UNIVERSIDADE
BEIRA INTERIOR

Purification of pre-miRNA-29 with ionic liquids-based supports

Versão Final Após Defesa

Sérgio Almeida Nunes

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

Orientador: Prof. Doutora Fani Pereira de Sousa
Co-orientador: Doutora Sandra Bernardo

janeiro, 2023

Declaração de Integridade

Eu, Sérgio Almeida Nunes, que abaixo assino, estudante com o número de inscrição M10884 de Biotecnologia da Faculdade Ciências, 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

Assinado por: **Sérgio Almeida Nunes**
Num. de Identificação: 15622616
Data: 2023.01.30 09:20:09+00'00'

Agradecimentos

O último ano foi a conclusão de uma aprendizagem contínua de 5 anos. Foi um ano de desafios e crescimento pessoal, pelo qual quero agradecer a todas as pessoas que me acompanharam.

Em primeiro lugar, às minhas orientadoras, Professora Fani Sousa, por me ter dado a oportunidade de integrar o grupo de investigação, e à Doutora Sandra Bernardo por me receber e acompanhar. Às duas, agradeço a orientação diária durante o ano e a partilha do enorme conhecimento científico que têm.

Quero também agradecer a todos os meus colegas de laboratório, pelas sugestões e conhecimento que me foram transmitindo, mas também pela boa disposição, o que fez com que a integração fosse bastante fácil.

Um obrigado muito especial a todos os meus amigos que me acompanharam nesta jornada, não só este ano, mas desde o início deste percurso na Covilhã. Sem os momentos que todos passámos juntos, bons e maus, teria sido mais difícil chegar até aqui.

À minha namorada, Patrícia, um obrigado enorme pelo apoio incondicional e pela paciência. Sem dúvida, uma parte fundamental da minha caminhada.

Por fim, quero agradecer à minha família pelo apoio e motivação incondicional. À minha irmã, Catarina um obrigado por estares sempre pronta a ajudar e pela preocupação de irmã mais velha, e aos meus pais, Maria da Conceição Albuquerque e Orlando Nunes, pelo apoio e compreensão. São peças fundamentais na minha vida, aos quais devo muito.

Este trabalho foi apoiado pelo projeto PTDC/BII-BBF/29496/2017 financiado pelo FEDER através do COMPETE2020-POCI e por fundos nacionais, através da FCT/MCTES. O trabalho decorreu ainda no âmbito dos projetos UIDB/00709/2020, UIDP/00709/2020, UIDB/50020/2020 e UIDP/50020/2020, e LA/P/0045/2020 (ALiCE) financiados por fundos nacionais, através da FCT/MCTES (PIDDAC).

Cofinanciado por:

COMPETE 2020 PROGRAMA OPERACIONAL COMPETITIVIDADE E INTERNACIONALIZAÇÃO

PORTUGAL 2020

UNIÃO EUROPEIA
Fundo Europeu
de Desenvolvimento Regional

FCT Fundação para a Ciência e a Tecnologia

Resumo

Os microRNAs (miRNAs) têm sido estudados relativamente à sua função de regulação da expressão génica a nível pós-transcricional, pensando na possibilidade de estabelecer novas estratégias de tratamento. O pré-miRNA-29 é a forma precursora do microRNA-29, um pequeno RNA não codificante que desempenha um papel importante no silenciamento génico, e que está envolvido em vários mecanismos celulares, nomeadamente nas doenças neurodegenerativas. Para considerar a aplicação terapêutica, são necessários processos de produção eficientes e economicamente viáveis para obter estes biofármacos, assegurando a sua integridade, pureza, estabilidade e bioatividade. Assim, neste trabalho é proposto um método de purificação utilizando suportes cromatográficos à base de Líquidos Iónicos (ILs). Os ILs são descritos como compostos iónicos que apresentam um ponto de fusão inferior a 100 ˚C, e são normalmente compostos por catiões orgânicos e assimétricos e de aniões orgânicos ou inorgânicos. Estes compostos podem ser selecionados e desenhados de acordo com as suas propriedades, tais como polaridade, densidade e viscosidade para alcançar uma maior eficácia no que diz respeito à sua aplicação na purificação e estabilização dos ácidos nucleicos. A aplicação de líquidos iónicos suportados (SILs) na purificação do RNA é altamente inovadora e pode levar ao estabelecimento de métodos de purificação vantajosos.

