UNIVERSIDADE
BEIRA INTERIOR

Recombinant pre-miRNA-29b: Extracellular production and purification

José Pedro Santos Bento

Dissertação para obtenção do Grau de Mestre em
Biotecnologia
(2º ciclo de estudos)

Orientadora: Profª. Doutora Fani Pereira de Sousa
Coorientadora: Mestre Ana Rita Mugeiro Carapito
Coorientadora: Profª. Doutora Carla Patrícia Alves Freire Madeira da Cruz

outubro de 2024

Declaração de Integridade

Eu, José Pedro Santos Bento, que abaixo assino, estudante com o número de inscrição 12825 de Ciências Biomédicas da Faculdade de Ciências da Saúde, declaro ter desenvolvido o presente trabalho e elaborado o presente texto em total consonância com o **Código de Integridades da Universidade da Beira Interior**.

Mais concretamente afirmo não ter incorrido em qualquer das variedades de Fraude Académica, e que aqui declaro conhecer, que em particular atendi à exigida referenciação de frases, extratos, imagens e outras formas de trabalho intelectual, e assumindo assim na íntegra as responsabilidades da autoria.

Universidade da Beira Interior, Covilhã 09 / 10 /2024

Assinado por: **José Pedro Santos Bento**
Num. de Identificação: 15068123
Data: 2024.10.09 17:37:18+01'00'

Agradecimentos

Gostaria de expressar a minha mais profunda gratidão à minha orientadora, Professora Fani Sousa, por ter me acolhido e aceitado nesta jornada académica. A sua orientação foi fundamental para o desenvolvimento deste trabalho, e sua disponibilidade, paciência e generosidade ao longo do processo foram inestimáveis. O seu conhecimento, rigor científico e dedicação não apenas enriqueceram a qualidade desta investigação, mas também contribuíram de forma crucial para o meu crescimento pessoal e académico. Serei eternamente grato por ter tido a oportunidade de trabalhar sob a sua orientação e se não fosse ela eu hoje não teria uma tese para entregar apesar do meu mau jeito com a escrita ela sempre acreditou em mim mesmo quando eu não acreditava!

Gostava de agradecer também á Professora Carla Cruz profundamente todo o acompanhamento prestado ao longo do percurso académico e durante o desenvolvimento deste projeto, bem como a constante disponibilidade e apoio demonstrados em cada etapa desta jornada.

Quero também agradecer à minha co-orientadora, Rita Carapito, onde expresso a mais sincera gratidão por toda a ajuda e apoio no laboratório. A sua generosidade ao transmitir conhecimento, a paciência infinita para esclarecer dúvidas e a constante disponibilidade foram cruciais para a realização deste trabalho. A sua orientação próxima e dedicada foi essencial em cada fase deste percurso. Sem o seu contributo, seria impossível concluir esta tese, e o seu apoio fez toda a diferença tanto no desenvolvimento científico quanto no meu crescimento pessoal. Muito obrigada por estar sempre presente, com dedicação e entusiasmo.

Aos meus colegas de laboratório, deixo um profundo agradecimento por todo o apoio, companheirismo e ajuda ao longo desta jornada. A vossa disposição para colaborar, compartilhar conhecimento e enfrentar desafios juntos fez com que este percurso fosse muito mais leve e gratificante. Cada conversa, troca de ideias e momento de descontração contribuíram para um ambiente de trabalho acolhedor e inspirador, tornando a caminhada muito mais fácil. Sou imensamente grato por ter tido a oportunidade de aprender e crescer ao vosso lado.

À pessoa mais importante à qual eu não poderia deixar de expressar o meu mais profundo agradecimento. Sem ti, eu nunca teria chegado até aqui. Foste o meu pilar, o meu maior apoio em todos os momentos desta jornada. Ter tido o privilégio de partilhar

o mesmo local contigo e de caminhar lado a lado durante este percurso fez toda a diferença. A tua presença tornou cada dia mais leve e aumentou a minha motivação para continuar, mesmo nos dias mais difíceis. Obrigado por acreditares sempre em mim, por me apoiares incondicionalmente, e por seres uma parte fundamental desta conquista. Sem ti, este mestrado não seria possível. Obrigado, Catarina!

O melhor fica para o fim, a todos aqueles que, no início, eram apenas colegas de trabalho, mas que ao longo do tempo se tornaram muito mais do que isso, deixo um agradecimento especial. Vocês foram fundamentais para tornar esta jornada mais leve e significativa, tanto nos bons momentos como nos mais desafiadores. A vossa amizade, apoio e companheirismo transformaram o local de trabalho em algo muito mais acolhedor e inspirador. Sei que as nossas experiências compartilhadas criaram laços que levarei comigo para a vida, e que posso sempre contar convosco. Obrigado por estarem ao meu lado e por fazerem parte desta etapa tão importante. Obrigado, Bernardo, Fábio, Tiago e Pedro.

Não podia faltar o agradecimento à minha família e amigos, o meu mais profundo agradecimento por todo o apoio e confiança inabalável que sempre depositaram em mim. Vocês acreditaram em mim, mesmo quando eu duvidava, e estiveram presentes nos momentos mais difíceis, oferecendo-me força, encorajamento e carinho. O vosso apoio constante foi fundamental para que eu conseguisse superar cada desafio e continuar em frente. Sou imensamente grato por ter-vos ao meu lado, e por fazerem parte desta conquista que não seria possível sem o vosso amor e amizade.

Resumo Alargado

A terapia baseada em RNA tem avançado rapidamente, trazendo desenvolvimentos promissores para o tratamento de diversas doenças. Os RNAs podem ser divididos em codificantes e não codificantes (ncRNAs), sendo que os não codificantes têm vindo a ganhar grande interesse visto que possuem a função de regulação da expressão génica. Os microRNAs (miRNAs ou miRs) são um exemplo de ncRNAs, que regulam a expressão génica a nível pós-transcricional e controlam processos biológicos com precisão notável, ligando-se de forma complementar aos mRNAs alvo, modulando assim a expressão génica. O pre-miRNA-29b é um desses miRNAs e já é reconhecido por estar envolvido nos mecanismos patológicos associados à doença de Alzheimer. Neste caso, os baixos níveis de miRNA-29b em doentes de Alzheimer, associado à expressão elevada da proteína BACE1, leva à acumulação dos péptidos Aβ, que podem depositar, conduzindo à degeneração dos neurónios, comprometendo a função cognitiva. Logo este pre-miRNA-29b pode ser considerado como potencial biomarcador ou agente terapêutico para esta doença. Mas neste momento a produção em larga escala de RNA para aplicações terapêuticas envolve um processo complexo que exige estratégias de produção e purificação eficientes e controlo de qualidade rigoroso para garantir a pureza e a atividade biológica do RNA. No entanto, os métodos mais comuns para a produção de RNA são a síntese química ou enzimática, que apresentam limitações em termos de quantidade e qualidade do RNA produzido. Nesse contexto, a produção recombinante pode ser uma alternativa para superar alguns desses desafios, uma vez que é um método mais económico e de fácil transposição para larga escala. Foi escolhida a bactéria *Rhodovulum sulfidophilum*, que é uma bactéria púrpura marinha, que apresenta a vantagem de secretar RNAs para o meio extracelular, sem libertar RNases. Com isso, a recuperação do RNA pode ser consideravelmente simplificada em comparação com os protocolos padrão de extração intracelular, facilitando o subsequente processo de purificação, que é a etapa mais crítica e dispendiosa de todo o processo de preparação de RNAs.

Este trabalho teve como foco a otimização da produção extracelular de pre-miRNA-29 assim como a sua subsequente purificação. As condições de crescimento bacteriano foram otimizadas de forma a garantir os níveis máximos da molécula alvo, monitorizando a produção em diferentes tempos de cultura (24 h, 36 h, 48 h e 60 h) e usando diferentes temperaturas (25 ºC, 30 ºC e 37 ºC) para o crescimento bacteriano. De modo geral, os níveis mais elevados de pre-miRNA-29b foram obtidos após 48 horas de crescimento a 25 ºC (0.23 ng/mL de meio), porém, para a continuidade dos estudos,

foi escolhida a segunda melhor condição, de 36 horas de cultura a 30 ºC (0.08 ng/mL de meio), uma vez que o erro associado às medições nesta condição era menor, e consistia num tempo de cultura mais curto, rentabilizando assim a duração do processo.

Quanto à purificação do pre-miRNA-29b, foi selecionado um suporte cromatográfico multimodal, Capto Q Impres, que demonstrou um desempenho promissor na interação seletiva com o RNA alvo. Realizou-se uma otimização inicial para avaliar o impacto da concentração de NaCl nas etapas de ligação e eluição, assim com o pH usado (7 ou 8) Adicionalmente, foi testado o efeito da pré-purificação da amostra por cromatografia de exclusão molecular e também analisada a importância de concentração da amostra antes de prosseguir para o passo de purificação. Foram ainda realizadas comparações quanto à massa de amostra aplicada na coluna, bem como a análise das diferenças entre as amostras intracelulares e extracelulares. Após a avaliação preliminar, foi aplicado um DoE na tentativa de melhorar o desempenho do processo de purificação. Os parâmetros de otimização incluíram a concentração de NaCl durante as etapas de equilíbrio e eluição, bem como o pH do tampão. Desta forma foi possível otimizar o processo de produção extracelular de pre-miRNA-29b em *Rhodovulum sulfidophilum* e também desenvolver um processo promissor para a sua posterior purificação.

Palavras-chave

Produção Recombinante; *Rhodovulum sulfidophilum*; pre-miRNA-29b; Purificação; Desenho Experimental.

Abstract

RNA-based therapies have advanced rapidly, offering promising developments for the treatment of various diseases. Pre-miRNA-29b is already recognized for its involvement in one of the mechanisms underlying Alzheimer's disease. Large-scale production of RNA for therapeutic applications involves a complex process that requires efficient production and purification strategies, alongside rigorous quality control, to ensure the purity and biological activity of the RNA. However, the most used methods face limitations in terms of the quantity and quality of the RNA produced. In this context, recombinant production offers a potential alternative to overcome some of these challenges, as it is a more cost-effective method and easier to scale up. *Rhodovulum sulfidophilum*, a marine purple bacterium, has the advantage of secreting RNAs into the extracellular medium without releasing RNases. This significantly simplifies the RNA recovery compared to standard intracellular extraction protocols, facilitating the subsequent purification process, which is the most critical and costly step of the global biomanufacturing process.

This work focused on optimizing the extracellular production of pre-miRNA-29b and its subsequent purification. Bacterial growth conditions were optimized to ensure maximum levels of the target molecule, monitoring production across different culture times (24 h, 36 h, 48 h, and 60 h) and temperatures (25 ºC, 30 ºC, and 37 ºC). In general, the highest levels of pre-miRNA-29b were achieved after 48 hours of growth at 25 ºC (0.23 ng/mL of medium). However, for the following studies, the second-best condition, 36 hours of culture at 30 ºC (0.08 ng/mL of medium), was chosen as it had a lower variability between assays and offered a shorter culture time, thereby improving process efficiency.

For the purification of pre-miRNA-29b, a multimodal chromatographic support, Capto Q ImpRes, was selected, showing promising performance in selectively interacting with the target RNA. An initial optimization was conducted to evaluate the impact of varying NaCl concentration during the binding and elution steps, as well as the effect of buffer pH (7 or 8). Additionally, the effect of a pre-purification by size-exclusion chromatography column was tested, along with assessing the importance of concentrating the sample before purification. Comparisons were also made regarding the sample mass loaded onto the column and the differences between intracellular and extracellular samples were also addressed. After this preliminary evaluation, a Design of Experiments (DoE) was applied to further refine the performance of the purification

process. The optimization parameters included NaCl concentration during the binding and elution steps, as well as buffer pH. This approach enabled the optimization of the extracellular production of pre-miRNA-29b in *Rhodovulum sulfidophilum*, while also developing a promising strategy for its subsequent purification.

Keywords

Recombinant Production; *Rhodovulum sulfidophilum*; pre-miRNA-29b; Purification; Design of Experiments.

Table of contents

List of Figures

List of Tables

List of Acronyms

A	Adenine
AC	Affinity chromatography
AD	Alzheimer's Disease
APS	Ammonium persulfate
APP	Amyloid precursor protein
Aβ	Amyloid-β
AEXC	Anion-exchange Chromatography
BBD	Box-Behnken design
CCD	Central Composite design
circRNA	Circular RNA
cDNA	Complementary DNA
C	Cytosine
DNA	Deoxyribonucleic acid
DoE	Design of Experiments
DEPC	Diethyl pyrocarbonate
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
BACE1	β-secretase
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid disodium salt 2-hydrate
EMA	European Medicines Agency
FrFD	Fractional Factorial design
FFD	Full Factorial design
gDNA	Genomic DNA
G	Guanine
Tris	Tris(hydroxymethyl)aminomethane
HIC	Hydrophobic interaction chromatography
siRNA	Small Interfering RNA
IEC	Ion-exchange chromatography
lncRNA	Long non-coding RNA
Ago	Mammalian Argonaute
mRNA	Messenger RNA
miRNA	MicroRNA
MMC	Mixed-mode chromatography
nt	Nucleotide
ncRNA	Non-coding RNA
OFAT	One-factor-at-a-time
OD	Optical density
piRNA	Piwi-interacting RNA
PBD	Plackett-Burman design
pDNA	Plasmid DNA
PCR	Polymerase chain reaction
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA

DGCR8	DiGeorge syndrome critical region protein 8
QbD	Quality by Design
qPCR	Quantitative polymerase chain Reaction
RNase	Ribonucleases
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
R. sulfidophilum	*Rhodovulum sulfidophilum*
pol II	RNA polymerase II
RISC	RNA-induced silencing complex
ssDNA	Single-stranded DNA
SEC	Size exclusion chromatography
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
TEMED	Tetramethylethylenediamine
T	Thymine
tRNA	Transfer RNA
TBE	Tris-Borate-EDTA
UV	Ultraviolet
U	Uracil

Chapter 1 - Introduction

1.1. Conventional Drugs vs Biopharmaceuticals

Biopharmaceuticals represent a significant advance in contemporary scientific research. Their application spans virtually all medical specialties, establishing them as one of the most efficient and versatile therapeutic tools available. These drugs have demonstrated remarkable effectiveness in the clinical treatment of a wide spectrum of diseases, thereby revolutionizing the current medical care [1]. Biopharmaceuticals include recombinant protein therapeutics, vaccines, and cellular and gene therapies (Figure 1) [2]. They offer numerous advantages, such as the targeting of specific molecules, thus reducing the side effects that are commonly associated with conventional small-molecule drugs [3]. Additionally, biopharmaceuticals exhibit high specificity and activity compared to conventional drugs, improving the treatment of patients who respond poorly to traditional synthetic medications [4].

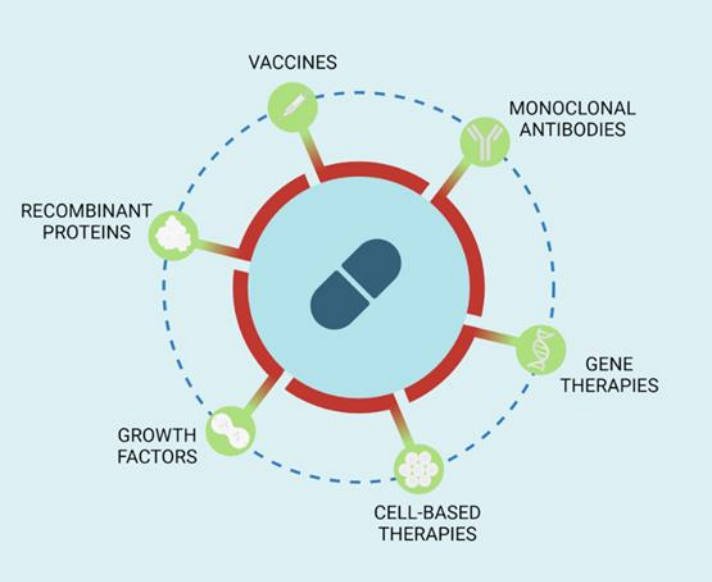

Figure 1 – Various types of biopharmaceuticals.

1.2. Nucleic Acids

Nucleic acids, meaning deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are two biomacromolecules chemically similar, mainly localized in the cell nucleus, although RNA is also abundantly present in the cytoplasm [5]. These molecules play crucial roles in many biological processes, like storage, regulation of gene expression, catalysis of biochemical reactions, and transfer of genetic information within the cell [6]. They can be used as important biomarkers for biological studies and medical diagnostics [7] or as therapeutic agents for the treatment of several diseases.

DNA and RNA primary structure is composed of monomers called nucleotides (nt). The DNA backbone is constituted by two complementary nt chains, while RNA is composed of one single chain. Nt structure is comprised of an organic base linked to a five-carbon sugar, which is, in turn, attached to a phosphate group at a specific carbon position. Notably, DNA sugar is a deoxyribose, whereas RNA has a ribose. The five different bases include adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U). These bases can be categorized into two groups: A and G are purines, characterized by a pair of fused rings, while C and T (found in DNA), and U (found only in RNA) are pyrimidines that feature a single ring structure (Figure 2). DNA and RNA have three bases, A, G, and C in common; however, T is found only in DNA, and U in RNA [7, 8].

In these nt, the 1' carbon atom of the sugar is attached to the nitrogen at position 9 of a purine or position 1 of a pyrimidine. Furthermore, the acidic character of nt is given by the phosphate group, which under normal cellular conditions releases H^+, leaving the phosphate negatively charged [8].

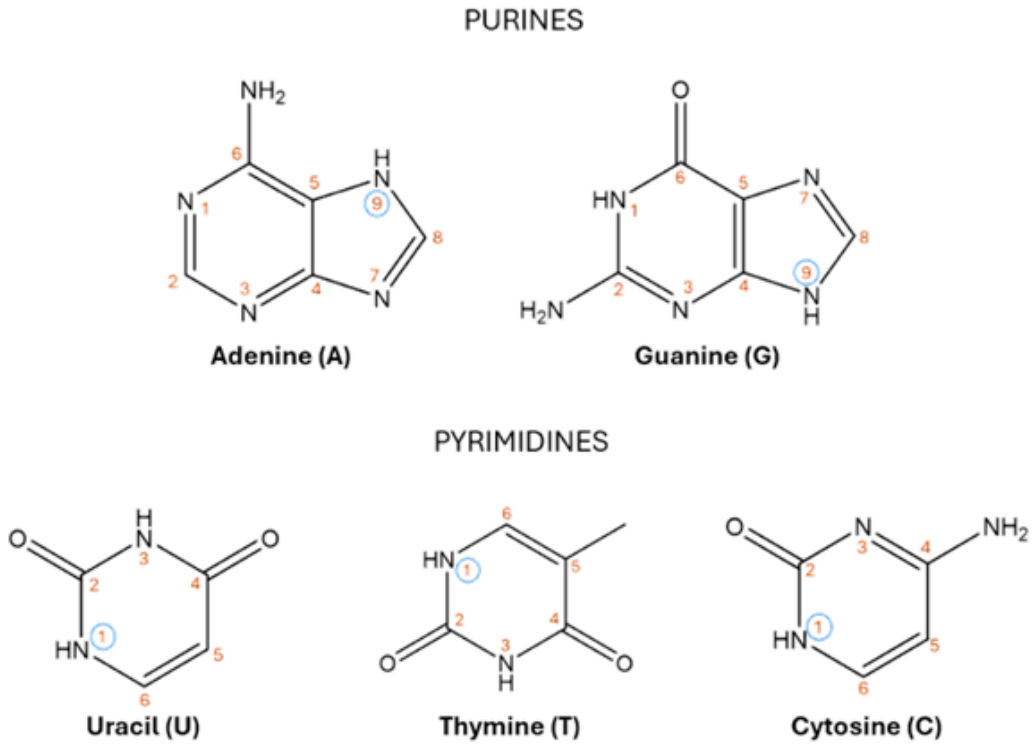

Figure 2 – The five nucleotide bases structure that constitute both DNA and RNA molecules.

