AMP
14	NAMP	NAMP	NAMP	NAMP
15	AMP	NAMP	AMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	NAMP
18	NAMP	NAMP	NAMP	NAMP
19	AMP	NAMP	NAMP	NAMP
20		NAMP	NAMP	NAMP
21	NAMP	NAMP	NAMP	NAMP
22		NAMP	NAMP	NAMP
23	NAMP	NAMP	NAMP	NAMP
24	NAMP	NAMP	AMP	NAMP
25	NAMP	NAMP	NAMP	AMP
26	AMP	NAMP	NAMP	NAMP
27	NAMP	NAMP	NAMP	NAMP
28	NAMP	NAMP	NAMP	NAMP
29		NAMP	NAMP	NAMP

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
30	AMP	NAMP	NAMP	NAMP
31	NAMP	NAMP	NAMP	AMP
32	NAMP	NAMP	NAMP	NAMP
33	NAMP	NAMP	NAMP	AMP
34	NAMP	NAMP	NAMP	AMP
35	AMP	NAMP	NAMP	NAMP
36	NAMP	NAMP	NAMP	NAMP
37	NAMP	NAMP	NAMP	NAMP
38	NAMP	NAMP	AMP	AMP
39		NAMP	NAMP	NAMP
40	NAMP	NAMP	NAMP	AMP
41	NAMP	NAMP	NAMP	NAMP
42	AMP	NAMP	NAMP	NAMP
43		NAMP	NAMP	NAMP
44	NAMP	NAMP	NAMP	NAMP
45	AMP	NAMP	NAMP	NAMP
46	NAMP	NAMP	NAMP	NAMP
47		NAMP	NAMP	NAMP
Total number of AMPs	9	0	5	10

Supplementary Table 7: AMP prediction result of S1_cDNA

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	NAMP
3		NAMP	NAMP	NAMP
4		NAMP	NAMP	NAMP
5	NAMP	AMP	NAMP	AMP
6	NAMP	NAMP	NAMP	AMP
7	NAMP	NAMP	NAMP	AMP
8	AMP	AMP	AMP	AMP
9	NAMP	NAMP	AMP	AMP
10	AMP	NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	AMP
12	AMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	AMP
14	NAMP	NAMP	NAMP	AMP

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
15		NAMP	NAMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	AMP
18	AMP	NAMP	NAMP	NAMP
19	NAMP	NAMP	NAMP	NAMP
20	NAMP	NAMP	NAMP	NAMP
21	AMP	NAMP	AMP	NAMP
22	NAMP	NAMP	AMP	NAMP
23	NAMP	NAMP	NAMP	NAMP
24	NAMP	NAMP	NAMP	NAMP
25	NAMP	NAMP	NAMP	NAMP
26	NAMP	NAMP	NAMP	NAMP
27	NAMP	NAMP	NAMP	NAMP
28	AMP	NAMP	NAMP	NAMP
29	NAMP	NAMP	NAMP	NAMP
30	NAMP	NAMP	NAMP	AMP
31	NAMP	NAMP	NAMP	NAMP
32	NAMP	NAMP	AMP	NAMP
33	NAMP	NAMP	NAMP	NAMP
34	NAMP	AMP	AMP	AMP
35	NAMP	NAMP	NAMP	NAMP
36	NAMP	NAMP	NAMP	NAMP
37	AMP	NAMP	NAMP	NAMP
38	NAMP	NAMP	NAMP	NAMP
39		NAMP	NAMP	NAMP
40	AMP	NAMP	NAMP	NAMP
41	AMP	NAMP	NAMP	NAMP
Total number of AMPs	9	3	6	11

Supplementary Table 8: AMP prediction result of S2_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	NAMP	NAMP
2	NAMP	AMP	AMP	AMP
3	NAMP	NAMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	NAMP
5	NAMP	NAMP	NAMP	AMP

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
6	NAMP	NAMP	NAMP	NAMP
7	AMP	NAMP	AMP	NAMP
8	NAMP	NAMP	NAMP	NAMP
9	AMP	NAMP	NAMP	NAMP
10	NAMP	NAMP	NAMP	AMP
11	NAMP	NAMP	NAMP	NAMP
12	AMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	NAMP
14	NAMP	NAMP	NAMP	NAMP
15	AMP	AMP	AMP	AMP
16	AMP	AMP	NAMP	NAMP
17	NAMP	AMP	NAMP	NAMP
18	NAMP	NAMP	NAMP	NAMP
19		NAMP	NAMP	NAMP
20	NAMP	NAMP	NAMP	AMP
21	AMP	NAMP	NAMP	NAMP
22	NAMP	NAMP	NAMP	AMP
23	NAMP	NAMP	NAMP	AMP
24	NAMP	AMP	NAMP	AMP
25	AMP	AMP	NAMP	NAMP
26	AMP	AMP	AMP	NAMP
27	NAMP	NAMP	NAMP	NAMP
28	NAMP	NAMP	NAMP	AMP
29	AMP	NAMP	NAMP	NAMP
30	NAMP	NAMP	NAMP	AMP
31	AMP	AMP	AMP	NAMP
Total number of AMPs	10	8	5	10