Neste trabalho, o SIL estudado resultou da modificação de sílica com o cloreto de 1-(3-aminopropil)imidazol, sendo denominado SilPrImPrACl. Inicialmente, foi realizado um estudo das melhores condições experimentais, em termos de pH e força iónica, para promover a ligação e eluição do RNA. Além disso, com o objetivo de desenvolver um desenho experimental, foram otimizados fatores-chave a 3 níveis, sendo eles a massa injetada de RNA, a concentração de NaCl no passo de ligação, e a concentração de NaCl no passo de eluição, utilizando o *Central Composite Design* (CCD) como modelo. A análise estatística dos resultados mostrou que este modelo não era o mais adequado para a otimização da purificação do pré-miRNA-29b, no entanto, foram efetuadas análises adicionais para compreender a influência dos fatores estudados na recuperação, fator de purificação e remoção de impurezas. Assim, foi possível estabelecer algumas condições que levaram a uma recuperação relativa de 35,1% e um fator de purificação de 1,55. Além disso, relativamente aos níveis de impurezas, verificou-se que o processo foi capaz de remover 78% das proteínas e 70% do gDNA presente na amostra inicial, provando a eficácia deste SIL para a purificação do RNA e particularmente do pré-miRNA-29b.

Palavras-chave

Pre-miRNA-29b, líquidos iónicos suportados, cromatografia, desenho experimental.

Abstract

MicroRNAs (miRNAs) have been studied regarding their biological role in gene expression regulation at post-transcriptional level, envisioning the establishment of new treatment strategies. Pre-miRNA-29 is the precursor form of microRNA-29, a small non-coding RNA that plays an important role in gene silencing, and proved to be involved in several cellular mechanisms, namely in neurodegenerative diseases. To consider the therapeutic application, efficient, and economically feasible manufacturing processes are required for obtaining these biopharmaceuticals, assuring their integrity, purity, stability, and bioactivity. Thus, in this work it is proposed an alternative purification method using Ionic Liquids (ILs)-based chromatographic supports. ILs are described as ionic compounds that present a melting point below 100 ˚C and are normally composed of large and asymmetric organic cations and organic or inorganic anions. These types of compounds can be tailored in terms of their properties, such as polarity, density, and viscosity to achieve their effectiveness regarding the purification and stabilization of nucleic acids. The application of Supported ILs (SILs) in RNA purification is highly innovative and can lead to the establishment of advantageous purification methods.

Herein, the SIL under study resulted from the functionalization of silica with the 1-(3-aminopropyl)imidazole chloride, which was named as SilPrImPrACl. First, a screening of the best experimental conditions, in terms of pH and ionic strength, to promote binding and elution of RNA was carried out. Moreover, in order to develop an Experimental Design, key factors were optimized at 3 levels, using the Central Composite Design (CCD) as model. The statistical analysis of the results showed that this model was not the best fit for the optimization of the purification of pre-miRNA-29b, however further analysis was carried out in order to understand the influence of the factors studied in the recovery, purification factor and removal of impurities. Thus, it was possible to establish some conditions conducting to a recovery and purification factor of 35.1% and 1.55, respectively. Furthermore, it was verified that the process was able to remove 78% of proteins and 70% of gDNA present in the initial sample, proving the effectiveness of this SIL for the purification of RNA and particularly the pre-miRNA-29b.

Keywords

Pre-miRNA-29b, supported ionic-liquids, chromatography, design of experiments.