1.3. RNA

Historically, RNA was thought to be a mere intermediary in the transfer of genetic information from DNA to proteins, with no other functions in the process of gene expression. However, contemporary research has revealed that various types of RNA are essential to various biological processes. These processes include epigenetic regulation, gene transcription, translation, RNA turnover, chromosomal organization, and genome defence, among others [9, 10]. One main characteristic of RNA is that it is more susceptible to alkaline hydrolysis, given the 2'-OH group on the pentose ring, contrarily to DNA that has a hydrogen atom. As a result, RNA is relatively unstable *in vivo* due to the presence of ribonucleases (RNases) that are ubiquitous in both serum and cells [11]. While RNA typically adopts a single-stranded configuration, short sequences that are complementary within the strand can base-pair and create secondary structures. These structures, frequently referred to as hairpins, hairpin-loops, or stem-loops, arise from specific nt sequences. Given that the structure of RNA molecules is closely linked to their function, these secondary structures enable RNA to carry out a wide range of biological functions [12]. As previously mentioned, RNA can perform various functions and can be divided into coding and non-coding RNAs (ncRNA) [13], which will be discussed in more detail.

1.3.1. Coding and non-coding RNAs

Coding RNA, specifically messenger RNA (mRNA), is the type of RNA responsible for encoding proteins. mRNA carries the genetic information for protein synthesis, meaning that the nt sequence of mRNA results in the sequence of amino acids of the given polypeptide chain during translation [14]. Due to this role, it can also correct gene expression defects or abnormalities through the exogenous introduction of a given mRNA molecule. For example, regarding vaccination, mRNA encoding for specific antigens can induce the production of antibodies and consequently activate an immunological response. This highlights the significant potential of mRNA as a biopharmaceutical for the treatment of various diseases. It allows the development of personalized medicine by inducing patients to synthesize their own therapeutic proteins [15]. As a matter of fact, mRNA vaccines revolutionized the gene therapy field by allowing a rapid and effective response to the COVID-19 pandemic, as they have shown to be very successful in the prevention of infection and control of symptoms [16, 17].

ncRNA refers to RNA that does not encode a protein but it is rather involved in the regulation of gene expression through many different mechanisms [18]. ncRNAs can be divided into housekeeping ncRNAs and regulatory ncRNAs, as it is represented in

Figure 3. Housekeeping ncRNAs are expressed constitutively and ubiquitously within the cell, playing crucial roles in a diverse range of cellular activities. These include transfer RNA (tRNA), ribosomal RNA (rRNA), small nucleolar RNA (snoRNA) and small nuclear RNA (snRNA). The tRNA acts as an adaptor molecule in the protein synthesis process, facilitating the transfer of amino acids and establishing a physical linkage between mRNA and the amino acid sequence of the protein. rRNA, on the other hand, constitutes a significant portion of cellular RNA and plays a central role in ribosome structure. It is approximately 80%-90% of the total mass of cellular RNA, making it the primary product of transcription [19]. Regulatory ncRNAs, in contrast, are only expressed in specific cell types and respond to developmental signals, internal conditions, and environmental stimuli. These comprise several types, including short-interfering RNA (siRNA), long non-coding RNA (lncRNA), piwi-interacting RNA (piRNA), circular RNA (circRNA), and microRNA (miRNA) [18].

siRNAs, typically between 21-25 nt in length, are a class of double-stranded RNAs (dsRNAs) specifically designed to silence the expression of a target gene [20]. LncRNAs, which are longer than 200 nt, play essential roles in the regulation of gene expression. Even though they bind to mRNA, they have little to no protein-coding capacity. They can be found in both nuclear and cytosolic fractions and share structural features with mRNA [21]. PiRNAs, ranging from 24-32 nt, are short single-stranded ncRNAs produced through a DICER-independent mechanism. These RNAs interact with piwi proteins to regulate gene expression by silencing transposons, modulating gene transcription, and influencing mRNA turnover and translation [22]. CircRNAs, comprise a wide range that can go from 100 nt to over 10 kb. They are a recently discovered type of endogenous ncRNA characterized by covalently closed loop structures without 5’ or 3’ polarities and non-polyadenylated tails, distinguishing them from linear RNAs. This unique structure confers higher resistance to exonuclease RNase R. While circRNAs in humans are generally classified as ncRNAs, some have been experimentally verified to encode specific proteins [23].

Regarding miRNAs, which are the most intensely studied type of ncRNAs, generally comprise about 19-25 nt, and are originated from hairpin-shaped precursor molecules. These single-stranded RNAs are encoded endogenously within the genome and exhibit a unique tissue-specific expression pattern. This specificity allows miRNAs to regulate a wide range of biological processes with outstanding precision. Through complementary binding to the target mRNAs, miRNAs have their regulatory effects in two ways: when the binding is perfect, with total complementarity the target mRNA degrades, consequently inhibiting the corresponding protein synthesis. On the other

hand, when the binding is imperfect, the translation of the target mRNA is suppressed, thus reducing the production of the protein with no degradation of the mRNA itself. This dual mechanism of action enables miRNAs to specifically regulate gene expression and maintain cellular homeostasis, making them crucial players in various physiological mechanisms [24].

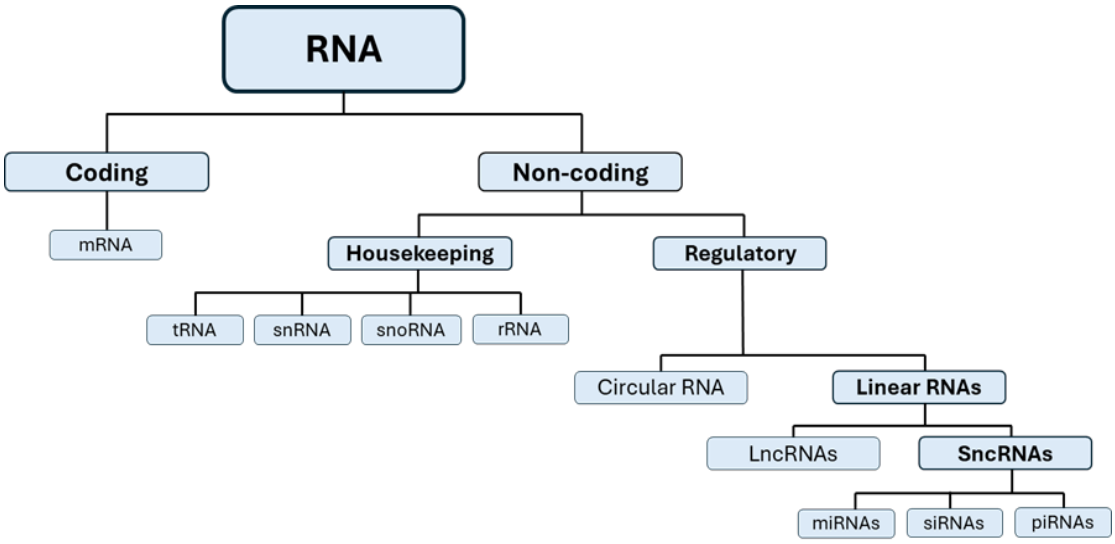

Figure 3 – Scheme of the classification of coding and non-coding RNAs (Adapted from [19, 25]).

The biogenesis of miRNAs is a complex, multi-step process initiated by transcription via RNA polymerase II (pol II) (Figure 4). Under the catalysis of pol II, primary miRNA (pri-miRNA) is transcribed. This then binds to DiGeorge syndrome critical region protein 8 (DGCR8) and undergoes nuclear processing by Drosha to form the precursor miRNA (pre-miRNA) with a hairpin structure. The pre-miRNA is subsequently exported to the cytoplasm by Exportin-5, where it is further processed by Dicer, resulting in the formation of a miRNA duplex. MiRNAs then assemble with the mammalian Argonaute (Ago) family of proteins to form an effector complex known as the RNA-induced silencing complex (RISC). This complex targets specific mRNAs. It is known that miRNAs can regulate a vast number of target mRNAs, and, on the other side, a single mRNA can have different miRNAs target sites. This intricate regulatory network underscores the critical role of miRNAs in gene expression control and cellular function [19].

Ultimately miRNAs can serve as biomarkers and are directly implicated in the pathogenesis and progression of various diseases, including cancer and Alzheimer's disease (AD). The recognition of the significance of functional ncRNAs in both physiological and pathological conditions has been transformative in the biopharmaceuticals research field. Notably, the discovery of miRNA and siRNA-

mediated post-transcriptional gene regulation and RNA interference mechanisms has revolutionized life sciences and biomedical research, highlighting their potential for diagnostic and therapeutic applications [26].

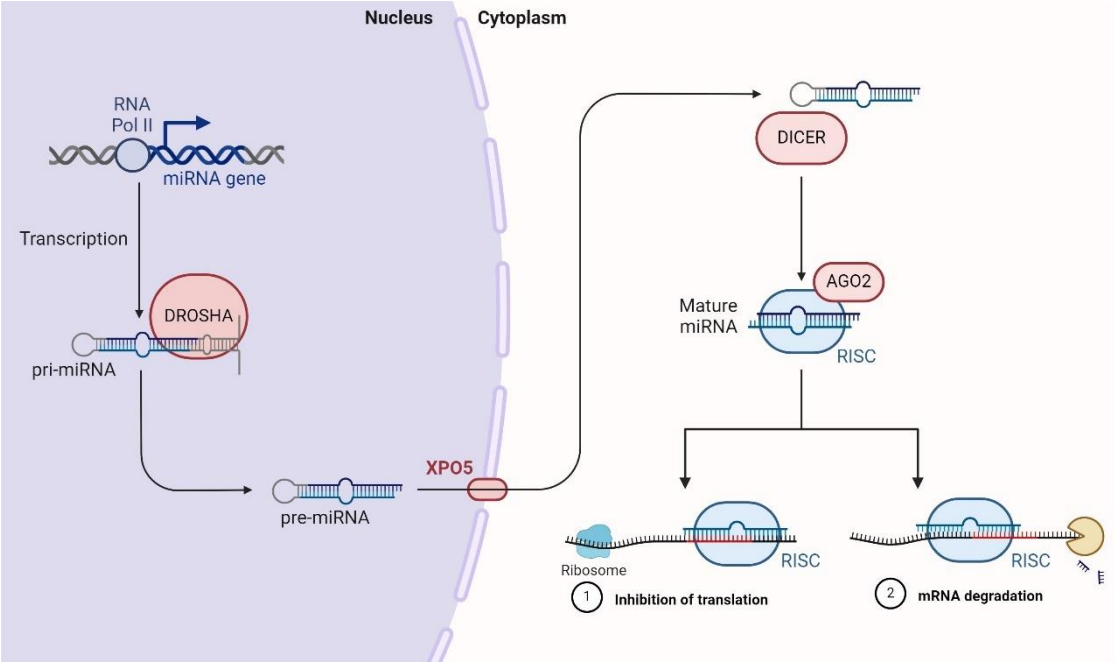

Figure 4 – Biogenesis and mechanism of action of miRNA.

1.4. Nucleic Acids based therapy

Gene therapy is considered a transformative advancement in medicine due to its unique approach of targeting the causes of several diseases, rather than only relieve their symptoms. This type of therapy involves the manipulation of genetic material within a patient's cells to alter defective genes that are responsible for development of the disease. The European Medicines Agency (EMA) defines a gene therapy medicinal product as any product that includes an active substance composed of recombinant nucleic acids. These are administered to humans with the intent to regulate, replace, repair, edit, add, or delete a genetic sequence [27]. This modulation is achieved through the administration of exogenous nucleic acids, including DNA, mRNA, miRNA, small siRNA, and antisense oligonucleotides. These molecules play distinct yet complementary roles in gene therapy. DNA and mRNA act as templates for protein production, which is fundamental for the proper cellular function, while miRNA and siRNA are involved in gene expression regulation by inhibiting the translation of specific mRNAs, based on a silencing mechanism. Antisense oligonucleotides, on the other hand, bind to complementary RNA sequences and can prevent their translation or promote their degradation [28].

Current gene therapy strategies encompass both *ex vivo* and *in vivo* approaches, each offering distinct methodologies for genetic modification. In the *ex vivo* approach, patient somatic cells are extracted and therapeutically modified outside the body within a controlled laboratory environment. These genetically altered cells are then reintroduced into the patient, where they can function with the desired modifications [29]. On the other hand, *in vivo* gene therapy involves the direct delivery of genetic material to the patient, typically through the injection of specifically engineered vectors into the patient's bloodstream [28]. Gene-based therapeutics hold significant promise for the treatment of a wide range of human diseases, among of which viral infections, cancer, and genetic disorders. However, realizing the full clinical potential of gene-based therapy requires overcoming some key challenges. These include the development of production methods that are feasible and cost-effective, thus achieving a stable, effective, and disease-specific biopharmaceutical, along with the creation of efficient and low-toxicity *in vivo* delivery systems. Addressing these bottlenecks will be crucial for the successful transfer of gene-based therapeutics from fundamental research to clinical practice [30].

1.4.1. MiRNA-29 in Alzheimer's disease

AD is recognized as the primary cause of dementia and stands among the most fatal and challenging diseases of the 21st century [31]. It is a neurodegenerative condition marked by a gradual, yet relentless decline in cognitive function, initially manifesting as memory loss and the inability to form new memories. As the disease progresses, it leads to significant impairments in the capacity to perform daily activities [32, 33]. Numerous studies suggest that the accumulation of amyloid-β (Aβ) peptides in neurons is a key factor in the development and progression of AD, as it contributes to synaptic dysfunction and neuronal death. Since similar patterns are observed in the brains of healthy older adults, it is widely believed that the biological processes underlying AD begin well before clinical symptoms become apparent [34, 35].

Aβ is a peptide consisting of 40 or 42 amino acids, generated through the sequential cleavage of amyloid precursor protein (APP) by β-site amyloid precursor protein cleaving β-secretase (BACE1). This enzyme is found to be overexpressed in individuals with AD, making it a primary target for therapies aimed at inhibiting Aβ production. However, treatments designed to reduce the production or eliminate existing Aβ proteins have not yet been proven effective. Currently available therapies can only slow the progression of the disease, rather than fully stopping it [33, 36].

Recent studies have revealed that patients with AD exhibit low levels of miRNA-29b alongside elevated expression of the BACE1 protein. This finding supports the idea of a

connection between miRNA-29b and AD, as miRNA-29b acts on the inhibition of BACE1 expression, thereby reducing Aβ accumulation in neuronal cells. Consequently, miRNA-29b is proposed as a potential biopharmaceutical candidate for the treatment of AD [36, 37].

1.5. RNA Production

RNA is becoming increasingly relevant as more functions, including regulatory and enzymatic roles, are being discovered. To investigate the diversity of RNA functions and conduct biochemical, biophysical, and genetic studies, the sample preparation process is crucial. Although, currently, several techniques are suitable for molecular biology procedures, the preparation of RNAs for therapeutic purposes presents additional challenges that must be addressed since they need to meet specific product requirements established by the regulatory agencies. Most miRNAs used in the development of novel therapeutic strategies are produced through chemical or enzymatic synthesis, or recombinant production [38, 39].

Chemical synthesis is the method of choice to produce oligonucleotides shorter than 80 nts, with an optimal length of fewer than 10 nts. The synthesis reaction starts from the 3' end to the 5' end and involves four different steps: deblocking, coupling, capping, and oxidation. These steps use 2'-OH and 5'-OH protecting groups of phosphoramidite monomers. The four-step cycle is repeated until the desired oligonucleotide length is achieved. This method is rapid and allows different chemical modifications; however, it implicates the usage of expensive equipment, the chain length is limited since it introduces changes difficult to control, and involves costly modified phosphoramidites [40].

Enzymatic synthesis, or *in vitro* transcription is the most common method and enables template-directed synthesis of RNA molecules with any sequence, ranging from short oligonucleotides up to several kilobases long. This process uses the promoter components of bacteriophage systems, with the T7 system being the most employed. The T7 system requires only Mg^{2+} as a co-factor for RNA synthesis. While this method can produce RNA molecules on a microgram scale, the scaling up is quite challenging due to the high quantities of enzyme that would be required, making the process incredibly expensive. Additionally, this method often results in low yields and still needs multiple and sequential enzymatic and purification steps to remove ubiquitous impurities [41, 42].

These two synthesis methods are the most used and established techniques for efficient RNA production. Nonetheless, several limitations must be overcome, such as

challenges associated with large-scale RNA production and complex purification procedures. These issues can result in non-targeted gene silencing, which not only reduces therapeutic effectiveness but also might provoke an immunological response, thus restricting the usage of these RNAs in preclinical or clinical trials [43]. To address this and consider the need for rapid production of promising biopharmaceuticals and achieving sustainable large-scale production processes, recombinant production emerges as the most promising method for this purpose [44].

Recombinant RNA techniques using *in vivo* agents are effective at preserving the structure, function, and properties of natural RNAs. This method is based on the introduction of a recombinant plasmid encoding for the target RNA into a host cell, which is then cultured under appropriate conditions. The host cell's transcription machinery then synthesizes the RNA of interest, which is accumulated in the cytosol. After, the cells are lysed and the RNA is purified using standard chromatographic techniques [44, 45]. Recombinant RNA expression has primarily been achieved and optimized using *Escherichia coli* (*E. coli*) as the host due to its ease of growth, cost-effectiveness, and the availability of numerous plasmids and strains. However, exploring broader possibilities for recombinant RNA production, namely research in using different expression hosts that could be more advantageous than *E. coli,* is already under consideration [39, 44]. A schematic representation of enzymatic synthesis and recombinant production is shown in Figure 5.

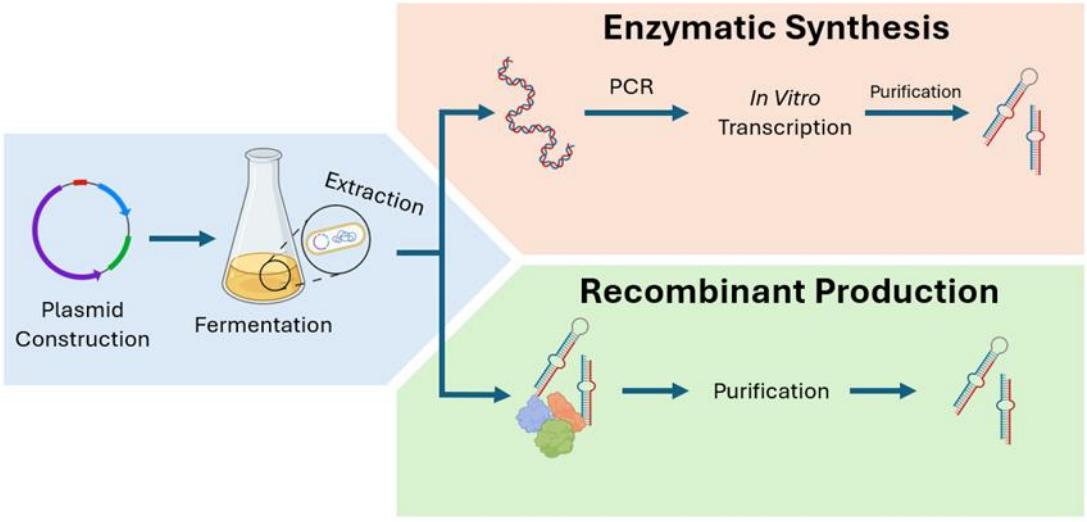

Figure 5 – Schematic representation of enzymatic synthesis and recombinant production of RNA.