Supplementary Table 9: AMP prediction result of S3_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	AMP	NAMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	AMP
3	NAMP	NAMP	NAMP	NAMP
4	NAMP	AMP	NAMP	AMP
5	NAMP	NAMP	NAMP	NAMP
6	NAMP	NAMP	NAMP	NAMP

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
7	NAMP	NAMP	NAMP	NAMP
8	NAMP	AMP	AMP	AMP
9	NAMP	NAMP	NAMP	NAMP
10	NAMP	NAMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	AMP
13	NAMP	AMP	AMP	NAMP
14	NAMP	NAMP	NAMP	NAMP
15	AMP	AMP	AMP	NAMP
16	NAMP	NAMP	NAMP	AMP
17	AMP	NAMP	NAMP	NAMP
18	NAMP	NAMP	NAMP	NAMP
19	NAMP	NAMP	NAMP	NAMP
20	AMP	NAMP	AMP	NAMP
21	AMP	NAMP	NAMP	NAMP
22	AMP	NAMP	NAMP	NAMP
23	NAMP	NAMP	AMP	NAMP
24		NAMP	NAMP	NAMP
25	NAMP	AMP	NAMP	NAMP
26	NAMP	NAMP	NAMP	NAMP
27	NAMP	NAMP	NAMP	NAMP
28	AMP	NAMP	NAMP	AMP
Total number of AMPs	7	5	5	6

Supplementary Table 10: AMP prediction result of S4_cDNA

Sequence number	CAMPR 3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	NAMP	AMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	NAMP
3	AMP	AMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	NAMP
5	NAMP	NAMP	AMP	NAMP
6	NAMP	AMP	NAMP	NAMP
7	NAMP	NAMP	NAMP	NAMP
8	NAMP	NAMP	NAMP	NAMP
9	AMP	AMP	NAMP	NAMP
Total number of AMPs	2	4	1	0

Supplementary Table 11: AMP prediction result of XSD0_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	AMP	NAMP	NAMP	AMP
2	NAMP	NAMP	NAMP	AMP
3	AMP	NAMP	NAMP	AMP
4	NAMP	NAMP	NAMP	AMP
5	NAMP	NAMP	NAMP	NAMP
6	NAMP	NAMP	NAMP	AMP
7	NAMP	NAMP	NAMP	NAMP
8	NAMP	NAMP	NAMP	AMP
9	AMP	NAMP	NAMP	AMP
10	AMP	AMP	NAMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	AMP
13	NAMP	NAMP	NAMP	NAMP
14	NAMP	NAMP	NAMP	NAMP
15	NAMP	NAMP	NAMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	AMP
18	NAMP	NAMP	NAMP	AMP
19	AMP	NAMP	NAMP	NAMP
20	AMP	NAMP	NAMP	AMP
21	NAMP	AMP	NAMP	NAMP
22	NAMP	NAMP	AMP	NAMP
23	NAMP	NAMP	NAMP	AMP
24	NAMP	NAMP	AMP	NAMP
Total number of AMPs	6	2	2	12

Supplementary Table 12: AMP prediction result of XSD1_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	AMP	NAMP	AMP	NAMP
2	NAMP	AMP	NAMP	AMP
3	NAMP	NAMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	AMP
5	NAMP	NAMP	NAMP	NAMP
6		NAMP	NAMP	NAMP

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
7	NAMP	AMP	NAMP	NAMP
8	NAMP	NAMP	NAMP	NAMP
9	NAMP	NAMP	NAMP	NAMP
10	AMP	AMP	AMP	NAMP
11	NAMP	AMP	AMP	AMP
12	NAMP	NAMP	NAMP	AMP
13		NAMP	NAMP	NAMP
14	NAMP	NAMP	NAMP	AMP
15	NAMP	NAMP	NAMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	NAMP	NAMP	NAMP	AMP
18	AMP	NAMP	NAMP	NAMP
19	NAMP	NAMP	NAMP	NAMP
20	AMP	NAMP	NAMP	NAMP
21	AMP	NAMP	NAMP	AMP
22	NAMP	NAMP	NAMP	NAMP
23	NAMP	NAMP	NAMP	AMP
24	NAMP	NAMP	NAMP	NAMP
25	NAMP	NAMP	NAMP	AMP
26	NAMP	NAMP	NAMP	NAMP
27	NAMP	NAMP	NAMP	AMP
28		NAMP	NAMP	NAMP
Total number of AMPs	5	4	3	10