Table of Contents

List of Figures

List of Tables

List of Acronyms

A	Adenine
ABS	Aqueous Biphasic System
AC	Affinity Chromatography
AEXC	Anion Exchange Chromatography
BBD	Box-Benheken
C	Cytosine
CCD	Central Composit Design
circRNA	Circular RNA
DNA	Deoxyribonucleic Acid
DVB	Divinylbenzene
E. coli	*Escherichia coli*
FFD	Full Factorial Design
FrFD	Fractional Factorial Design
G	Guanine
gDNA	Genomic DNA
HIC	Hydrophobic Interaction Chromatography
HPLC	High-Performance Liquid Chromatography
IEC	Ion Exchange Chromatography
IL	Ionic Liquid
IVT	*In Vitro* Transcription
lncRNA	Long non-coding RNA
MD	Mixture Design
miRNA	Micro RNA
mRNA	Messenger RNA
ncRNA	Non-coding RNA
PBD	Plackett-Burman Design
pDNA	Plasmid DNA
piRNA	Piwi-interacting RNA
pre-miRNA	miRNA precursor
pri-miRNA	Primary miRNA
PSILs	Polimeric Supported Ionic Liquids
RdV	*Rhodovulum sulfidophilum*
RISC	RNA-induced silencing complex
RNA	Ribonucleic Acid

RNase	Ribonuclease
rRNA	Ribosomal RNA
SEC	Size Exclusion Chromatography
SILs	Supported Ionic Liquids
siRNA	Short-interfering RNA
snoRNA	Small Nucleoar RNA
snRNA	Small Nuclear RNA
SSILs	Silica Supported Ionic Liquids
U	Uracil
VBC	Vinylbenzyl Chloride

Chapter 1 - Introduction

1.1. Nucleic Acids

Nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), are natural biopolymers of nucleotides that store, encode, transmit and express genetic information [1]. DNA is composed of a phosphate and a 2-deoxy-α-D-ribose sugar backbone and the nitrogenous bases adenine (A), cytosine (C), guanine (G), and thymine (T), while RNA has a α-D-ribose sugar backbone and another base, uracil (U), instead of T, as described in Table 1. Notice that nitrogenous bases can be classified into two classes: purines, that include adenine and guanine, and pyrimidines, which include cytosine, thymine and uracil (Figure 1). The combination of the nitrogenous base, sugar and phosphate group are called nucleotide, which is the repeating unit of the nucleic acids, linked together by phosphodiester bonds. In DNA, single stranded chains can form double helix structures when hybridizing to complementary sequences, following the Watson-Crick base-pairing rules (A:T, G:C) [2, 3].

DNA and RNA play crucial roles in diverse biological events, thus it is unsurprising that there are already about 17 nucleic acid-based biopharmaceutical products approved by Food and Drug Administration (FDA) or/and the European Medicines Agency (EMA). These approved drugs comprise a variety of DNA and RNA agents with different functions, which can either directly act on the pathogenic target gene or in a target messenger RNA. Regarding gene therapy application, therapeutic nucleic acids can be also used for gene silencing or activation therapy at the post-transcriptional level. Compared with traditional drugs that act at a protein level, nucleic acid-based biopharmaceuticals have obvious advantages of having high specificity, high efficiency, and long-term effect [4].

Table 1 – General characterization of DNA vs RNA (Adapted from [5]).

Characteristics		**DNA**	**RNA**
Type of sugar		Deoxyribose	Ribose
Presence of 2´-OH group		No	Yes
Bases	Purines	A, G	A, G
	Pyrimidines	C, T	C, U
Nucleotides joined by phosphodiester bonds		Yes	Yes
Double or single stranded		Double	Single
Secondary structure		Double helix	Variable
Stability		Stable	Easily degraded