1.5.1. RNA biosynthesis in *Rhodovulum sulfidophilum*

The production of recombinant products in *E. coli* is both rapid and efficient, facilitated by the availability of numerous plasmids, as already mentioned above. However, the process requires cell lysis and extraction, which are time-consuming given their complexity. Also, the potential release of endotoxins can contaminate the target biomolecule preparation and lead to *in vivo* toxicity. Besides this, *E. coli* expresses several endonucleases, which implicate additional challenges in maintaining the integrity of the target RNA [39]. For these reasons, *Rhodovulum sulfidophilum* (*R. sulfidophilum*) can be considered as an alternative host, as it presents different characteristics that can overcome some of these issues associated with the usage of *E. coli*. The marine bacterium *R. sulfidophilum* is a gram-negative organism that grows under anaerobic conditions with light [46], or aerobic conditions in the dark [47]. This non-pathogenic bacterium offers several advantages over *E. coli*, among which is the ability to secrete nucleic acids directly into the culture medium and, at the same time, not secreting RNases into the extracellular medium [47, 48]. The first researchers to describe the possible use of *R. sulfidophilum* as a host for the production of an RNA aptamer were the research group of Kikuchi, and later our research group was able to produce human pre-miRNA-29b in the same host [46, 47].

The developed method enabled high-yield intra- and extracellular production of human pre-miRNA-29b, presenting a promising alternative to conventional methods. Furthermore, this research demonstrates that recombinant technology can be used to produce this biomolecule, paving the way for further studies on the use of natural pre-miRNA molecules in the biomedical field [39, 47].

1.6. RNA Purification Strategies

RNA molecules are typically extracted from complex mixtures, always needing to undergo a robust purification process to ensure the target molecule's integrity, stability, and high purity. Chromatography techniques are the method of choice for purifying biomolecules, as they effectively isolate various high-value products [49, 50].

Chromatography is a robust, efficient, and versatile technique that can separate a mixture into its individual components in a single step while providing quantitative characterization of each component. This separation is achieved through the implication of two main phases: the stationary phase and the mobile phase. The stationary phase consists of a solid support with ligands, or a layer of liquid adsorbed on its surface, whereas the mobile phase is either a liquid or a gas [50, 51]. Compounds of the mobile phase that interact strongly with the stationary phase are retained longer in the system, while those with less interaction are selectively distributed in the mobile

phase. Consequently, different compounds elute from the system in increasing order of their affinity to the column, meaning that those with the weakest interaction elute first, and those with the strongest interaction elute last [52]. There are various types of liquid chromatography, each one distinguished essentially by the properties of the stationary phase used. For example, size exclusion chromatography (SEC) separates biomolecules according to their size and shape, ion-exchange chromatography (IEC) relies on the overall net charge of the molecule, while hydrophobic interaction chromatography (HIC) is based on the hydrophobic properties of the molecule, and affinity chromatography (AC) depends on specific interactions [49, 53]. These will be discussed in more detail in the next topics.

1.6.1. Size Exclusion Chromatography (SEC)

The fundamental principle of SEC, or gel filtration, is to separate macromolecules based on their molecular sizes and shape using porous materials (Figure 6). The column matrix consists of inert materials with pores of defined sizes. When a sample solution that contains different molecules of varying dimensions is passed through the column at a constant flow rate, molecules larger than the pore size are unable to enter the gel particles and will pass through the column more quickly. Contrarily, molecules smaller than the pores diffuse into the pores and are retained longer, resulting in proportionally longer retention times [54, 55].

SEC is commonly employed as a separation process when the sample has been previously clarified and the goal is to determine the presence of high-order impurities such as aggregates or degradation products. In nucleic acid purification, SEC can be used to separate genomic DNA (gDNA) or plasmid DNA (pDNA), which usually elute first, given their high molecular weight, from other smaller molecules such as RNA, proteins, and endotoxins, that elute lastly. In the context of RNA purification, different researchers have been working with this method, demonstrating fast performance and high-resolution purification of RNA [56, 57]. Nonetheless, SEC has inherent limitations, including low resolution and the need of higher sample dilutions. Consequently, SEC is typically employed as a final step in the downstream processing [55, 58].

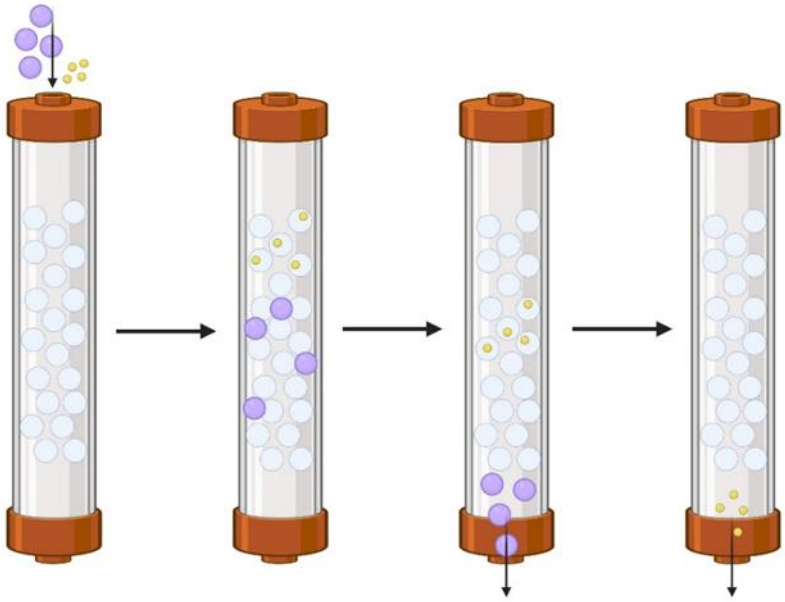

Figure 6 – Scheme of SEC separation showing the separation of smaller (yellow) and larger (purple) molecules.

1.6.2. Ion-exchange chromatography (IEC)

IEC is a technique that relies essentially on electrostatic interactions between charged groups of the mobile phase compounds and a solid chromatographic material with an opposite charge. These interactions are often easily reversible and can be altered by adjusting the pH or ionic strength of the mobile phase. Matrices with a positive charge, known as anion-exchange matrices, will be used to absorb negatively charged compounds. Conversely, matrices with a negative charge, referred to as cation-exchange matrices, will be used to adsorb positively charged molecules [49, 59].

Anion-exchange chromatography (AEXC) is an extremely popular technique and can be used for nucleic acids purification. Since nucleic acids are polyanionic molecules due to the phosphate groups on their backbone, they are effectively captured on supports functionalized with positively charged groups. AEXC provides several advantages, including rapid separation and analysis, it does not require the usage of organic solvents, it can be regenerated with sodium hydroxide, and there are available numerous commercial stationary phases.

A research group developed a method for purifying siRNA using AEXC. The authors also suggest that this approach could also be applied to the purification of other nucleic acids, such as miRNA [60]. Another group produced milligram-scale quantities of pure RNA from crude transcription reactions in under four hours, using weak anion-exchange [61]. Nonetheless, there are some drawbacks; for example, some RNA species

may co-elute with pDNA, and the technique can exhibit relatively low resolution between pDNA and endotoxins [62, 63].

1.6.3. Hydrophobic interaction chromatography (HIC)

HIC is a crucial technique in bioseparation, applicable across laboratory and industrial scales. HIC explores the hydrophobicity of biomolecules to achieve separation through interactions between the immobilized hydrophobic ligands and the non-polar regions of the biomolecules. In an aqueous environment, high salt concentrations enhance the retention of polar biomolecules by displacing water molecules around both the biomolecules and the ligands, thereby increasing entropy and promoting the attraction of non-polar groups to the stationary phase [49, 64]. Then, elution is achieved by reducing the salt concentration in the mobile phase, which weakens the established hydrophobic interactions. For nucleic acids, HIC exploits the greater hydrophobicity of single-stranded molecules, which have the hydrophobic aromatic bases more exposed compared to double-stranded ones [53, 64]. Despite its advantages, the requirement for high salt concentrations during the binding step represents a significant impact on the environment, consequently limiting its application in an industrial scale [58]. In the context of nucleic acid purification, HIC is primarily used for the preparative purification of pDNA, since RNA is typically regarded as an impurity in this context [65]. This process can involve negative chromatography, in which the objective is to bind impurities such as RNA, gDNA, oligonucleotides, and denatured pDNA, rather than the target biomolecule itself. The target it then eluted with the decrease of salt concentration [49, 58]. Nonetheless, considering the more hydrophobic character of RNA, compared to pDNA, HIC can also be exploited in RNA purification.

1.6.4. Affinity chromatography (AC)

AC is recognized as one of the most powerful and versatile chromatographic techniques, particularly effective even when the target biomolecule is present as a minor component in a complex mixture. The separation of the target biomolecule is based on specific interactions, such as those found between antigens and antibodies, enzymes and substrates, or nucleic acids and proteins (Figure 7). These naturally occurring interactions are explored in AC by immobilizing one compound of the interaction pair on the stationary phase, while the other compound is present in the mobile phase and can be selectively adsorbed onto the solid phase [66].

Since its application in various fields, AC has evolved to incorporate a diverse range of ligands for both analytical and preparative applications. However, the biological origin of some ligands can be a limitation, as these ligands are often fragile and exhibit low binding capacities, besides greatly increasing the process cost. To address this issue,

the development of synthetic ligands that combine the selectivity of natural ligands with the high capacity and robustness of synthetic systems has emerged as a promising approach for enhancing AC [58, 62]. The specific interactions between ligands and target molecules can involve different types of interactions, such as electrostatic and hydrophobic interactions, van der Waals forces, π-π stacking, and hydrogen bonding. Due to this variety of potential interactions, the elution step can be tailored to be either specific using a competitive ligand, or non-specific, by adjusting the buffer composition regarding the pH, ionic strength, or polarity [67]. The choice of elution strategy depends on the properties of the stationary phase and the chemical characteristics of the biomolecules involved [49, 58].

AC has been effectively applied in RNA purification processes by leveraging the natural biorecognition between RNA and amino acids. Histidine and arginine have been used as affinity ligands, demonstrating their ability to isolate different RNA species. This versatility is valuable in molecular biology analysis and RNA therapeutics preparation, showcasing the potential of these methods to overcome RNA purification challenges [68]. One of the most common AC strategies for RNA purification is the use of oligo(dT) for mRNA purification. This approach relies on the base pairing specificity between A and T nucleotide bases. AC allows the purification of mRNA directly from biological samples or from previously isolated total RNA. Overall, the method is simple, reliable, and frequently preferred over other techniques due to its efficiency and ease of use [66, 68, 69]. Some examples of this technique used in RNA purification are based on arginine, Oligo(dT) and histidine ligands, which enable the separation of RNA from other contaminants [70-72].

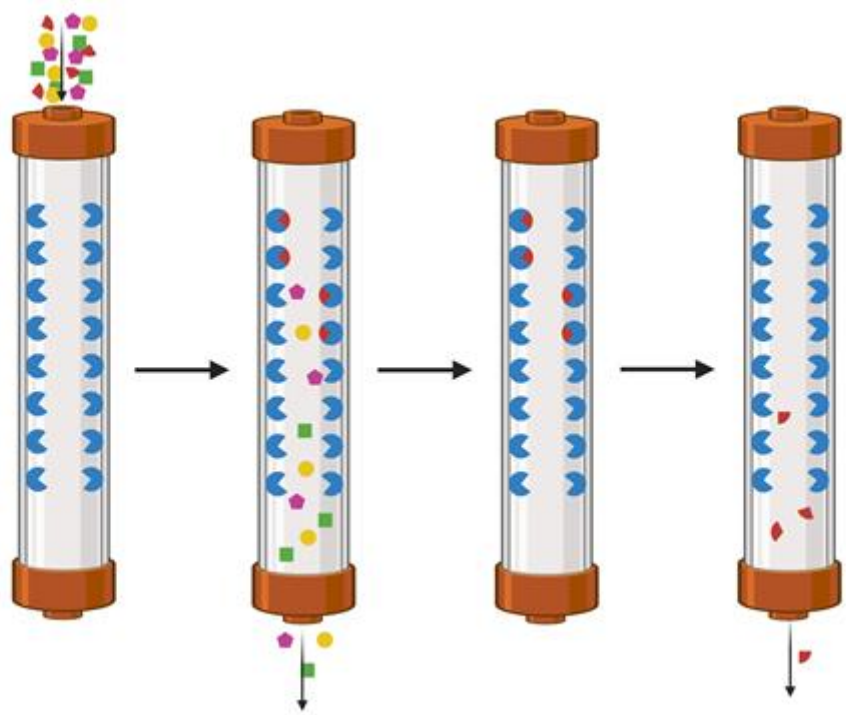

Figure 7 – Scheme of AC showing the separation of different molecules (target biomolecule - red triangle with affinity to the specific ligand - blue circle).

Another emerging type of purification is through multimodal or mixed-mode chromatography (MMC), which enables multiple binding interactions simultaneously. Presents a more refined alternative to traditional chromatography methods involving single-interaction steps. For example, MMC can simultaneously leverage ion exchange and hydrophobic interactions, often enhancing both selectivity and specificity in the separation process. This approach has been employed in various applications, such as the separation of proteins [73], oligosaccharides [74], antibody purification [75], and pDNA purification [76]. Additionally, MMC media can be tailored to function as "single mode" chromatography resins by adjusting conditions to minimize or modify certain interactions. This adaptability underscores the versatility and potential utility of MMC in a wide range of biomolecular separation processes [77].

1.7. Design of Experiments (DoE)

Experimental design is a fundamental aspect of scientific research and industrial processes, crucial for ensuring the validity, reliability, and applicability of collected data. Effective experimental design allows researchers to systematically investigate the effects of multiple factors on a given response variable, thereby facilitating a comprehensive understanding of complex systems. This methodology not only enhances the accuracy of the research but also optimizes its efficiency by ensuring an optimal resource utilization [78, 79].

Quality by Design (QbD) is connected to experimental design since it is a systematic approach that starts with predefined objectives and focuses on understanding and controlling both the product and process, grounded in scientific principles and quality risk management. Emphasized by all pharmaceutical regulatory agencies, quality is essential for ensuring customer satisfaction in terms of product performance, reliability, and absence of defects, such as impurities. The principles of QbD have been widely adopted across industries to enhance product and process quality. In response to the demand for safe and effective drugs, pharmaceutical companies are heavily investing in drug discovery and process development to design consistently high-quality products [79, 80].

One-factor-at-a-time (OFAT) experimentation is the traditional approach in which one factor is varied while all other factors are maintained constant. The advantages of OFAT include its simplicity, ease of understanding, and implementation, which makes it suitable for researchers without extensive statistical training allowing a clear examination of the effect of each specific factor independently [78, 81].

However, while OFAT is straightforward, it has notable limitations, such as failure to consider interactions between factors, which can lead to incomplete or misleading conclusions. In contrast, Design of Experiments (DoE) methodologies can address multiple factors simultaneously and reveal interactions, providing a more holistic view of the system under study (Figure 8) [78, 82].

DoE is a robust statistical methodology that simplifies the planning, execution, analysis, and interpretation of controlled tests to evaluate the factors that influence a particular outcome. DoE provides a structured approach to experimentation, which is instrumental in identifying cause-and-effect relationships between the different tested variables. Employing DoE helps in optimizing processes by identifying the optimal conditions for achieving a given maximum outcome, usually efficiency and effectiveness. It is useful for understanding and minimizing sources of variability, leading to more consistent and reliable outcomes. Additionally, DoE reduces the number of experiments required, thereby saving time, materials, and, consequently, financial resources. By thoroughly understanding the contributing factors, DoE enhances the quality and reliability of products or processes [78, 83].

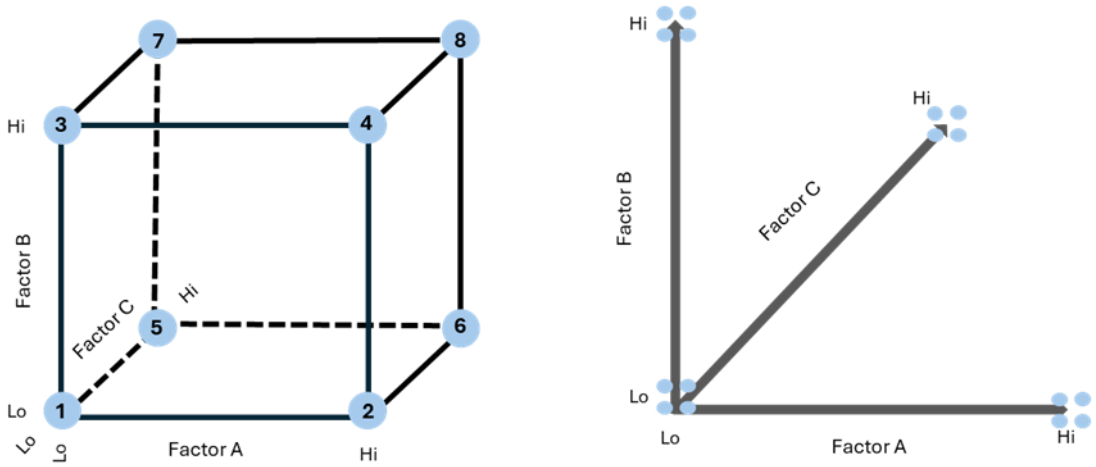

Figure 8 – Schematic of the difference between DoE (left) and OFAT (right).

The quality of an experimental design significantly impacts the validity of the research conclusions. High-quality design ensures that the results are statistically significant, reproducible, and generalizable. Essential attributes of quality design include randomization, which ensures that experimental units are randomly assigned to treatments, thereby minimizing bias. Reproducibility, which involves repeating the experiment multiple times, helps estimate variability and improves the precision of the results. Blocking controls for the effect of nuisance variables, thereby enhancing the accuracy and validity of the conclusions [83, 84].

The designs can be classified into two primary categories based on the experimental aim: screening designs and optimization designs (response surface designs). The most used screening designs are the Plackett-Burman design (PBD) and the Fractional Factorial design (FrFD). In contrast, the Full Factorial design (FFD), Box-Behnken design (BBD), and Central Composite design (CCD) are examples of widely used optimization designs [83].

1.7.1. Plackett-Burman Design

PBD is one of the most frequently employed screening designs due to its versatility and efficiency. It is used primarily in two distinct ways. Firstly, as a screening design, in which the most critical factors can be identified from a large pool of potential variables. Secondly, PBD can be used in ruggedness or robustness studies during method validation, ensuring that methods remain reliable under varying conditions [83, 85]. The principal advantage of PBD lies in its ability to evaluate a large number of factors with a minimal number of experiments, making it a cost-effective and time-efficient option. However, its utility is limited to estimating the main effects of the factors, since it does not account for interactions among them. Despite this limitation, PBD remains

a powerful tool for preliminary screening in complex experimental setups, laying the groundwork for more detailed optimization studies [85].