Supplementary Table 13: AMP prediction result of XSD2_cDNA

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	NAMP	AMP
2	AMP	NAMP	NAMP	NAMP
3		NAMP	NAMP	NAMP
4		NAMP	NAMP	NAMP
5	NAMP	NAMP	NAMP	NAMP
6	NAMP	NAMP	NAMP	NAMP
7		NAMP	NAMP	NAMP
8	NAMP	NAMP	AMP	NAMP
9	NAMP	AMP	AMP	AMP
10	NAMP	AMP	AMP	AMP
11	NAMP	NAMP	NAMP	NAMP

Sequence number	CAMPR3 (RF)	AMP scanner	AMPA	ADAM (HMM)
12	NAMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	AMP
14	NAMP	NAMP	NAMP	AMP
15	NAMP	NAMP	NAMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	NAMP	AMP	AMP	NAMP
18	NAMP	NAMP	NAMP	NAMP
19	AMP	AMP	AMP	NAMP
20	NAMP	NAMP	NAMP	AMP
Total number of AMPs	2	4	5	6

Supplementary Table 14: AMP prediction result of XSD3_cDNA

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1		NAMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	NAMP
3		NAMP	NAMP	NAMP
4	AMP	NAMP	NAMP	NAMP
5	AMP	NAMP	NAMP	NAMP
6	NAMP	NAMP	NAMP	NAMP
7	NAMP	NAMP	NAMP	NAMP
8		NAMP	NAMP	NAMP
9		NAMP	NAMP	NAMP
10	AMP	NAMP	NAMP	NAMP
11	AMP	NAMP	NAMP	NAMP
12	NAMP	NAMP	NAMP	NAMP
13		NAMP	NAMP	NAMP
14		NAMP	NAMP	NAMP
15	NAMP	NAMP	NAMP	NAMP
16		NAMP	NAMP	NAMP
17	NAMP	NAMP	AMP	NAMP
18	NAMP	AMP	NAMP	NAMP
19	NAMP	NAMP	NAMP	NAMP
20		NAMP	NAMP	NAMP
21	AMP	AMP	AMP	NAMP
22	AMP	NAMP	NAMP	AMP
23	NAMP	AMP	AMP	NAMP
24	NAMP	NAMP	AMP	NAMP

Sequence number	CAMPR3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
25	AMP	NAMP	NAMP	NAMP
26	AMP	NAMP	NAMP	NAMP
27	AMP	NAMP	AMP	NAMP
28	NAMP	NAMP	NAMP	NAMP
29		NAMP	NAMP	NAMP
Total number of AMPs	9	3	5	1

Supplementary Table 15: AMP prediction result of XSD4_cDNA

Sequence number	CAMPR 3 (RF)	AMP Scanner	AMPA	ADAM (HMM)
1	NAMP	NAMP	NAMP	NAMP
2	NAMP	NAMP	NAMP	NAMP
3	AMP	NAMP	NAMP	NAMP
4	NAMP	NAMP	NAMP	NAMP
5	AMP	AMP	NAMP	AMP
6	NAMP	NAMP	AMP	NAMP
7	AMP	NAMP	NAMP	NAMP
8	NAMP	NAMP	NAMP	NAMP
9	AMP	NAMP	NAMP	NAMP
10	NAMP	AMP	AMP	NAMP
11	NAMP	NAMP	NAMP	NAMP
12	AMP	NAMP	NAMP	NAMP
13	NAMP	NAMP	NAMP	NAMP
14	AMP	AMP	NAMP	NAMP
15	NAMP	NAMP	NAMP	NAMP
16	NAMP	NAMP	NAMP	NAMP
17	NAMP	AMP	AMP	NAMP
18	NAMP	NAMP	NAMP	NAMP
19	AMP	NAMP	NAMP	NAMP
20	AMP	AMP	NAMP	AMP
21	AMP	NAMP	NAMP	NAMP
22	NAMP	NAMP	NAMP	NAMP
23	NAMP	NAMP	NAMP	NAMP
24	NAMP	NAMP	NAMP	NAMP
25	NAMP	AMP	AMP	NAMP
Total number of AMPs	9	6	4	2