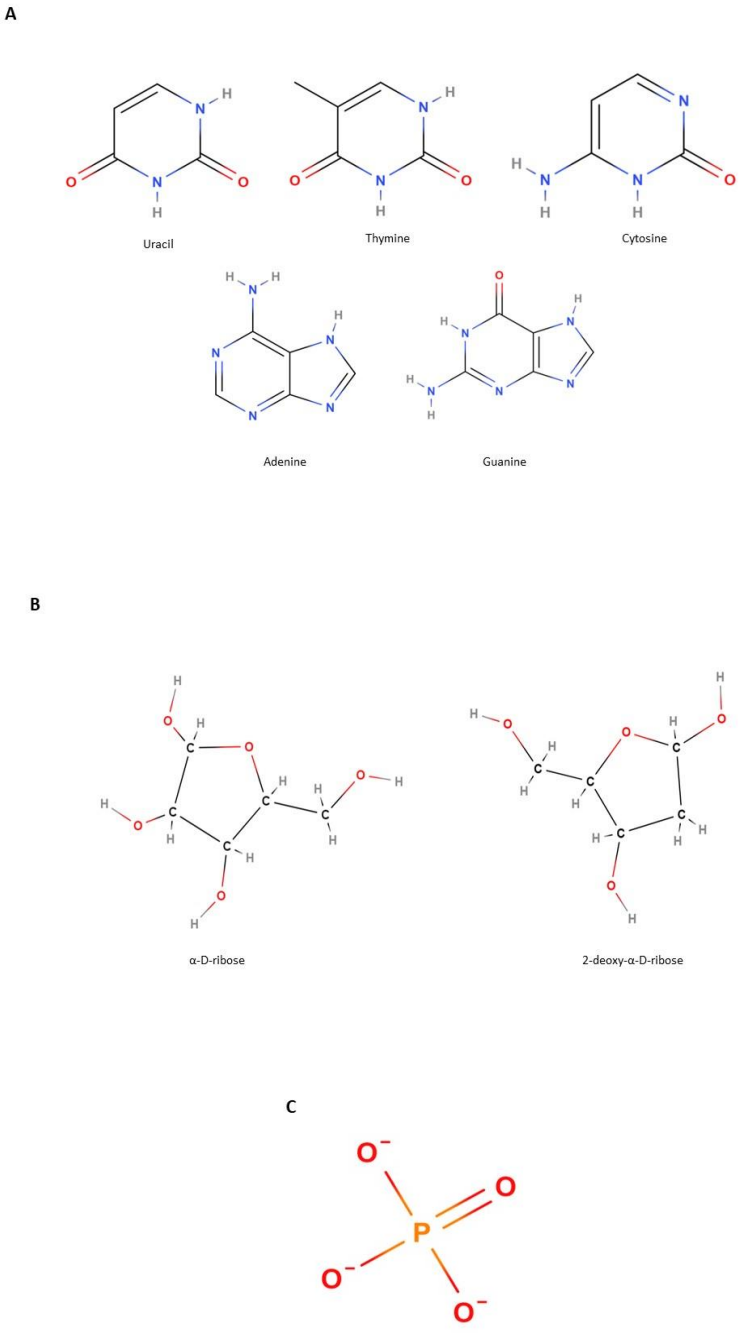


Figure 1 - The elements constituting both DNA and RNA molecules: A) The five nitrogenous bases divided in purines (A, G) and pyrimidines (U, T, C). B) The two pentoses that constitute RNA (α-D-ribose) and DNA (2-deoxy-α-D-ribose). C) Phosphate group, common to both molecules (Adapted from [2]).

1.1.1. Nucleic acids-based therapeutics

Gene therapy is regarded as a potential revolution in medicine since it is focused on treating or eliminating the causes of disease, whereas most current treatments are focused to treat the symptoms. EMA defines a gene therapy medicinal product as a product that contains an active substance that consists of a recombinant nucleic acid used in or administered to human beings to regulate, repair, replace, add, or delete a genetic sequence; or its therapeutic, prophylactic or diagnostic effect relates directly to the recombinant nucleic acid sequence it contains, or to the product of gene expression of this sequence [6, 7]. Essentially, this form of therapy involves the use of nucleic acids such as DNA, mRNA, microRNA, small interfering RNA, and also antisense oligonucleotides in the treatment, cure, or prevention of diseases by delivery of a functional therapeutic gene as a substitute for the defective or missing endogenous counterpart or by reducing the levels of a harmful defective gene product [8, 9].