1.7.2. Factorial Designs

This category includes FFD and FrFD. In FFD, every combination of all levels of all factors is tested at least once. This comprehensive approach has the significant advantage of being able to estimate all the main effects and the interaction among factors. However, the total number of experiments increases exponentially with the number of factors, making FFD more suitable for studies involving a small number of factors [83, 86]. On the other hand, FrFD requires only a fraction of all possible combinations, making it more practical when dealing with a larger number of factors. This design is particularly useful for initial screenings, where the objective is to identify the most influential factors without the need for exhaustive testing. The insights gained from FrFD can be then used to inform subsequent optimization designs, such as CCD or BBD. These follow-up designs ensure that the factors being screened have a significant impact on the responses, thereby refining the experimental focus and enhancing the efficiency of the overall study [83, 87].

1.7.3. Central Composite Design

The CCD is a widely used and effective optimization design that includes three types of design points: FFD or FrFD points, central points, and star points. This design leverages the preliminary implementation of two-level FFD or FrFD to assess factors, providing a comprehensive evaluation of the experimental space. However, CCD requires additional central points and examines extreme conditions, which leads to an increased number of experiments, and, potentially, unsatisfactory results. The inclusion of central points helps to provide a more precise estimation of the response surface curvature, but the need to explore extreme conditions can sometimes result in inefficiencies or data that are less useful. Despite these drawbacks, CCD still remains a robust choice for optimization, particularly when a detailed exploration of the experimental space is needed [86-88].

1.7.4. Box-Behnken Design

The BBD is another type of optimization design. Unlike the CCD, BBD does not investigate extreme factor combinations. This characteristic gives BBD the unique advantage of requiring fewer experimental runs than CCD for the same number of factors. However, BBD also has a notable drawback because the design is not sequential. This means that the experimental results from one batch are almost useless for the next, having to repeat each batch of experiments [83, 89]. The application of BBD is particularly suitable for systems involving more than two factors where the

optimal values are expected to lie within the mid-range of the factor levels. This makes BBD a valuable tool in optimization studies where experimental efficiency and the avoidance of extreme conditions are priorities [90].

Overall, considering the challenges associated with the production and purification of RNA, exploiting new biotechnological strategies and using adequate tools can effectively help on the progress of this research field and on the successful implementation of RNA in biomedical applications.

Chapter 2 - Objectives

Given the therapeutic potential of miRNAs, there is a critical need to develop efficient, selective, and robust production and purification processes. Consequently, it is essential to design novel strategies for producing highly pure and biologically active miRNAs. One promising approach involves the recombinant production of these biomolecules in prokaryotic hosts, such as *R. sulfidophilum*.

The expression system *R. sulfidophilum* facilitates the production of human pre-miR-29b, enabling a straightforward recovery process for low molecular weight RNA from the culture media. To achieve high yield and purity, the production and purification processes must be optimized.

The primary objective of this work is dual-sided with the optimization of the production of pre-miRNA-29b and the design of an experimental approach for the chromatographic purification of pre-miRNA-29b. To achieve these goals, the following tasks have been undertaken:

- Evaluation of the best conditions (temperature of growth and time-point for recovery) for the production of pre-miRNA-29b in *R. sulfidophilum*.
- Evaluation and understanding of binding and elution profile of pre-miRNA-29b at different pH and ionic strength conditions, using the commercial chromatographic support Capto Q ImpRes.
- Development of a FFD, establishing the key factors under study and their range of values, according to the initial screening, to improve the purification performance.

Chapter 3 - Materials and Methods

3.1. Materials

For solid culture media preparation, the following reagents were used: yeast extract, D-glucose anhydrous (C6H12O6), and sodium chloride (NaCl) from Thermo Fisher Scientific Inc. (Waltham, USA), polipeptone, and magnesium chloride ($MgCl_2$) from Acros (Waltham, USA), LB-Agar from Oxoid (Waltham, USA), zinc sulfate heptahydrate ($ZnSO_4.7H_2O$) and manganese chloride tetrahydrate ($MnCl_2.4H_2O$), both from Sigma Aldrich (St. Louis, USA), and iron sulfate heptahydrate ($FeSO_4.7H_2O$) from Thermo Fisher Scientific Inc. (Waltham, USA).

For liquid culture media preparation the following reagents were used: tryptone, polipeptone, yeast extract, NaCl, potassium dihydrogen phosphate (KH_2PO_4), and D-glucose anhydrous ($C_6H_{12}O_6$) from Thermo Fisher Scientific Inc. (Waltham, USA), dipotassium hydrogen phosphate (K_2HPO_4) from Panreac AppliChem (Barcelona, Spain), magnesium sulfate heptahydrate ($MgSO_4.7H_2O$) from Labkem (Barcelona, Spain), calcium chloride dihydrate ($CaCl_2.2H_2O$), ammonium sulfate ($(NH_4)_2SO_4$) from Panreac (Barcelona, Spain), and kanamycin from Thermo Fischer Scientific Inc. (Waltham, USA).

The extraction of low molecular weight RNA used N-lauroylsarcosine sodium salt (Sarcosil), guanidine thiocyanate, sodium citrate, and isoamyl alcohol, all from Sigma-Aldrich (St. Louis, Missouri, USA), β-Mercaptoethanol from Merck (Whitehouse Station, USA), chloroform, phenol, and isopropanol from Thermo Fisher Scientific Inc. (Waltham, USA). All solutions were prepared with ultra-pure grade water, purified with a Milli-Q system from Merck Millipore (Darmstadt, Germany) treated with 0.01% of diethyl pyrocarbonate (DEPC) from Sigma-Aldrich (St. Louis, Missouri, USA).

Chromatographic assays were carried out by using Tris-HCl buffer that is prepared with tris(hydroxymethyl)aminomethane (Tris) acquired from Merck (Darmstadt, Germany) and NaCl, acquired from Panreac (Barcelona, Spain). Milli-Q water treated with DEPC with a concentration of 0.01%, was used to prepare these buffers. The chromatographic support used in this work was Capto™ Q ImpRes and HiScale™ 16/20 column (Cytiva, Sweden).

The urea-PAGE gel was prepared with urea from Acros (Waltham, USA), acrylamide, tetramethylethylenediamine (TEMED), ethylenediaminetetraacetic acid disodium salt 2-hydrate (EDTA), ammonium persulfate (APS), Tris-Borate-EDTA (TBE) buffer from

VWR (Leuven, Belgium) and formamide for sample preparation from Merck (Darmstadt, Germany).

For quantitative polymerase chain reaction (qPCR), the Maxima SYBR Green/Fluorescein qPCR Master Mix (2X) purchased from Thermo Fisher Scientific Inc. (Massachusetts, USA), Xpert cDNA Synthesis Kit from Grisp (Porto, Portugal), and specific primers for pre-miRNA-29b-1 amplification synthesized by STAB VIDA (Caparica, Portugal) were used. NZY DNase I was obtained from NZYtech (Lisbon, Portugal) and MEGAclear Transcription Clean-Up Kit was from Thermo Fisher Scientific Inc. (Massachusetts, USA).

3.2. Methods

3.2.1. Optimization of pre-miR-29b production

3.2.1.1. Nucleic acids production in *R. sulfidophilum*

RNA was obtained from *R. sulfidophilum* culture transformed with the plasmid pNZY28-pre-miR-29b-1 [91] containing the sequence of human pre-miRNA-29b. *R. sulfidophilum* was cultured in a solid medium composed of 5 g/L of yeast extract, 10 g/L of glucose, 20 g/L of NaCl, 10 g/L of polypeptone, 4.10 g/L of $MgCl_2$, 15 g/L of agar, 1 mg/L of $ZnSO_4.7H_2O$, 10 mg/L of $MnCl_2.4H_2O$, 10 mg/L of $FeSO_4.7H_2O$, supplemented with 50 μg/mL of kanamycin for 48h. Hereafter, pre-fermentation and fermentation were performed using a culture media composed of 10 g/L of tryptone, 5 g/L of polypeptone, 0.6 g/L of yeast extract, 30 g/L of NaCl, 4 g/L of K_2HPO_4, 1 g/L of KH_2PO_4, 5% (v/v) of glucose, 1% (v/v) of $MgSO_4.7H_2O$, 1% (v/v) $CaCl_2.2H_2O$, 1% (v/v) of $(NH_4)_2SO_4$, 0.1% (v/v) of TES and 0.1% (v/v) of a 50 mg/L kanamycin stock. The medium/O2 ratio was established as 1:5. To start the pre-fermentation process, the inoculum was transferred from the solid medium to the Erlenmeyer flask and left in an orbital shaker for approximately 24 h until it reached an optical density (OD) of 2.6, at 600 nm. After pre-fermentation, the volume needed to start the fermentation with an OD of 0.3 was transferred to another Erlenmeyer flask where fermentation took place. This volume was calculated by the equation below:

$$\mathrm{V\ (pre\ -\ fermentation\ needed)} = \frac{OD\ (fermentation)\ \times\ V\ (fermentation)}{\mathrm{OD\ (pre\ -\ fermentation)\ -\ OD\ (fermentation)}}$$

To measure the OD during the culture, a spectrophotometer Pharmacia Biotech Ultraspec 3000 UV/Visible (Cambridge, England) was used.

The bacteria growth was kept for approximately 36 h and then the medium was centrifuged at 4 °C and 3800 g, for 10 min to obtain cell pellets which were stored at -20°C. Moreover, since *R. sulfidophilum*can produce the biomolecule extracellularly, the resulting liquid medium was also stored at -20 °C.

To optimize the pre-miRNA-29b production, in some particular experiments, the temperature of growth was varied, evaluating the bacteria growth and target RNA production at 25 ºC, 30 ºC, and 37 ºC, also monitoring the production across different culture times (24 h, 36 h, 48 h, and 60 h).

3.2.1.2. Low molecular weight RNA extraction

Intracellular RNA extraction was performed using the acid guanidinium thiocyanate-phenol-chloroform method. Initially, cell pellets were thawed and resuspended in a 0.8% NaCl solution. The suspension was centrifuged at 6000 g at 4 °C for 10 minutes. The supernatant was discarded, and 5 mL of solution D (4 M guanidinium thiocyanate, 0.025 M sodium citrate at pH 7, 0.5% sodium N-lauroylsarcosinate, and 0.1 M β-mercaptoethanol) were added. The mixture was incubated on ice for 10 min. After, 500 µL of 2 M sodium acetate (pH 4) and 5 mL of phenol-concentrated solution were added to the lysis tubes, with careful homogenization. Subsequently, 1 mL of a chloroform/isoamyl alcohol mixture (49:1) was added to the lysate, followed by vigorous shaking. The sample was incubated on ice for 15 min and centrifuged at 10,000 g at 4 °C for 20 min. The addition of phenol/chloroform induces phase separation due to differences in density between the components. Furthermore, the buffer at pH 4 facilitated the effective separation of nucleic acids resulting in three distinct phases: an upper aqueous phase containing RNA, an interphase containing DNA, and a lower organic phase enriched in proteins, lipids, and cellular debris. The next step involved carefully transferring the upper aqueous phase to a new tube to avoid contamination with DNA. The RNA-enriched phase was transferred to a new tube and 5 mL of isopropanol were added to precipitate RNA. After homogenization, the sample was centrifuged at 10,000 g at 4 °C for 20 min. Then, 1.5 mL of solution D and 1.5 mL of isopropanol were added to the pellet. A second centrifugation was then performed under the same conditions (10,000 g, 4 °C, 10 min), and the supernatant was discarded again. Following this, 2.5 mL of 75% ethanol in DEPC-treated water was added, and the mixture was incubated at room temperature for 10–15 min. A final centrifugation at 10,000 g, 4 °C, for 5 min was carried out. After discarding the supernatant, the pellet was air-dried at room temperature for 5–10 min before being resuspended in 1 mL of DEPC-treated water and incubated at room temperature for 10–15 min. The integrity of the RNA samples was assessed through agarose gel electrophoresis, and their concentration was determined using a NanoDrop™ One spectrophotometer (Thermo Scientific, Canada).

3.2.1.3. Nucleic Acids Extracellular Extraction

For extracellular nucleic acids recovery, 10 mL of culture medium were collected and centrifuged at 4304 rpm, for 10 min at 4 ºC. The supernatant was transferred to a new tube and centrifuged at 6600 rpm, for 10 min at 4 ºC to remove the remaining cells. For nucleic acid precipitation, 1 volume of ice-cold isopropanol was added followed by an incubation on ice for 10 min. After, the sample was centrifuged at 6600 rpm, for 10

min, at 4 ºC, and the pellet was air-dried for 10 min at room temperature and solubilized in 1 mL 0.01% of DEPC-treated water. The sample was then stored at -80 ºC, and after defrosting, it was centrifuged at 6600 rpm, for 1 min at room temperature, to remove the salt precipitate. After nucleic acids solubilization, the samples were purified by SEC. For this, a column packed with 5 mL of Sephacryl SF-1000 was used to separate RNA and DNA. The column was equilibrated with 20 mL of buffer (10 mM Tris-EDTA, 150 mM NaCl pH 7), and 500 µL of sample was added. The elution was carried out by adding 8 mL of the same buffer. Fractions of 1 mL were collected for further evaluation. At the end of each assay, columns were washed with DEPC-treated water. The RNA fractions were desalted and concentrated with concentrators Vivaspin 10000 Da (Vivascience).

3.2.1.4. Agarose gel electrophoresis

The analysis of nucleic acids extracted from samples, as well as the peaks obtained from chromatographic assays, was conducted using horizontal electrophoresis on a 1% agarose gel. To visualize the nucleic acids, the gel was prepared with the addition of 0.012 µL/mL Green Safe dye (Porto, Portugal). Electrophoresis was run at 130 V for 40 min in a TAE buffer, consisting of 40 mM Tris base, 20 mM acetic acid, and 1 mM EDTA, adjusted to a pH of 8.0. The gels were subsequently visualized by exposure to ultraviolet (UV) light using a Uvitec Cambridge Fire-Reader equipped with a UV chamber (UVITEC Cambridge, United Kingdom).

3.2.1.5. pre-miR-29b-1 Analysis by Quantitative Polymerase Chain Reaction

pre-miR-29b production was assessed by qPCR. Complementary DNA (cDNA) synthesis was performed using Xpert cDNA Synthesis Kit (Grisp, Portugal), according to the manufacturer's instructions. A total of about 26 ng of RNA sample was used to initiate cDNA synthesis, with a 20 pmol gene-specific primer. Then, polymerase chain reaction (PCR) reactions were carried out, in a 96-well PCR plate (Bio-Rad, USA), using 1 µL of synthesized cDNAs in a 20 µL reaction containing 10 µL of Maxima SYBR Green/Fluorescein qPCR Master Mix (2X) and 0.3 µM of each specific primer. The PCR program was carried out, in a CFX Connected Real-Time PCR Detection System (Bio-Rad, USA), as follows: after an initial denaturation at 95 °C for 5 min, 40 cycles consisting of denaturation at 95 °C for 30 s, annealing at 63 °C for 30 s, and final extension at 72 °C for 30 s. A melting curve analysis from 55 to 95 °C for 10 s with an increment of 0.5 °C was also done. Specific primers for cDNA and qPCR were reconstituted in nuclease-free water to a final concentration of 100 µM.

3.2.2. Chromatography assays

3.2.2.1. Screening of RNA binding/elution with different conditions

Initial chromatographic assays to study the separation of pre-miR-29 from other RNAs were performed in the equipment AKTA Avant with the software UNICORNTM 6.3 (GE Healthcare Biosciences, Uppsala, Sweden). The chromatographic support used was Capto Q ImpRes. This support was packed in a HiScale™ 16/20 column with 16 mm diameter x 10 mm of height. For this, the support was dispersed in Milli-Q water and then added to the column, being extremely important to avoid drying the matrix during this process.

For each chromatographic assay, the column was equilibrated with 10 mM Tris-HCl buffer (pH 7 or 8), previously filtered and sonicated, using a flow rate of 1 mL/min. After equilibration, samples of 100 μg of RNA extracted from *R. sulfidophilum*, were independently injected using a 100 μL loop. Completed the binding step, several steps of increasing salt concentration from 0 to 3 M of NaCl in 10 mM Tris-HCl buffer were applied to analyze different retention patterns and eventual RNA species separation. Different binding and elution conditions were screened in these assays, namely the working pH values (7 or 8), the amount (50 or 100 μg) of loaded RNA sample, and the NaCl concentration used to perform the elution steps. All experiments were performed at room temperature and the absorbance of eluted species was continuously monitored at 260 nm. The fractions regarding the elution peaks were recovered and further desalted with concentrators Vivaspin 10000 kDa (Vivascience) until reaching about 100 μL, being finally analysed by Urea-PAGE.

3.2.2.2. Experimental design for optimization of pre-miR-29 purification

The two-level factorial design with tree variables (2^3) was used as the model for the optimization of pre-miRNA-29b purification. The factors studied were the NaCl concentration in the binding step, NaCl concentration in the elution step, and buffer pH (Table 1). Each variable was coded by letters A, B, and C, and the first two had 3 levels: -1, 0, and 1, and one had only two levels -1 and 1. The software used was Design Expert (version 13), and it generated 28 experiments to be performed, and one response was established: purity percentage (%) of pre-miRNA-29b (R1). The purity was calculated by measuring the intensity of the band of pre-miRNA-29b relative to all the bands present in the analyzed lane. To measure the intensity of the bands, it was used the ImageJ 1.54g software.

Table 1 – Factors and respective low, medium, and high values for the optimization of pre-miRNA-29b purification by DoE.

Code	Factor	Levels		
		Low	Medium	High
A	[NaCl] binding step (M)	0.540	0.585	0.630
B	[NaCl] elution step (M)	0.630	0.690	0.750
C	Buffer pH	7	-	8

3.2.2.3. Urea-PAGE analysis

Urea-PAGE was performed in order to analyze if the separation conditions used in the chromatographic assays were able to properly separate the pre-miR-29b from the remaining contaminants. Gels were prepared by firstly mixing 6 g of urea, 5 mL of MilliQ-Water, and 1.5 mL of 10× TBE Buffer (0.89 M tris base, 0.89 M boric acid and 0.02 M EDTA), followed by heating at 40 °C. Afterwards, 5 mL of Acrylamide/Bisacrylamide Solution (37.5:1) was added to the mixture, followed by the addition of 25 µL of TEMED and 100 µL of PSA 10% to initiate the polymerization process. Once the polymerization of the gels occurred, a pre-run was performed at 133 V for 15 min to wash the wells. Samples were prepared by mixing 10 µL of RNA-containing sample and 10 µL of formamide followed by heating at 60 °C, for 5 min. Then, 20 µL of each sample was injected into the wells. The electrophoresis was run at 133 V for 60 min. For the visualization of nucleic acids, the gel was incubated with 0.01% of Green Safe in TBE buffer (Porto, Portugal) and it was further revealed using UV light exposure in the Bio-Rad equipment using the ChemiDoc software.

Chapter 4 - Results and Discussion

4.1. Optimization of pre-miRNA- 29b production

4.1.1. Growth conditions optimization

Optimizing the growth of *R. sulfidophilum* is crucial for maximizing its potential, particularly in the production of extracellular RNA. As previously mentioned, this bacterium not only synthesizes substantial amounts of RNA but also secretes it into the extracellular environment. From a biotechnological perspective, optimizing growth temperature and time is essential to enhance RNA yield. Proper temperature regulation boosts metabolic activity, while also adjusting growth time, thus ensuring an enhanced RNA production and secretion. It's possible to achieve higher RNA output by fine-tuning these factors, making production processes more efficient and cost-effective.