Gene therapy can take place either *ex vivo* or *in vivo*. In the *ex vivo* approach, cells are removed from the patient for transfection with the therapeutic nucleic acid, and the therapeutic entity comprises the engineered cells, offering the advantages of more-efficient gene transfer and the possibility of cell propagation to generate higher cell doses. Although, it is mostly patient-specific because of cell immunogenicity and is more expensive because of the manufacturing and quality control that cell manipulation implies. On the other hand, *in vivo* gene therapy involves direct administration of the gene by transfer vectors to patients, conferring a reduced cost [6].

For a long time, gene therapy was ineffective because of immune reactions, off-target effects, developmental obstacles, and mutagenesis, but recently has emerged as one of the most significant innovations in the pharmaceutical sector. Cancer is the most studied disease reaching to clinical trials, followed by genetic, cardiovascular, and infectious diseases [10]. As gene therapy is becoming more notorious, how to effectively manufacture and administer gene products is a major challenge, which is being addressed.

1.2. RNA

The general properties of RNA molecule were already described, but it should be noted that RNA is often characterized as an unstable molecule, which is partially due to the free hydroxyl group on the 2′-carbon atom of the ribose sugar, highly susceptible to degradation under alkaline conditions. Although RNA is usually single stranded, short complementary regions can pair and form secondary structures. These RNA secondary structures are often called hairpin-loop or stem-loop structures and are determined by the base sequence of the nucleotide chain. The formation of secondary structures plays a

significant role in RNA function, and for this reason RNA molecules have the tremendous potential for alternative functions and, almost without exception, it acts throughout proteins' interactions [5]. RNA molecules are responsible for a variety of functions within the cells, such as, carrying coding instructions for polypeptide chains from DNA to a ribosome, incorporation of amino acids into polypeptide chains, processing, and assembly of rRNA, translation, gene expression regulation, among others [5, 11]. The types of RNA molecules involved in these processes can be divided into coding and non-coding RNAs (Figure 2), which will be discussed further in the next chapters [12, 13].

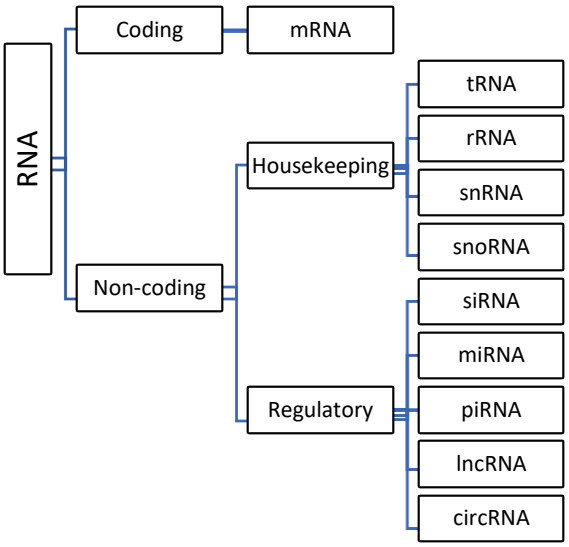


Figure 2 – General classification of RNAs (adapted from [17, 21]).

1.2.1. Coding and Non-coding RNAs

mRNA is a type of RNA that codes for proteins and can correct gene expression defects or abnormalities by exogenous introduction of mRNA, or, in case of vaccination, the mRNA expressing antigen can make the body produce antibodies and activate immune response. This molecule has an enormous potential in life science and biomedical fields as it can facilitate personalized medicines and allow patients to produce their own therapeutic proteins, surpassing some limitations associated to currently used recombinant proteins. Additionally, once the target is identified, the discovery and design of mRNA drugs are almost pragmatic. For example, mRNA vaccines allowed us to respond much quicker to the COVID-19 pandemic as their production is much faster and more flexible than the production of the traditional vaccines. Moreover, a single mRNA can express multiple proteins at the same time, which provides unique convenience for the development of multiprotein combination therapy [4, 14, 15].