With this in mind, *R. sulfidophilum* growth conditions were optimized by focusing on temperature as a key variable. For this, was selected three temperatures for comparative analysis: 25 °C, 30 °C, and 37 °C. The temperature of 30 °C was chosen as a reference, as previous studies in the literature have demonstrated its suitability [47, 91-93]. However, to broaden our understanding and explore the potential for enhanced RNA production, it was included two additional temperatures, 25 °C and 37 °C, that could offer new insights into optimal growth conditions.

A preliminary test was conducted in which the bacteria were grown on agar plates at each of the three temperatures. This approach provided a rapid and visual analysis to assess whether the chosen temperatures would support continued experimental investigation (Figure 9). By comparing overall growth across the different temperature conditions, we aimed to determine whether it would be worthwhile to pursue further optimization at these specific temperatures, ultimately guiding the selection of the most promising conditions for enhanced *R. sulfidophilum* growth.

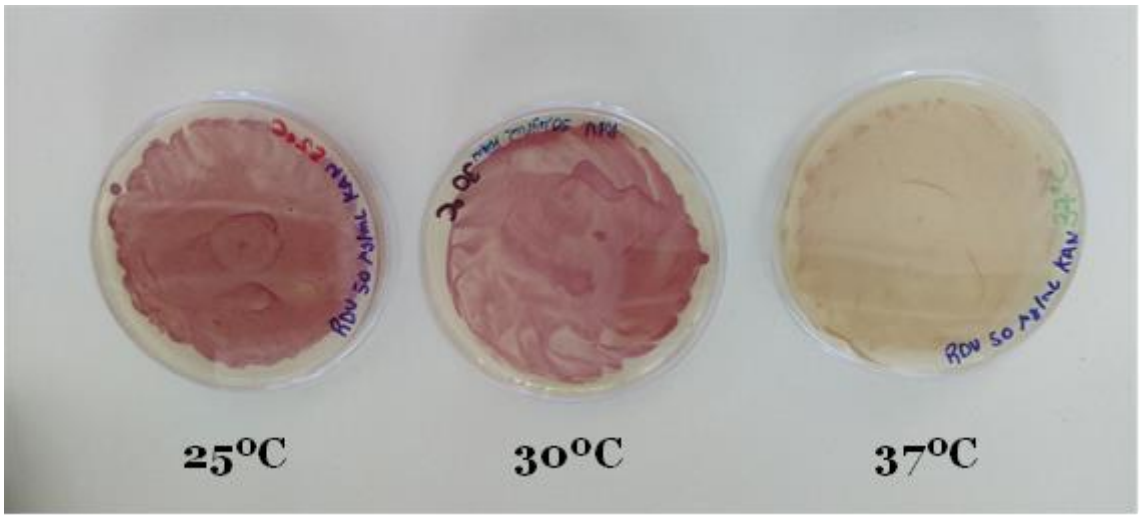

Figure 9 – Growth of *R. sulfidophilum* on agar plates at three different temperatures (25 °C, 30 °C and 37 °C respectively) with 48 h of growth time.

Upon examining the agar plates, was able to draw immediate conclusions regarding the suitability of the selected temperatures. Notably, the 37 °C condition showed minimal bacterial growth, with colonies failing to develop at the same rate as those at lower temperatures. This lack of growth led us to exclude 37 °C from further experimentation, as it did not support the intended enhanced growth conditions.

Between the remaining two temperatures, 25 °C and 30 °C, it was not immediately clear which would be more beneficial in promoting bacterial growth and maximizing RNA production. Therefore, it was decided to proceed with further optimization studies focusing on these two temperatures, to monitor the growth rate for a longer period of time in liquid medium. To achieve this, it was conducted a time-course study by taking samples at regular 12-hour intervals, from 0 up to 72 h of cultivation. These samples were used to construct detailed growth curves for *R. sulfidophilum* under both temperature conditions (Figure 10).

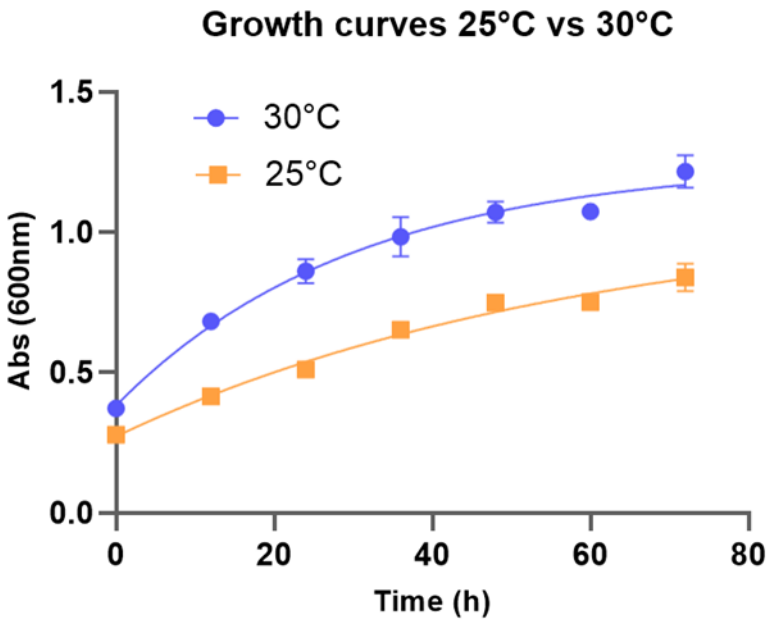

Figure 10 – Growth curves of *R. sulfidophilum* obtained by the culture at 25 ºC and 30 ºC, at 250 rpm, under dark aerobic conditions.

Examining the growth curves, it is revealed a clear distinction in bacterial proliferation at the two tested temperatures. From a starting OD_{600nm} of 0.3, the growth at 30 °C consistently surpassed the growth at 25 °C. Additionally, the growth curve at 25°C appeared almost linear, lacking the identification of the growth phases slightly observed at 30 °C, where there is some definition of the exponential and stationary phases, which is indicative of more robust metabolic activity [94].

Although the sampling intervals were sufficient for our preliminary assessment, taking more frequent samples over a longer period of time would offer a more precise growth curve. However, it is evident that 30 °C supports a higher overall growth, which suggests it is the temperature of choice for bacterial proliferation. However, the ultimate goal is to maximize pre-miRNA-29b production, regardless of the growth rate. Therefore, even though 30 °C resulted in more substantial growth, this does not necessarily guarantee higher pre-miRNA-29b production. Given this, the study proceeded to evaluate the relationship between bacterial growth and RNA yield, considering that higher biomass might not directly correlate with increased RNA secretion.

4.1.2. Nucleic Acids Extracellular Extraction

The next step in the process is RNA extraction, in which the samples (extracellular medium) resulting from the cultivation are subjected to an extraction protocol. This step is crucial for the recovery of the target RNA from the bacterial culture. Following the extraction, agarose gel electrophoresis is performed to assess the quality and integrity of the extracted nucleic acids sample. This initial electrophoresis provides a visual confirmation of successful RNA extraction, allowing us to evaluate the presence and size distribution of both DNA and RNA in the samples (Figure 11, A).

After this, the samples undergo purification using a benchtop column with a size exclusion matrix, the Sephacryl SF-1000. This clarification step is essential for removing any remaining DNA, which could interfere with the downstream process. After this purification, the samples are concentrated and desalted, ensuring RNA quality for further analysis and chromatographic procedures. A second agarose gel electrophoresis is then conducted (Figure 11, B) to verify the efficiency of DNA removal. By comparing the pre- and post-purification electrophoresis results, we can confirm the effective removal of DNA, ensuring that the samples contain only RNA for the development of subsequent pre-miRNA-29b purification strategies.

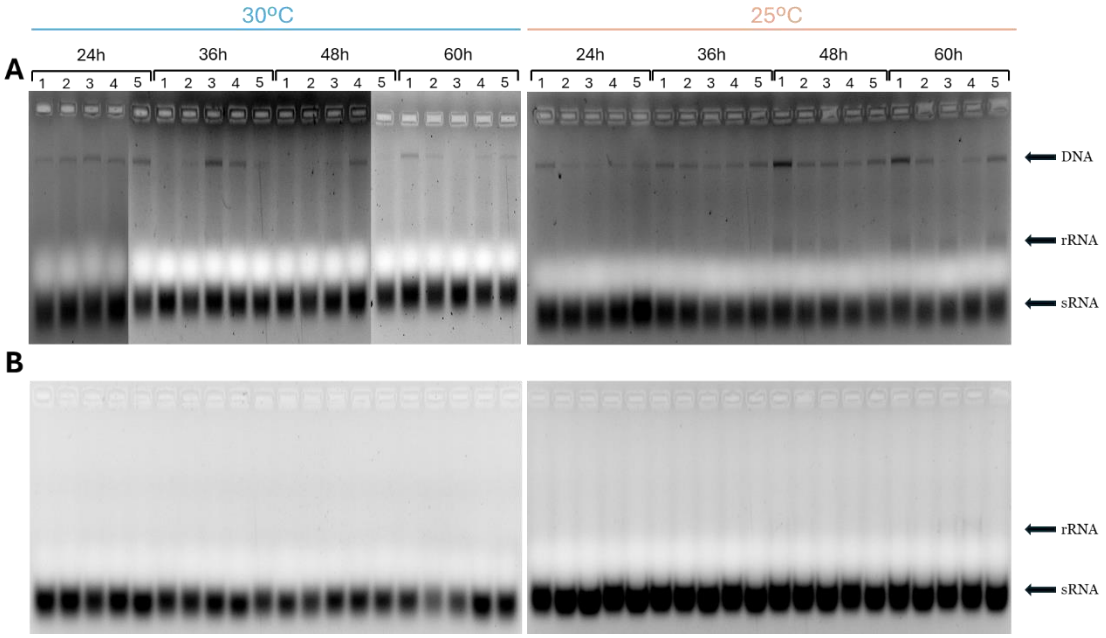

Figure 11 – Agarose gels of low molecular weight RNA extractions. A – Analysis of the samples recovered from the extraction, before purification by SEC; B – Analysis of the samples after purification by SEC.

As previously mentioned, agarose gel electrophoresis provides valuable information regarding the presence and distribution of RNA, as well as potential contaminants like DNA. After performing size exclusion chromatography, was observed that the DNA was

successfully eliminated from the samples, as indicated by the absence of DNA bands in the post-purification gels. Additionally, the distribution of RNA in the gels was consistent with the expected size range for low molecular weight RNA, confirming the effectiveness of our extraction and purification process.

However, an interesting observation was made in the samples from cultivation at 25 °C, particularly at the longer fermentation times of 48 and 60 hours. In these samples, there was noticeable secretion of rRNA, which was not observed in samples from cultivation at 30 °C. This rRNA accumulation at 25 °C suggests that the bacteria may experience stress due to the suboptimal growth temperature, leading to the overproduction of rRNA [95]. This stress response highlights the importance of optimizing growth conditions, as non-ideal temperatures like 25 °C can negatively affect the desired RNA profile, particularly during extended fermentation periods.

4.1.3. Quantification assays

While agarose gel electrophoresis provides useful insights into the presence and distribution of different types of RNA, it is ultimately a qualitative technique. This means it offers only general information about which RNA types are present, without giving precise quantitative data. In this case, the target pre-miRNA-29b falls within the low molecular weight RNA band, making it impossible to determine its specific concentration or distinguish it from other small RNAs using the agarose gel electrophoresis alone.

Given these limitations, a more precise technique is required to accurately quantify the amount of pre-miRNA-29b. For this purpose, a quantitative method such as qPCR is necessary to measure the exact levels of pre-miRNA-29b. Figure 12 A and B, presents the qPCR results for pre-miRNA-29b levels achieved in the cultures conducted at 25 °C and 30 °C, respectively. For each temperature, five independent replicates were prepared to reduce variability and minimize errors. However, the results exhibited a relatively high standard deviation, indicating significant variability across the replicates.

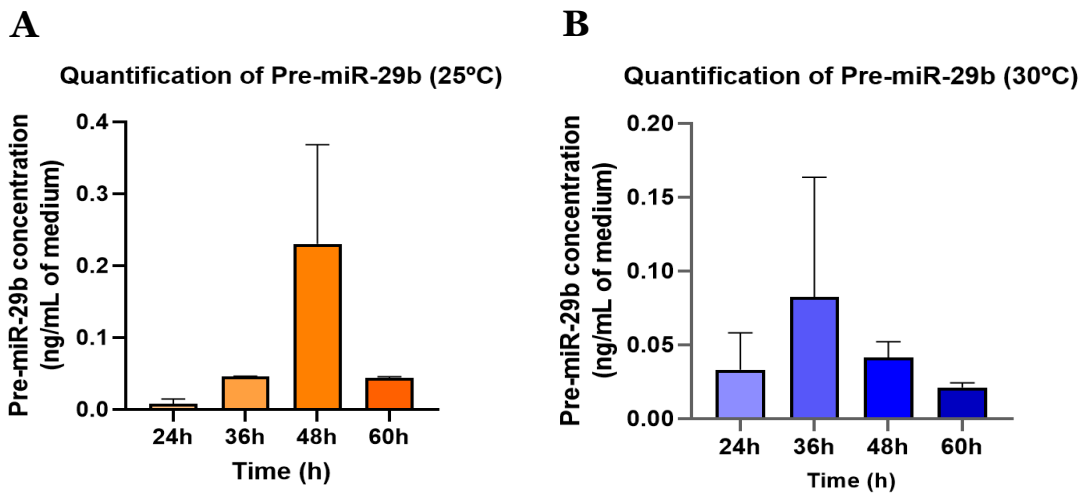

Figure 12 – Intracellular pre-miR-29b levels at 24, 36, 48 and 60 hours of cultivation. A - pre-miR-29b levels obtained at 25 ºC. B - pre-miR-29b levels obtained at 30 ºC.

As previously noted, the results obtained showed a high standard deviation, essentially in time points in which pre-miRNA-29b levels were increased. Since this work is working with live bacterial cells, it is challenging to eliminate all sources of variation, even when using the same experimental setup, bench, and pre-fermentation conditions for all replicates. This biological variability likely accounts for the noticeable standard deviation bars in the data. Despite these fluctuations, these results still provide valuable insights into the trends and overall production of pre-miRNA-29b under the tested conditions.

Analysing Figure 12 A, which displays the quantification of pre-miRNA-29b at 25 ºC, was observe that the highest production occurs at 48 hours (0.23 ng/mL of medium). There is a noticeable increase in pre-miRNA-29b levels from the start, reaching its highest value at 48 hours, followed by a decline at 60 hours. Regarding the Figure 12 B, the results at 30 ºC show a different tendency. Here, there is a steady increase in pre-miRNA-29b production, peaking earlier at 36 hours (0.08 ng/mL of medium), followed by a gradual reduction through the following time of the culture. Comparing the values for the two temperatures, it is suggested that there is a higher production of pre-miRNA-29b at 48 hours at 25 ºC, but the error bar must be considered as it can influence this analysis. Moreover, as previously noted, at 25 ºC there is increased production of other RNA types, like rRNA, particularly for 48 and 60 hours, as shown by the agarose electrophoresis results (Figure 11 A). Given that the literature supports 30 ºC as an optimal growth temperature [47, 57, 69, 93], and considering that at this temperature the pre-miRNA-29b production peak occurs earlier, at 36 hours of culture,

making the process faster than at 25 ºC, it was selected the temperature of 30 ºC to proceed for the purification studies.

4.2. Screening of optimal conditions for pre-miRNA-29b purification

After identifying the ideal conditions for producing pre-miRNA-29b, it was necessary to proceed to the most critical phase, which is the purification process. The primary goal of this step was to determine the optimal conditions for purifying the target RNA. Since obtaining substantial quantities of pre-miRNA-29b extracellularly is more challenging, we initially focused on the screening of favourable purification conditions using intracellular samples, where larger quantities of RNA could be obtained.

A commercial Capto Q ImpRes resin was selected for this purification testing that presents multimodal characteristics to maximize the interactions with the target pre-miRNA-29b. The multimodal nature of this column allows the combination of ionic and hydrophobic interactions, which was expected to enhance the selectivity for pre-miRNA-29b while effectively eliminating other contaminants.

Figure 13 illustrates the Capto Q ImpRes ligand, which features a quaternary ammonium group, a strong anion exchanger. This positively charged group interacts with negatively charged molecules like the phosphate backbone of RNA. The ligand is stable across a wide pH range, providing reliable binding under diverse conditions. The quaternary ammonium's strong ionic interactions make it particularly effective for binding nucleic acids such as pre-miRNA-29b, while allowing selective elution during purification.

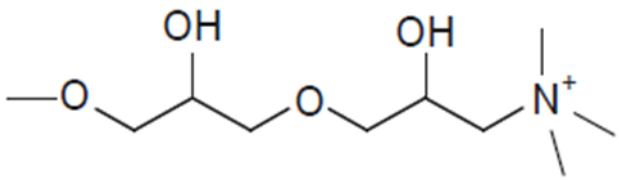

Figure 13 – Capto Q ImpRes ligand structure.

The initial screening was initiated with the implementation of a similar procedure already described in the literature for this matrix, originally designed to separate single-stranded DNA (ssDNA) and DNA (dsDNA) [96]. Although our target, the pre-miRNA-29b, differs from these nucleic acids, this initial strategy could provide insights into the

matrix behaviour with our specific sample. This approach allowed us to assess the resin performance and optimize conditions for the effective purification of pre-miRNA-29b. Thus, for the first experiment, a linear gradient of NaCl between 0 and 3 M was established, and the results are schematically presented in Figure 14.

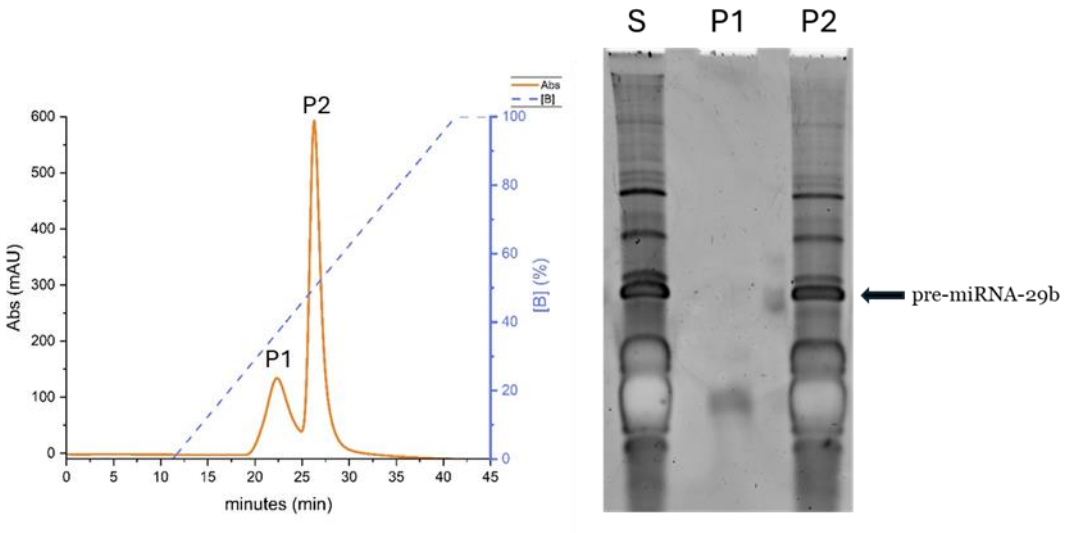

Figure 14 - Representative chromatograms of 50 µg of low molecular weight RNA sample injected in a Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 8; elution buffer 10 mM Tris-HCl, pH 8 with 3 M NaCl; Linear gradient from 0 to 3 M of NaCl. S – Sample; P1 -First peak; P2- Second peak.