Gene expression is an essential mechanism required for all biological processes, and its regulation has a certain importance since it defines development and homeostasis of all cells and tissues [16]. This regulation is highly dependent on several RNAs, known as non-coding RNAs [17] (figure 2), which can be divided into housekeeping and regulatory ncRNAs [18, 19]. Housekeeping ncRNAs are abundantly and ubiquitously expressed in cells to primarily regulating generic cellular functions, and include rRNAs, tRNAs, small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs). The tRNAs are ubiquitous nucleic acid entities that are one of the most abundant non-coding RNA molecules. They are a fundamental component of the translation machinery, responsible for the delivery of amino acids to the ribosome to translate the genetic information in an mRNA template-directed manner into a corresponding polypeptide chain [20]. Other important housekeeping ncRNA is rRNA, which is a key component of the ribosome. The snRNAs are important in RNA splicing and snoRNAs in RNA modifications [21].

Regulatory ncRNAs are usually considered as key RNA, functioning as regulators of gene expression at epigenetic, transcriptional, and post-transcriptional levels. This class of ncRNAs includes piRNA, lncRNA, circRNA, siRNA and miRNA [21, 22]. PIWI-interacting RNAs (piRNAs) are a class of small RNAs that are 24–31 nucleotides in length. They possess 2´-O-methyl modification sites at the 3´ terminus and are processed from single-stranded precursor transcripts expressed from intergenic regions termed piRNA clusters via a Dicer-independent mechanism. When associated with PIWI proteins, which constitute a germline-specific subclass of the Argonaute family, they form effector complexes known as piRNA-induced silencing complexes, which repress transposons via transcriptional or post-transcriptional mechanisms and maintain germline genome integrity [23, 24].

Small interfering RNAs (siRNAs) are 20-25 nt in length, structurally comprising two strands. Both strands have their 3´- and 5´-ends oriented oppositely with free phosphate groups at the 5´-ends and 2-nucleotide overhangs at the 3´-ends. The main function of this class of ncRNA is the cleavage of mRNA and consequently gene silencing [21, 25].

Circular RNAs (circRNAs) were discovered as a special novel type of endogenous non-coding RNAs and represent a research hotspot in the field. Unlike linear RNAs that are terminated with 5′ caps and 3′ tails, circRNAs form covalently closed loop structures with neither 5′–3′ polarities or polyadenylated tails. These are molecules with a wide range of sizes, ranging from 100 nt to over 10,000 nt. circRNAs are the circularization product from splicing events, and can be generated from exons, introns, intergenic regions, untranslated regions (UTRs) or even tRNAs [21, 26].

Long non-coding RNAs (lncRNAs) are a family that are usually longer than 200 nt, which regulate various developmental and physiological processes. For example, lncRNAs can

modulate transcription, epigenetic modifications, protein/RNA stability, translation, and post-translational modifications by interacting with DNA, RNAs and proteins [27, 28]. Micro RNAs (miRNAs) are the shorter non-coding regulatory RNAs with approximately 19-25 nt and derived from hairpin precursors. miRNAs are small, single-stranded, and endogenously encoded; their peculiarity of tissue-specific expression converges in the regulation of various biological processes in more accurate way via specific complementary binding to target mRNA, and results in either mRNA degradation (perfect binding) or translational suppression (imperfect binding). The biogenesis of miRNAs involves a complex process with multiple steps (Figure 3), starting by the transcription carried out by the RNA polymerase II (pol II). Under the catalysis of pol II, pri-miRNA is transcribed. Subsequently, pri-miRNA binds to diGeorge syndrome critical region protein 8 (DGCR8) and is then processed in nucleus to form pre-miRNA with hairpin structure, through the cleavage by Drosha. When pre-miRNA is exported to cytoplasm by Exportin-5, is further processed by DICER, resulting the miRNA duplex. miRNAs assemble with the mammalian Argonaute (Ago) family of proteins into an effector complex known as the RNA-induced silencing complex (RISC). By interacting with Ago proteins, one strand of the resulting duplex is associated with the RISC that is a nuclease complex. The miRNA-RISC complexes target specific mRNA through perfect or imperfect complementary binding to the 3'untranslated regions (3'-UTRs), leading to degradation of mRNA or translation repression. Furthermore, it is known that miRNAs can regulate a high number of target mRNAs and, conversely, a given mRNA can contain target sites for many miRNAs [12, 29, 30].

These classes of ncRNAs can be potential biomarkers and may be directly involved in the pathogenesis and progression of various diseases, like cancer and Alzheimer's disease. For that reason, the recognition of the importance of functional ncRNAs in physiological and pathological conditions, especially the discovery of miRNA and siRNA-mediated post-transcriptional gene regulation and RNA interference mechanisms, has revolutionized life sciences and biomedical research [19, 29-33].

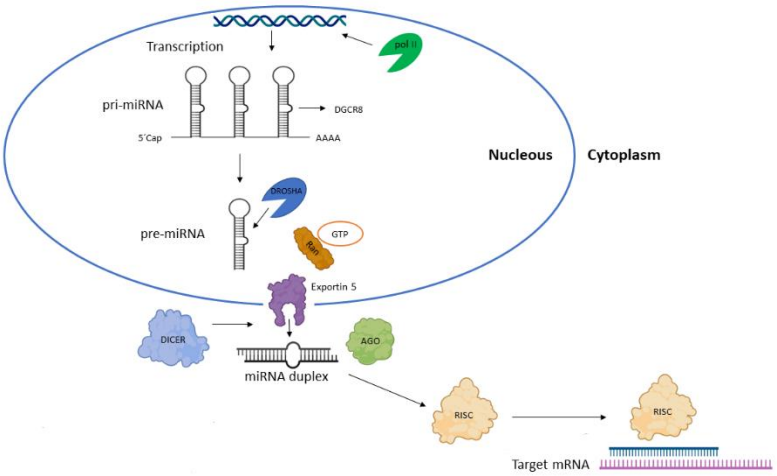


Figure 3 - Biogenesis and mechanism of action of miRNA (adapted from [29]).

1.3. Small RNA production

1.3.1. Chemical Synthesis

This method is ideal when it comes to produce oligonucleotides shorter than 10 nucleotides, as these cannot be efficiently produced by the T7 system, but is suitable for relativity short oligomers, up to 80 nt. The synthesis involves four main steps, first the 5´-hydroxyl protecting group of the nucleotide that is bound to the solid support is removed to allow the 5´-hydroxyl to be attacked by the activated 3´-hydroxyl group of the incoming nucleoside phosphoramidite monomer in the second step or coupling step. In the third step, unreacted 5´-hydroxyl groups are blocked from participating in subsequent reactions to avoid synthesis of side products. The fourth, and last step, the unstable phosphite triester formed during the coupling step is converted to a stable species by an oxidation step, and the whole cycle can be repeated to obtain the desired length of the oligonucleotide polymer [34, 35].

1.3.2. *In vitro* transcription (IVT)

In vitro transcription is the most common method used for synthesizing RNA molecules. It utilizes a DNA template with a promoter sequence for T7 RNA polymerase, which is the most extensively studied and widely used, followed by a sequence encoding the target molecule. The enzyme binds the template at the promoter region and starts the RNA synthesis. The elongation of the RNA chain is finalized when the enzyme drops off 3´-end of the DNA template. *In vitro* synthesis can be efficient, resulting in milligrams of

RNA sample, but a high yield depends on several factors, for example, the type of nucleotides located immediately downstream of the T7 promoter. The transcript should start from at least one guanosine residue, but having two or three consecutive Gs is better. The RNA sequence can be adjusted, by simply introducing the respective changes into the DNA template [36, 37].