The chromatogram and Urea-PAGE results show two distinct peaks, with the second peak containing nearly the entire sample, indicating that most of the material eluted during this phase. Based on the NaCl concentrations at the point of elution, was estimate that a suitable concentration to consider as a first step in a stepwise gradient could be around 20% of the 3 M elution buffer, which corresponds to approximately 0.6 M NaCl, to optimize further experiments. With this information, we proceeded to confirm the binding and elution profile with this new gradient, considering the use of 0.6 M NaCl as the first step. To expand the hypotheses of separation, it was also tested the condition of 0.75 M (Figure 15) in the binding step. This additional test would provide more insight into the elution profile and the behaviour of the sample under varying salt conditions, helping to better understand the optimal conditions for the pre-miRNA-29b purification.

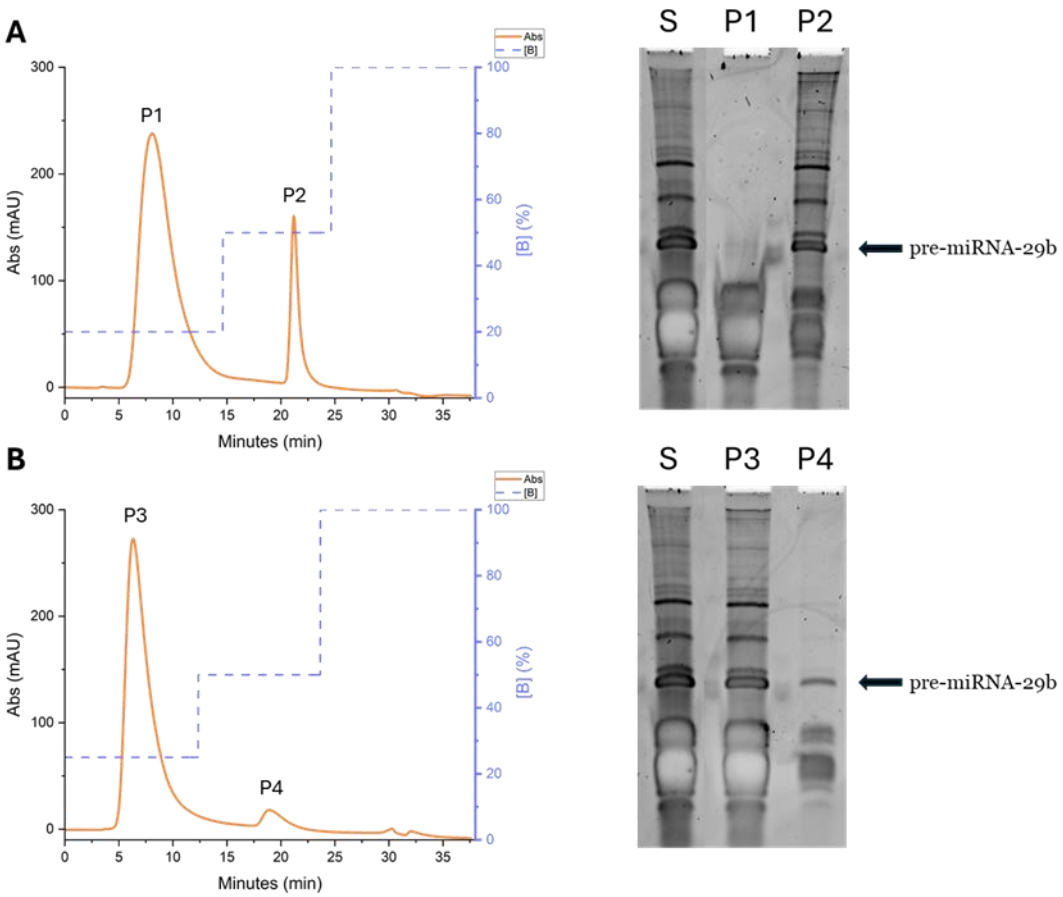

Figure 15 – Representative chromatogram of 50 µg of low molecular weight RNA sample injected in a Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 8; elution buffer 10 mM Tris-HCl, pH 8 with 3 M NaCl. **A**: S– Sample; P1 -First peak obtained at 0.6 M NaCl; P2- Second peak obtained at 1.5 M NaCl. **B**: S – Sample; P3 -First peak obtained at 0.75 M NaCl; P4 - Second peak obtained at 1.5 M NaCl.

Referring to Figure 15, which illustrates the mentioned gradients, it is evident that 0.75 M NaCl is not the ideal concentration for the binding step. Although contaminant removal is effective at the first peak, too much of the target is lost at this stage. Meanwhile, with 0.6 M of NaCl some contaminants are removed in the first peak, but the second peak still contains many impurities that co-elute with the target RNA. Both conditions highlight the need for further optimization to balance contaminant elimination and target retention.

With the previous experiments, it was realized that a single elution step was insufficient for effective separation, so an intermediate step was introduced. This approach aimed to remove contaminants during the first step, aiming for the target to be more isolated in the second peak, as has been done previously in the literature [69]. By refining the elution process, it is expected to achieve better separation of pre-miRNA-29b from

impurities while maximizing recovery. Following this strategy, two assays were performed using slightly different NaCl concentrations to assess the matrix sensitivity to the changes in NaCl concentration. So, 0.675 M and 0.705 M NaCl were established for the binding step, respectively to fine-tune the elution process. The results of these tests are presented in Figure 16, providing insights into how a slight concentration adjustment impacts the separation and purification of our target.

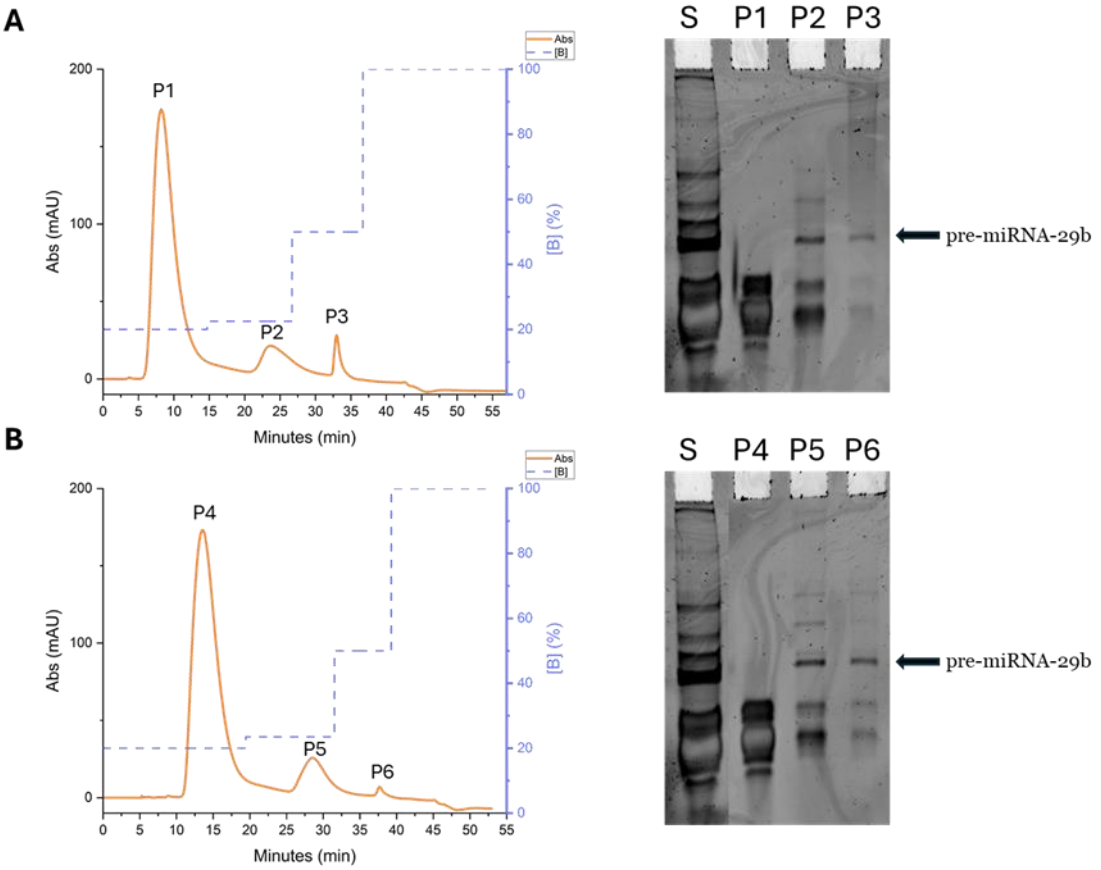

Figure 16 - Representative chromatograms of 50 µg of low molecular weight RNA sample injected into the Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 8; elution buffer 10 mM Tris-HCl, pH 8 with 3 M NaCl. **A**: S – Sample; P1 -First peak obtained at 0.6 M NaCl; P2- Second peak obtained at 0.675 M NaCl; P3- Third peak obtained at 1.5 M NaCl; **B**: S – Sample; P4 - First peak obtained at 0.6 M NaCl; P5- Second peak obtained at 0.705 M NaCl; P6- Third peak obtained at 1.5 M NaCl.

The results revealed that the matrix responds well to small changes in NaCl concentrations, as shown by the improved isolation of the RNA target in the third peak. However, when increasing the NaCl concentration from 0.675 M to 0.705 M, it is confirmed a higher concentration of the pre-miR-29b in the peak of interest, but more contaminants are also co-eluting in this peak. Despite this, the first peak shows no trace of our target, suggesting that the 0.6 M NaCl successfully removes a considerable

amount of contaminants in the initial stage. With two elution steps, it was almost achieved the isolation of the target RNA, though some contaminants remain. This may indicate already a promising way to find the best conditions for efficient pre-miRNA-29b purification.

With this in mind, it was decided to increase the binding step by 0.03 M, as it was not previously identified any loss of the target in this phase, aiming to eliminate additional contaminants. Alongside this adjustment, it was also tested a different buffer pH. While all the previous experiments were performed at pH 8, it was now evaluated the use of buffers with pH 7, to assess potential improvements in the target purification. The results of these modifications are presented in Figure 17.

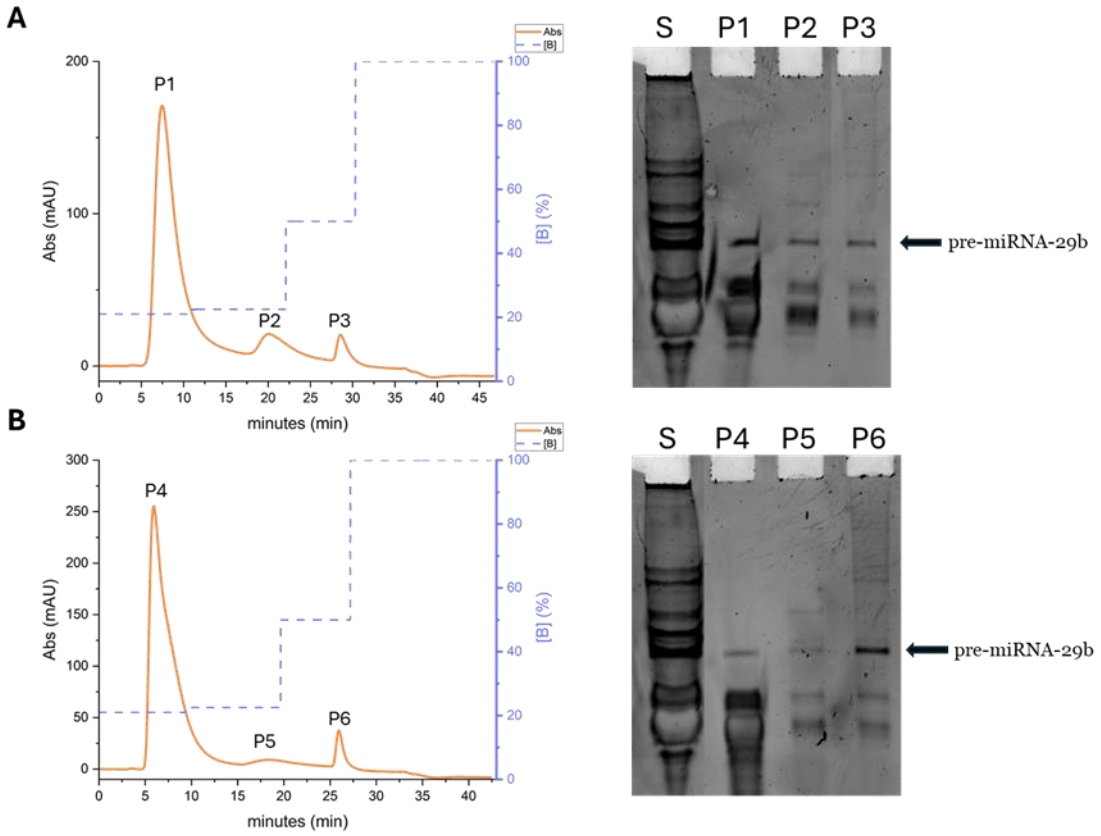

Figure 17 - Representative chromatograms of 50 µg of low molecular weight RNA sample injected in a Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 7 and pH 8; elution buffer 10 mM Tris-HCl, pH 7 and pH 8 with 3 M NaCl. **A**: Buffer with pH 7; S – Sample; P1 - First peak obtained at 0.63 M NaCl; P2- Second peak obtained at 0.675 M NaCl; P3- Third peak obtained at 1.5 M NaCl; **B**: Buffer with pH 8; S – Sample; P4 -First peak obtained at 0.63 M NaCl; P5- Second peak obtained at 0.675 M NaCl; P6- Third peak obtained at 1.5 M NaCl.

The image above shows the impact of pH and the small change in the equilibrium buffer on the retention of different species in the column. There was observed some loss of

pre-miRNA-29b during the binding step, especially at pH 7 (Figure 17 A), although the intensity of the band was higher at pH 8. At pH 8, our target molecule is concentrated in the peak of interest, though there is still some distribution to the other peaks, mainly in the binding step. This could be mitigated by slightly reducing the NaCl concentration. Interestingly, although the matrix is designed to maintain its charge across different pH levels, we noticed a clear difference between pH 8 and pH 7 in its performance.

In a previous study developed by our research group, the resin L-arginine-Sepharose 4B was used for pre-miRNA-29b purification. With Tris buffer at pH 8, it was achieved the effective separation of pre-miRNA-29b using a NaCl concentration of 0.360 M [69]. When compared to the 1.5 M concentration used in this study, it is evident that the interactions between our target molecule and the column are stronger, as a higher NaCl concentration is required for the elution of the target RNA. Furthermore, regarding the usually obtained second peak in the chromatograms of this work (Figure 17, A, P2 and B, P5), higher NaCl concentrations were needed compared to those in the elution step of the reference study (0.675 M to 0.360 M). In fact, the concentration selected for the equilibration step in this study is nearly double the concentration used to achieve the best elution results in the arginine-based study, further corroborating that the interactions with this ligand are significantly stronger. Nonetheless these retention differences, the elution behaviour is similar for both resins and the kind of interaction established, is also similar, with electrostatic forces playing a crucial role. With arginine being an affinity chromatography approach, this similarity in behaviour emphasizes the multimodal character of Capto Q Impres.

4.2.1. Extracellular optimization

After establishing a starting point regarding salt gradient with intracellular samples, extracellular samples for pre-miRNA-29b purification were tested under the previously optimized conditions. The main goal was to determine whether the extracellular sample behaved similarly to the intracellular one. It was initially noticed that while the extracellular sample appeared clarified, it was more challenging to detect on urea-PAGE, causing difficulties in visualizing the species eluting in each peak. This represented a new challenge in assessing the purification process for extracellular pre-miRNA-29b.

This was not an issue when working with intracellular samples but obtaining substantial amounts of extracellular samples presented a challenge. Running all the samples through the size exclusion chromatography column would be too time-consuming for the required volumes. Therefore, it was tested whether this step was

mandatory. A comparison was performed by injecting samples recovered from the extraction, after being subjected to the concentration step, and after both SEC and concentration to evaluate the differences and determine the need of this additional step. Results are presented in Figure 18.

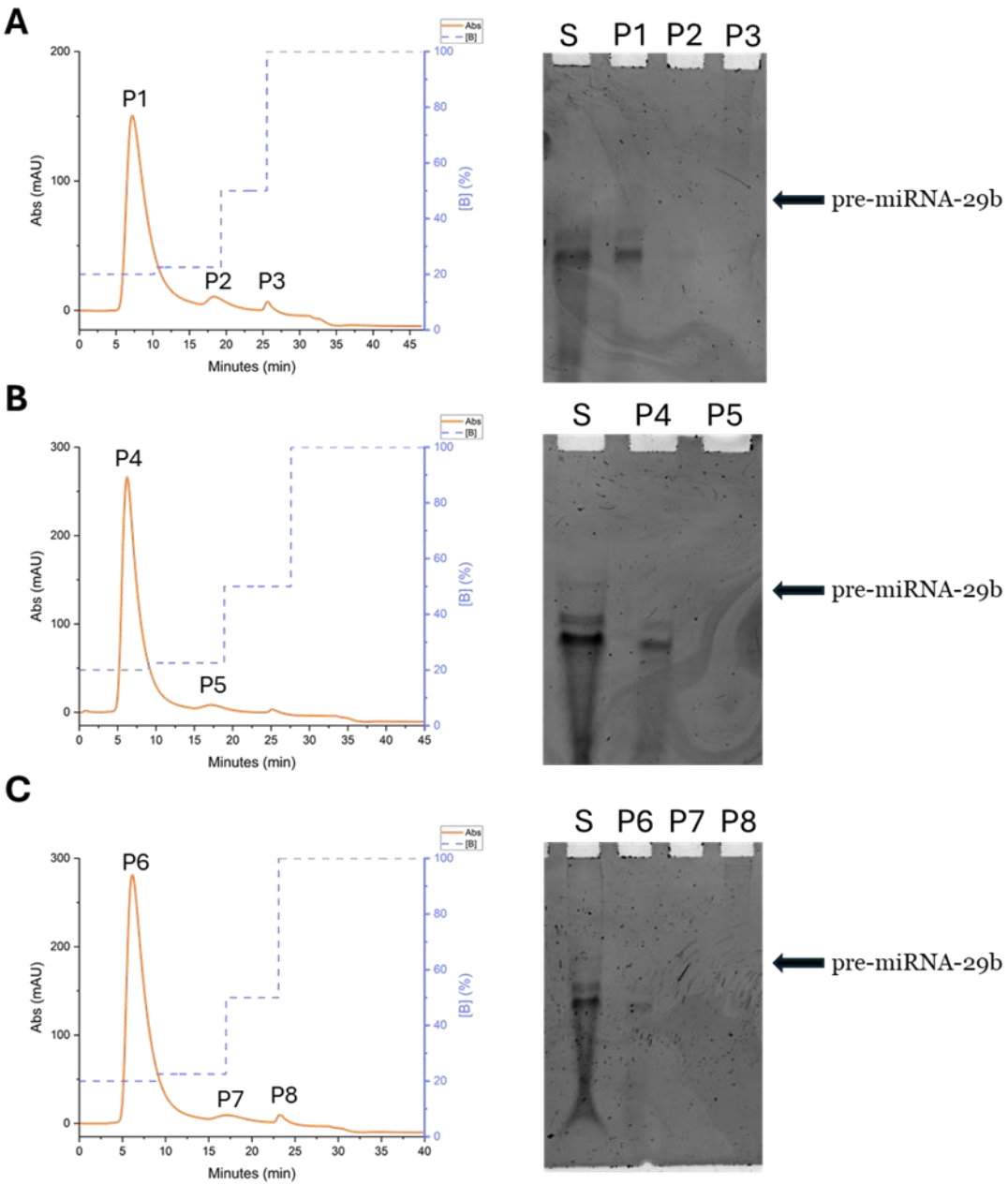

Figure 18 - Representative chromatograms of 50 µg of low molecular weight RNA sample injected in a Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 8; elution buffer 10 mM Tris-HCl, pH 8 with 3 M NaCl. **A**: Sample after extraction; S – Sample; P1 -First peak obtained at 0.60 M NaCl; P2- Second peak obtained at 0.675 M NaCl; P3- Third peak obtained at 1.5 M NaCl; **B**: Sample after SEC and concentration; S – Sample; P4 -First peak obtained at 0.60 M NaCl; P5- Second peak obtained at 0.675 M NaCl; **C**: Sample after concentration; S – Sample; P6 -First peak obtained at 0.60 M NaCl; P7- Second peak obtained at 0.675 M NaCl; P8- Third peak obtained at 1.5 M NaCl.

The image above highlights the differences in sample preparation. As previously noted, visualizing our target in the gel remains challenging. However, there is a noticeable improvement in band intensity for the sample that underwent SEC and concentration (Figure 18 B) compared to the others (Figure 18 A and C). Despite this, the detection of the RNA target was not possible at the expected peak, even when the target was visible in the initial sample. This issue occurred consistently across all three tests.

Despite the difficulty in detecting the bands, the concentration step could be helpful, and since concentrating the sample does not substantially increase the duration of the process, it was decided to concentrate the sample immediately after extraction, thereby eliminating the need for the SEC step.

For carefulness, the test was repeated, this time adding an intracellular sample to the gel to compare the migration of the target band in the gel and changing the first step of the gradient from 0.60 M to 0.54 M of NaCl, to try to induce elution in the second step. Moreover, it was also tested the process of concentrating the sample prior to injection. This involved passing the sample through concentrators to decrease its dilution factor, allowing us to inject the same mass with smaller volumes. Additionally, this approach was used to observe the potential effect of concentration on the visualization of the bands. By reducing the volume, we aimed to enhance the intensity of the bands, facilitating the visualization and achieving a more accurate analysis of pre-miRNA-29b during the subsequent steps. These results are shown in Figure 19, offering clearer insight into the bands position and the impact of concentration on the injection process.

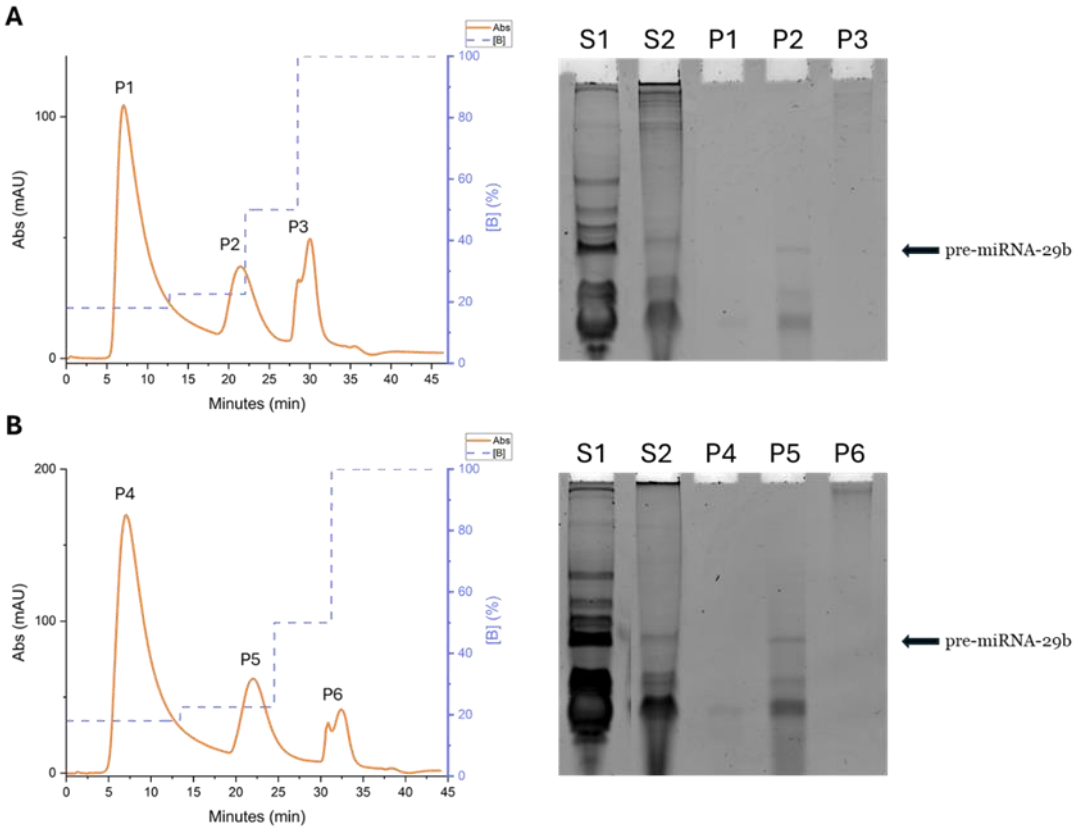

Figure 19 - Representative chromatograms of 50 µg of low molecular weight RNA sample injected in a Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 8; elution buffer 10 mM Tris-HCl, pH 8 with 3 M NaCl. **A**: Sample before extraction; S1 – Intercellular sample; S2 – Extracellular sample; P1 -First peak obtained at 0.54 M NaCl; P2- Second peak obtained at 0.675 M NaCl; P3- Third peak obtained at 1.5 M NaCl; **B**: Sample after extraction; S1 – Intercellular sample; S2 – Extracellular sample; P4 -First peak obtained at 0.54 M NaCl; P5- Second peak obtained at 0.675 M NaCl; P6- Third peak obtained at 1.5M.

With these new trials, it was possible to visualize the pre-miR-29b band in both assays. Unlike the previous experiments, these results highlighted the importance of concentrating the sample before injection, as the target band intensity increased significantly. Overall, the intensity of the bands also improved. This reinforced our decision to concentrate the sample before injection, eliminating any previous doubts about the effectiveness of this step.

After this, the study was carried out whether increasing the injected mass would enhance the analysis and band visualization on urea-PAGE. The concentration step gave us the flexibility to increase the injected mass, so we explored whether this would improve sample detection. At the same time, two binding conditions were evaluated, namely by using 0.54 M and 0.60 M NaCl. The previous tests indicated 0.60 M was ideal, but the modification to 0.54 M aimed to elute contaminants at lower

concentrations while still ensuring a sufficiently pure pre-miR-29b in the second peak. So, this experiment aimed to determine which buffer worked best and whether increasing the injected mass from 50 µg to 100 µg would improve results (Figure 20).

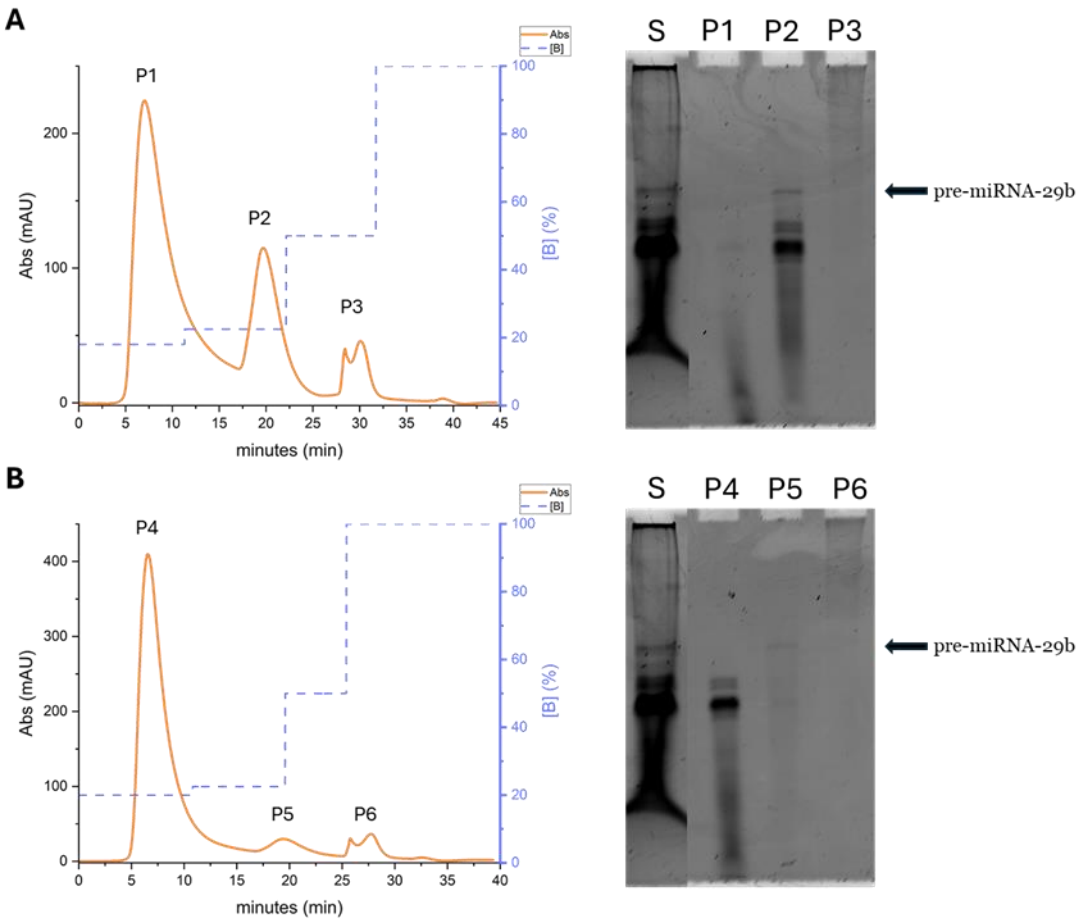

Figure 20 - Representative chromatograms of 100 µg of low molecular weight RNA sample injected in a Capto Q ImpRes column and respective UREA-PAGE. Equilibration buffer 10 mM Tris-HCl, pH 8; elution buffer 10 mM Tris-HCl, pH 8 with 3 M NaCl. **A**: Equilibrium with 0.54 M NaCl; S – Sample; P1 -First peak obtained at 0.54 M NaCl; P2- Second peak obtained at 0.675 M NaCl; P3- Third peak obtained at 1.5 M NaCl; **B**: Equilibrium with 0.60 M NaCl; S – Sample; P4 -First peak obtained at 0.60 M NaCl; P5- Second peak obtained at 0.675 M NaCl; P6- Third peak obtained at 1.5M.

The image above shows that using 0.54 M NaCl in the first gradient step causes the majority of contaminants to co-elute in peak 2 with the pre-miRNA-29b (Figure 20 A). This shows that a lower NaCl concentration is not suitable for eliminating contaminants, comparatively to what is needed for the intracellular sample. This is evident from the lower intensity of the first peak and the stronger intensity of the second peak. The gel supports this observation, with nearly all the bands concentrated in the second peak and displaying good intensity.

Increasing the NaCl concentration from 0.54 M to 0.60 M had a noticeable effect on the pre-miR-29b elution. Almost all contaminants were eliminated in the first gradient step, leaving pre-miR-29b practically isolated in the elution step (peak 2). Unlike the previous example, the slightly higher NaCl concentration allowed the matrix to effectively elute contaminants in the first peak, leaving the second peak for our target. However, when comparing the band intensity of the initial sample with the band of the second peak, it's possible to see some loss in other peaks, even though no visible bands appear in those peaks. Quantifying band intensity revealed that the purity percentage of the second peak was 25.39% for the assay with 0.54 M NaCl and 58.30% for the assay with 0.6 M NaCl. These results align with the observations from the gel, confirming our expectations regarding the impact of NaCl concentration on purity.

Despite the improvement and the significant isolation of our target, we decided to proceed with further optimization. This time, the parameters that appeared most impactful during previous optimizations for effective pre-miR-29b purification with this matrix were prioritized. An experimental study was designed around these parameters to explore potential further improvements.

4.3. Design of Experiments (DoE)

For the experimental design, it was selected a two-level factorial design because it offers a systematic, efficient, and insightful method for optimizing the purification of pre-miR-29b by evaluating the combined effects of NaCl concentrations in both the equilibrium and elution steps and pH. It provides a clear understanding of the interactions between these variables, identifies the most critical factors, and allows for the fine-tuning of conditions for maximum purification efficiency.

The analysis of the FFD 2^3 model resulted in 28 experiments, in which there were obtained purity percentage results ranging from 3.07% to 100% (Table 2).

Table 2 – FFD runs and obtained responses (A: [NaCl] in binding, B: [NaCl] in elution, C: Buffer pH, R1: purity percentage (%) of pre-miRNA-29b).

Run	[NaCl] in binding (A)	[NaCl] in elution (B)	Buffer pH (C)	Purity percentage (%) of pre-miRNA-29b (R1)
1	0.54	0.75	7	4.78
2	0.585	0.69	8	0
3	0.63	0.63	7	0
4	0.585	0.69	8	9.58
5	0.63	0.75	7	0
6	0.585	0.69	7	0
7	0.54	0.63	7	3.07
8	0.585	0.69	8	0
9	0.54	0.63	8	12.05
10	0.63	0.63	8	0
11	0.585	0.69	7	0
12	0.63	0.75	8	0
13	0.585	0.69	7	0
14	0.54	0.75	8	0
15	0.54	0.63	7	0
16	0.585	0.69	8	0
17	0.54	0.63	8	6.51
18	0.585	0.69	8	0
19	0.63	0.75	7	0
20	0.585	0.69	8	30.36
21	0.585	0.69	7	100
22	0.63	0.75	8	0
23	0.63	0.63	8	0
24	0.63	0.63	7	0
25	0.54	0.75	7	11.73
26	0.585	0.69	7	0
27	0.585	0.69	7	0
28	0.54	0.75	8	8.38

There are multiple results of zero purity of pre-miRNA-29b due to the low detection of the bands and even the bands that could be seen were too faint to achieve a good quantification (Figure 21). As a result, the program was unable to generate any reliable predictions due to numerous tests lacking a measurable response, and even those with quantifiable responses showed low values. Although there was a point with 100% purity, a very low sample concentration was obtained.

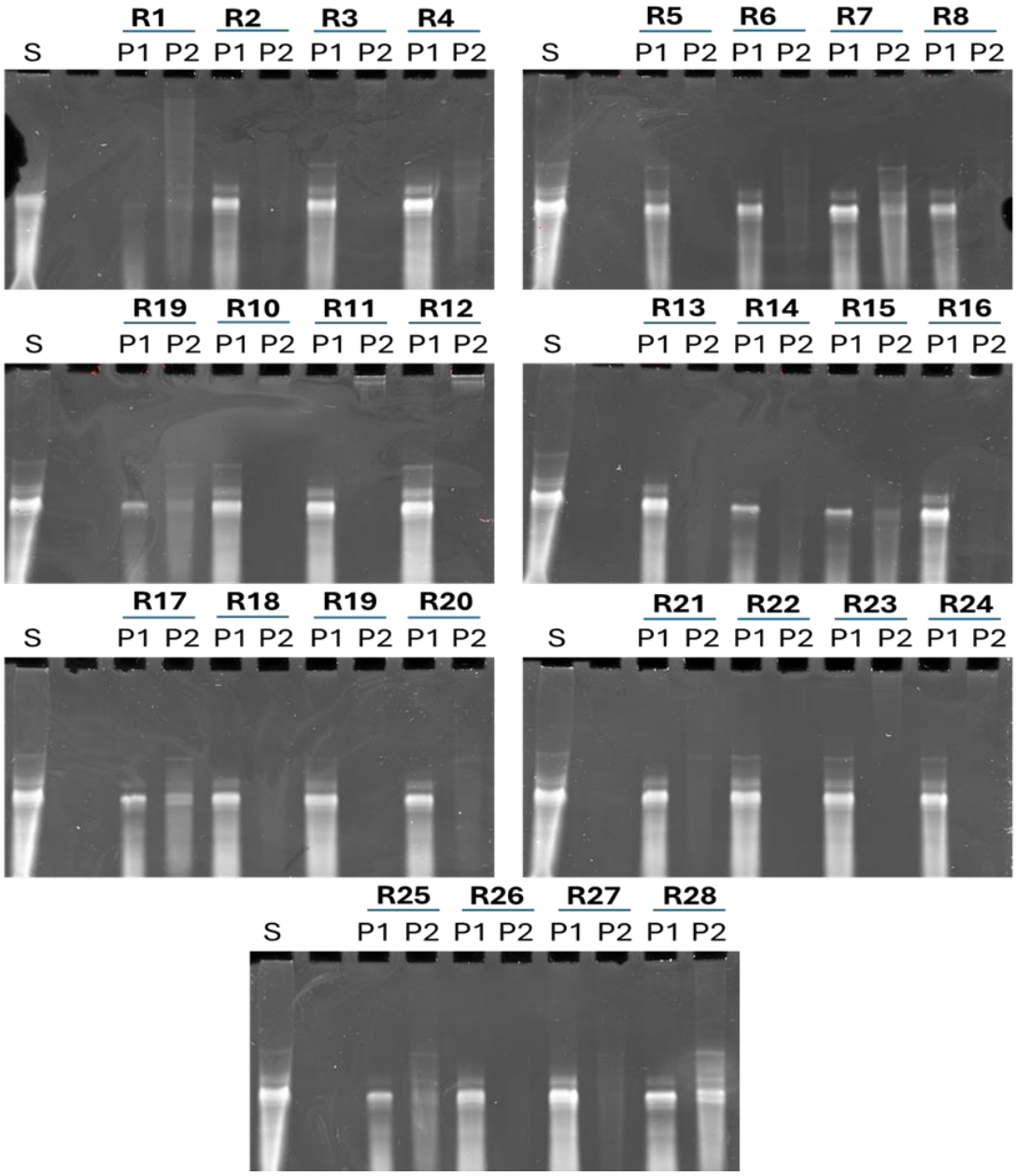

Figure 21 – Urea-PAGE of the 28 generated runs from DoE.

The limited visibility of the gels made it extremely challenging to quantify pre-miRNA-29b based on band intensity. This lack of clear resolution directly impacted the effectiveness of the experimental design, preventing it from yielding the desired outcomes and offering little assistance in further optimizing the purification process. Despite this setback, we can rely on the highly promising results from earlier optimization steps, which had already demonstrated significant improvements in the purification of pre-miRNA-29b.

Chapter 5 - Conclusions and Future Perspectives

The production and purification of pre-miRNA-29b hold significant importance, particularly due to its potential implication in AD. As a precursor to miRNA-29b, a miRNA involved in regulating various cellular processes, including those related to neurodegeneration, its role in Alzheimer's pathogenesis has drawn considerable attention. Efficient production and purification of pre-miRNA-29b are essential for studying its molecular functions and therapeutic potential. A well-optimized process is imperative to obtain high-purity samples for research, exploring how pre-miRNA-29b influences gene expression pathways linked to AD, potentially paving the way for new treatments.

In recent years, the demand for efficient and economically feasible processes to produce and purify therapeutic biomolecules has risen, driven by the need for methods that ensure that these biomolecules remain active while being free from impurities. In this work, we aimed to optimize the conditions for both the production and purification of pre-miRNA-29b, with the goal of developing an efficient and reliable protocol for producing and purifying our target molecule.

In the initial screening for pre-miRNA-29b production optimization, it was verified that at 30ºC the bacterial growth was higher. However, after analysing the quantification of pre-miRNA-29b obtained from the cultures at 25 ºC and 30 ºC, we encountered some challenges due to low concentrations and high variability. Nonetheless, we observed that at 25 ºC, there was higher production at 48 hours (0.23 ng/mL of medium), while at 30 ºC, the highest production occurred at 36 hours of culture (0.08 ng/mL of medium). Both temperatures followed a similar production pattern: initial production, followed by a peak, and then a decline in the latter hours. Despite some evidence of the higher production of pre-miRNA-29b at 25 ºC and 48 hours, the variability in the results could not assure this phenomenon. Thus, based on these results, and also supported by the literature, it was decided to proceed with a temperature of 30 ºC and 36 hours for the recovery of pre-miRNA-29b.

After completing production optimization, the aim was to purify pre-miR-29b. Using intracellular extraction samples, the column was quickly identified to have sensitivity to low NaCl concentrations and its selectivity towards pre-miRNA-29b. The equilibrium, elution conditions, and buffer pH were optimized. Our findings suggested that 0.6 M NaCl in the binding step, followed by 0.675 M NaCl in the first elution step,

1.5 M NaCl in the second elution step, and maintaining a pH of 8 would serve to obtain a significant purity of the pre-miR-29b. These conditions were also considered a solid starting point for optimizing the purification of the target RNA from the extracellular samples.

Regarding the extracellular tests, it was quickly noticed improved sample clarification on the urea-PAGE gels, though the band visualization remained a challenge at first. It was tested concentrations from previous optimizations and found that maintaining 0.6 M NaCl in the equilibrium step and 0.675 M in the elution step was ideal. Additionally, it was observed that passing samples through the molecular exclusion column was unnecessary, and concentrating the samples before injection proved to be more effective. For DoE, it was selected a NaCl range of 0.54 M to 0.63 M for the binding step, 0.63 M to 0.75 M for elution, and pH levels of 7 and 8 for the buffer. The experimental design results did not meet expectations and did not contribute to further optimization. The primary issue was the difficulty in visualizing the bands, which led to no purity quantification and consequently hindered the program's ability to generate predictions.

As a future perspective, it will be very important to develop alternative detection and quantification methods for low-concentration samples, particularly given the limitations of urea-PAGE gels in this regard. While the desired purity for the use of biopharmaceuticals has not yet been achieved, this work lays the groundwork for a purification strategy that could be coupled with additional polishing phases (SEC or AC). Future studies should focus on optimizing these strategies to enhance the purification process and improve the yield and quality of pre-miRNA-29b, thereby advancing its potential therapeutic applications. Exploring alternative detection methods will also be crucial in overcoming the current challenges and achieving the necessary purity for biopharmaceutical use.

Chapter 6 - References

1. Kesik-Brodacka, M., *Progress in biopharmaceutical development*. Biotechnology and Applied Biochemistry, 2018. **65**(3): p. 306-322.
2. Rader, R.A., *(Re)defining biopharmaceutical*. Nature Biotechnology, 2008. **26**(7): p. 743-751.
3. Craik, D.J., et al., *The Future of Peptide-based Drugs*. Chemical Biology & Drug Design, 2013. **81**(1): p. 136-147.
4. Mitragotri, S., P.A. Burke, and R. Langer, *Overcoming the challenges in administering biopharmaceuticals: formulation and delivery strategies*. Nature Reviews Drug Discovery, 2014. **13**(9): p. 655-672.
5. Khedkar, G., et al., *Nucleic acids*. 2016.
6. Wang, F., et al., *Nucleic Acids and Their Analogues for Biomedical Applications*. Biosensors, 2022. **12**(2): p. 93.
7. Zhao, Y., et al., *Nucleic Acids Analysis*. Science China Chemistry, 2021. **64**(2): p. 171-203.
8. Lodish, H.F., *Molecular cell biology*. 2008: Macmillan.
9. Cech, Thomas R. and Joan A. Steitz, *The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones*. Cell, 2014. **157**(1): p. 77-94.
10. Morris, K.V. and J.S. Mattick, *The rise of regulatory RNA*. Nature Reviews Genetics, 2014. **15**(6): p. 423-437.
11. Burnett, J.C. and J.J. Rossi, *RNA-based therapeutics: current progress and future prospects*. Chem Biol, 2012. **19**(1): p. 60-71.
12. Svoboda, P. and A.D. Cara, *Hairpin RNA: a secondary structure of primary importance*. Cellular and Molecular Life Sciences CMLS, 2006. **63**(7): p. 901-908.
13. Mattick, J.S. and I.V. Makunin, *Non-coding RNA*. Hum Mol Genet, 2006. **15 Spec No 1**: p. R17-29.
14. Pierce, B., *Genetics: A Conceptual Approach*. 2017.
15. Talap, J., et al., *Recent advances in therapeutic nucleic acids and their analytical methods*. Journal of Pharmaceutical and Biomedical Analysis, 2021. **206**: p. 114368.
16. Wang, Y., et al., *mRNA vaccine: a potential therapeutic strategy*. Molecular Cancer, 2021. **20**(1): p. 33.
17. Baptista, B., et al., *mRNA, a Revolution in Biomedicine*. Pharmaceutics, 2021. **13**(12).

18. Hombach, S. and M. Kretz, *Non-coding RNAs: Classification, Biology and Functioning*, in *Non-coding RNAs in Colorectal Cancer*, O. Slaby and G.A. Calin, Editors. 2016, Springer International Publishing: Cham. p. 3-17.
19. Baptista, B., et al., *Non-coding RNAs: Emerging from the discovery to therapeutic applications*. Biochem Pharmacol, 2021. **189**: p. 114469.
20. Nikam, R.R. and K.R. Gore, *Journey of siRNA: Clinical Developments and Targeted Delivery*. Nucleic Acid Therapeutics, 2018. **28**(4): p. 209-224.
21. Bridges, M.C., A.C. Daulagala, and A. Kourtidis, *LNCcation: lncRNA localization and function*. J Cell Biol, 2021. **220**(2).
22. Kim, K.W., *PIWI Proteins and piRNAs in the Nervous System*. Molecules and Cells, 2019. **42**(12): p. 828-835.
23. Zhang, P., et al., *Non-Coding RNAs and their Integrated Networks*. J Integr Bioinform, 2019. **16**(3).
24. Guil, S. and M. Esteller, *RNA-RNA interactions in gene regulation: the coding and noncoding players*. Trends Biochem Sci, 2015. **40**(5): p. 248-56.
25. Brandenburger, T., et al., *Noncoding RNAs in acute kidney injury*. Kidney Int, 2018. **94**(5): p. 870-881.
26. Cogswell, J.P., et al., *Identification of miRNA Changes in Alzheimer's Disease Brain and CSF Yields Putative Biomarkers and Insights into Disease Pathways*. Journal of Alzheimer's Disease, 2008. **14**: p. 27-41.
27. Mountain, A., *Gene therapy: the first decade*. Trends Biotechnol, 2000. **18**(3): p. 119-28.
28. Maslova, O., A. Koliada, and A. Vaiserman, *Gene Therapy*, in *Encyclopedia of Biomedical Gerontology*, S.I.S. Rattan, Editor. 2020, Academic Press: Oxford. p. 136-146.
29. van Haasteren, J., S.C. Hyde, and D.R. Gill, *Lessons learned from lung and liver in-vivo gene therapy: implications for the future*. Expert Opinion on Biological Therapy, 2018. **18**(9): p. 959-972.
30. Chen, C., Z. Yang, and X. Tang, *Chemical modifications of nucleic acid drugs and their delivery systems for gene-based therapy*. Med Res Rev, 2018. **38**(3): p. 829-869.
31. Scheltens, P., et al., *Alzheimer's disease*. The Lancet, 2021. **397**(10284): p. 1577-1590.
32. Angelucci, F., et al., *MicroRNAs in Alzheimer's Disease: Diagnostic Markers or Therapeutic Agents?* Frontiers in Pharmacology, 2019. **10**.
33. Jagust, W., *Imaging the evolution and pathophysiology of Alzheimer disease*. Nature Reviews Neuroscience, 2018. **19**(11): p. 687-700.

34. Hunter, S., N. Smailagic, and C. Brayne, *Aβ and the dementia syndrome: Simple versus complex perspectives*. European journal of clinical investigation, 2018. **48**(12): p. e13025.
35. Roher, A.E., et al., *APP/Aβ structural diversity and Alzheimer's disease pathogenesis*. Neurochemistry International, 2017. **110**: p. 1-13.
36. Pereira, P.A., et al., *Recombinant pre-miR-29b for Alzheimer´s disease therapeutics*. Sci Rep, 2016. **6**: p. 19946.
37. Zong, Y., et al., *miR-29c regulates BACE1 protein expression*. Brain Res, 2011. **1395**: p. 108-15.
38. Ho, P.Y. and A.-M. Yu, *Bioengineering of noncoding RNAs for research agents and therapeutics*. WIREs RNA, 2016. **7**(2): p. 186-197.
39. Pereira, P., et al., *New insights for therapeutic recombinant human miRNAs heterologous production: Rhodovolum sulfidophilum vs Escherichia coli*. Bioengineered, 2017. **8**(5): p. 670-677.
40. Caruthers, M.H., *A brief review of DNA and RNA chemical synthesis*. Biochemical Society Transactions, 2011. **39**(2): p. 575-580.
41. Ryczek, M., et al., *Overview of Methods for Large-Scale RNA Synthesis*. Applied Sciences, 2022. **12**(3): p. 1543.
42. Beckert, B. and B. Masquida, *Synthesis of RNA by In Vitro Transcription*, in *RNA: Methods and Protocols*, H. Nielsen, Editor. 2011, Humana Press: Totowa, NJ. p. 29-41.
43. Pereira, P., et al., *Advances in time course extracellular production of human pre-miR-29b from Rhodovulum sulfidophilum*. Appl Microbiol Biotechnol, 2016. **100**(8): p. 3723-34.
44. Ponchon, L. and F. Dardel, *Large scale expression and purification of recombinant RNA in Escherichia coli*. Methods, 2011. **54**(2): p. 267-273.
45. Pereira, P., et al., *Current progress on microRNAs-based therapeutics in neurodegenerative diseases*. WIREs RNA, 2017. **8**(3): p. e1409.
46. Umekage, S. and Y. Kikuchi, *In vitro and in vivo production and purification of circular RNA aptamer*. J Biotechnol, 2009. **139**(4): p. 265-72.
47. Pereira, P., et al., *Advances in time course extracellular production of human pre-miR-29b from Rhodovulum sulfidophilum*. Applied Microbiology and Biotechnology, 2016. **100**(8): p. 3723-3734.
48. Kikuchi, Y. and S. Umekage, *Extracellular nucleic acids of the marine bacterium Rhodovulum sulfidophilum and recombinant RNA production technology using bacteria*. FEMS Microbiology Letters, 2017. **365**(3).

49. Coskun, O., *Separation techniques: Chromatography*. North Clin Istanb, 2016. **3**(2): p. 156-160.
50. Scott, R.P., *Principles and practice of chromatography*. Chrom-ed book series, 2003. **1**.
51. Engel, R.G., et al., *Introduction to organic laboratory techniques: a small-scale approach*. Chemicke Listy, 2010. **104**: p. 950-953.
52. Harris, D.C., *Exploring chemical analysis*. 2012: Macmillan.
53. Ismail, B. and M. Alzeer, *Basic Principles of Chromatography*. 2010. p. 473-498.
54. Skoog, D.A., F.J. Holler, and S.R. Crouch, *Principles of instrumental analysis*. Seventh edition ed. 2018, Australia: Cengage Learning.
55. Mori, S. and H.G. Barth, *Size exclusion chromatography*. 1999: Springer Science & Business Media.
56. Fujii, S.-i., et al., *Separation and quantification of RNA molecules using size-exclusion chromatography hyphenated with inductively coupled plasma-mass spectrometry*. ELECTROPHORESIS, 2014. **35**(9): p. 1315-1318.
57. Kim, I., et al., *Rapid purification of RNAs using fast performance liquid chromatography (FPLC)*. Rna, 2007. **13**(2): p. 289-94.
58. Sousa, F., L. Passarinha, and J.A. Queiroz, *Biomedical application of plasmid DNA in gene therapy: a new challenge for chromatography*. Biotechnol Genet Eng Rev, 2010. **26**: p. 83-116.
59. Cummins, P.M., K.D. Rochfort, and B.F. O'Connor, *Ion-Exchange Chromatography: Basic Principles and Application*, in *Protein Chromatography: Methods and Protocols*, D. Walls and S.T. Loughran, Editors. 2017, Springer New York: New York, NY. p. 209-223.
60. Noll, B., et al., *Purification of Small Interfering RNA Using Nondenaturing Anion-Exchange Chromatography*. Nucleic Acid Therapeutics, 2011. **21**(6): p. 383-393.
61. Easton, L.E., Y. Shibata, and P.J. Lukavsky, *Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography*. Rna, 2010. **16**(3): p. 647-53.
62. Sousa, F., D.M.F. Prazeres, and J.A. Queiroz, *Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA*. Trends in Biotechnology, 2008. **26**(9): p. 518-525.
63. Stadler, J., R. Lemmens, and T. Nyhammar, *Plasmid DNA purification*. The Journal of Gene Medicine, 2004. **6**(S1): p. S54-S66.

64. Eriksson, K.O., *Chapter 19 - Hydrophobic Interaction Chromatography*, in *Biopharmaceutical Processing*, G. Jagschies, et al., Editors. 2018, Elsevier. p. 401-408.
65. Bonturi, N., et al., *Sodium citrate and potassium phosphate as alternative adsorption buffers in hydrophobic and aromatic thiophilic chromatographic purification of plasmid DNA from neutralized lysate*. J Chromatogr B Analyt Technol Biomed Life Sci, 2013. **919-920**: p. 67-74.
66. Magdeldin, S., *Affinity chromatography*. 2012: BoD–Books on Demand.
67. Urh, M., D. Simpson, and K. Zhao, *Affinity chromatography: general methods*. Methods Enzymol, 2009. **463**: p. 417-38.
68. Martins, R., J.A. Queiroz, and F. Sousa, *Ribonucleic acid purification*. J Chromatogr A, 2014. **1355**: p. 1-14.
69. Pereira, P., et al., *Purification of pre-miR-29 by arginine-affinity chromatography*. Journal of Chromatography B, 2014. **951-952**: p. 16-23.
70. Martins, R., et al., *A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography*. Journal of Separation Science, 2012. **35**(22): p. 3217-3226.
71. Minkner, R., et al., *Oligonucleotide separation techniques for purification and analysis: What can we learn for today's tasks?* Electrophoresis, 2022. **43**(23-24): p. 2402-2427.
72. Batey, R.T. and J.S. Kieft, *Improved native affinity purification of RNA*. Rna, 2007. **13**(8): p. 1384-1389.
73. Gilar, M., et al., *Mixed-mode chromatography for fractionation of peptides, phosphopeptides, and sialylated glycopeptides*. J Chromatogr A, 2008. **1191**(1-2): p. 162-70.
74. Neville, D.C., R.A. Dwek, and T.D. Butters, *Development of a single column method for the separation of lipid- and protein-derived oligosaccharides*. J Proteome Res, 2009. **8**(2): p. 681-7.
75. Guerrier, L., I. Flayeux, and E. Boschetti, *A dual-mode approach to the selective separation of antibodies and their fragments*. J Chromatogr B Biomed Sci Appl, 2001. **755**(1-2): p. 37-46.
76. Matos, T., J.A. Queiroz, and L. Bülow, *Plasmid DNA purification using a multimodal chromatography resin*. Journal of Molecular Recognition, 2014. **27**(4): p. 184-189.
77. Kallberg, K., H.-O. Johansson, and L. Bulow, *Multimodal chromatography: An efficient tool in downstream processing of proteins*. Biotechnology Journal, 2012. **7**(12): p. 1485-1495.

78. Montgomery, D.C., *Design and Analysis of Experiments*. 10. edition ed. 2021, Hoboken: John Wiley & Sons, Inc.
79. S, N.P., et al., *Design of experiments (DoE) in pharmaceutical development*. Drug Dev Ind Pharm, 2017. **43**(6): p. 889-901.
80. Sangshetti, J., et al., *Quality by Design Approach: Regulatory Need*. Arabian Journal of Chemistry, 2014. **101**.
81. Saltelli, A., et al., *Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index*. Computer Physics Communications, 2010. **181**(2): p. 259-270.
82. Tian, W., *A review of sensitivity analysis methods in building energy analysis*. Renewable and Sustainable Energy Reviews, 2013. **20**: p. 411-419.
83. Peng, L., et al., *Design of experiment techniques for the optimization of chromatographic analysis conditions: A review*. Electrophoresis, 2022. **43**(18-19): p. 1882-1898.
84. Dejaegher, B., A. Durand, and Y. Heyden, *Experimental Design in Method Optimization and Robustness Testing*. 2009. p. 11-74.
85. Vanaja, K. and R. Shobha Rani, *Design of experiments: concept and applications of Plackett Burman design*. Clinical research and regulatory affairs, 2007. **24**(1): p. 1-23.
86. Gunst, R.F. and R.L. Mason, *Fractional factorial design*. WIREs Computational Statistics, 2009. **1**(2): p. 234-244.
87. Ebrahimi, H., R. Leardi, and M. Jalali-Heravi, *Experimental Design in Analytical Chemistry-Part II: Applications*. Journal of AOAC International, 2014. **97**: p. 12-8.
88. Sankha, B., *Central Composite Design for Response Surface Methodology and Its Application in Pharmacy*, in *Response Surface Methodology in Engineering Science*, K. Palanikumar, Editor. 2021, IntechOpen: Rijeka. p. Ch. 5.
89. Ferreira, S.L.C., et al., *Box-Behnken design: An alternative for the optimization of analytical methods*. Analytica Chimica Acta, 2007. **597**(2): p. 179-186.
90. Hibbert, D.B., *Experimental design in chromatography: a tutorial review*. J Chromatogr B Analyt Technol Biomed Life Sci, 2012. **910**: p. 2-13.
91. Riscado, M., et al., *A new approach for extracellular RNA recovery from Rhodovulum sulfidophilum*. Analytical Biochemistry, 2025. **696**: p. 115681.
92. HIRAISHI, A. and Y. UEDA, *Intrageneric Structure of the Genus Rhodobacter: Transfer of Rhodobacter sulfidophilus and Related Marine Species to the Genus Rhodovulum gen. nov*. International Journal of Systematic and Evolutionary Microbiology, 1994. **44**(1): p. 15-23.

93. Nagao, N., et al., *Short hairpin RNAs of designed sequences can be extracellularly produced by the marine bacterium Rhodovulum sulfidophilum*. The Journal of General and Applied Microbiology, 2014. **60**(6): p. 222-226.
94. Peleg, M. and M.G. Corradini, *Microbial Growth Curves: What the Models Tell Us and What They Cannot*. Critical Reviews in Food Science and Nutrition, 2011. **51**(10): p. 917-945.
95. Njenga, R., et al., *Coping with stress: How bacteria fine-tune protein synthesis and protein transport*. J Biol Chem, 2023. **299**(9): p. 105163.
96. Matos, T., et al., *Capto™ Resins for Chromatography of DNA: A Minor Difference in Ligand Composition Greatly Influences the Separation of Guanidyl-Containing Fragments*. Chromatographia, 2016. **79**(19): p. 1277-1282.