1.3.3. Recombinant RNA

The chemical and enzymatic or *in vitro* methods mentioned before are the most frequently used to produce RNAs efficiently. However, some limitations must be overcome, like the constrains regarding the large-scale production of RNA or complex purification procedures. Actually, several purification protocols to remove the contaminants, such as impurities of plasmid DNA template, enzymes, nucleotides, chemicals, salts, or buffers (common reagents used in chemical and enzymatic synthesis) must be employed. The presence of these impurities can lead to non-targeted gene silencing, which is associated with a decrease in therapeutic effectiveness and restricts the implementation of these RNAs in pre-clinical or clinical trials [38, 39]. To overcome this problem, recombinant expression of RNA in different microorganisms have emerged. The expression of RNA, *in vivo*, follows the principle that a recombinant plasmid encoding the target RNA is introduced into a host cell, which is then grown in appropriate conditions. The host transcription machinery will synthesize the RNA of interest, which will accumulate in the cytosol. At the end of the process the cells are pelleted, lysed and the RNA is purified by standard chromatographic techniques. Recombinant RNA expression has been mostly achieved using *Escherichia coli* as host, once it can be grown easily and economically and many plasmids and strains are available, but the possibility of using different expression hosts has already started to be explored. Thus, lately, some strategies emerged to produce stable RNAs in *E. coli*, for example, circular RNA aptamer [40], a transfer RNA scaffold [41], and a human pre-miRNA-27b [42]. The production in *E. coli* is quick and efficient, and a large number of plasmids is available, however, it is necessary to perform cell lysis and extraction, which is time-consuming because it involves a certain number of complex steps. Another step down is the possible release of endotoxins, which can lead to the contamination of the target biomolecule and toxicity *in vivo*. The final consideration is that this bacteria expresses several endonucleases, which make difficult to maintain the integrity of the target RNA [33, 39]. For these reasons, other bacteria were also described as alternative hosts for production of RNAs.

1.3.4. RNA biosynthesis in *Rhodovulum sulfidophilum*

In our research group, the use of *Rhodovulum sulfidophilum* to produce recombinant RNA has been compared with *E. coli*. The study revealed that even being necessary several purification steps to obtain RNA from intracellular fractions of *E. coli,* it produced higher quantity of RNA in shorter culture time, while the use of *R. sulfidophilum* as an expression host could drastically simplify the downstream processes due to RNA secretion ability [30, 35, 41]. The marine bacteria *R. sulfidophilum* is a Gram-negative bacteria and cells divide by binary fission under anaerobic conditions in the presence of light or under aerobic conditions in the dark [43, 44]. This bacteria is non-pathogenic and presents several advantages over *E. coli*, such as the capacity of secretion of nucleic acids directly to the culture medium and the absence of detectable host RNases in the extracellular medium [44, 45]. The research group of Kikuchi was the first to describe the production of an RNA aptamer in *R. sulfidophilum*, and later, our research group was able to produce human pre-miRNA-29b in the same host [39]. The study was carried out by growing the transformed cells in dark-aerobic conditions, and the production was performed during 72h, which is a longer culture time than the established for *E. coli* (8h). The RNA levels were monitored during fermentation and the results showed that the maximum level of extracellular pre-miRNA-29b was 182 μg/L at 40h and the intracellular quantity was about 358 μg/L at 32h, which is an excellent result since the maximum level of intracellular production matches the exponential phase of extracellular production. On the other hand, results suggest that after intracellular pre-miRNA-29b production, this RNA is targeted to *R. sulfidophilum* secretory pathway, thus explaining why the peak for the extracellular pre-miR-29b accumulation is attained after the maximum levels of intracellular pre-miR29b, and at 40 h of fermentation, the intracellular levels of pre-miRNA-29b are low, enhancing the fact that the majority of the intracellular production in the first 32 h of fermentation is processed by *R. sulfidophilum* secretory machinery. Shortly, this method allowed the intra and extracellular production of human pre-miRNA-29b with high yields, proving to be an interesting substitute to the usual methods. Moreover, this research demonstrate that this biomolecule can be produced by recombinant technology which allows further studies in the production of natural pre-miRNA molecules in the biomedical field [39].

1.4. RNA purification by chromatography

Liquid Chromatography is based on the principle that molecules present in a mixture can be applied onto a stationary phase, and a mobile phase promotes the flow and separation of the different molecules, depending on the interaction with both phases [46]. Molecules with a high interaction with stationary phase will move through the system at a lower

velocity than those that interact preferentially with the mobile phase. The most conventional physical configuration is column chromatography, in which the stationary phase is packed or molded into a column, through which a mobile phase is pumped. The sample to be separated is introduced into one end of the column and the various components of the sample travel at different velocities through the column and are subsequently detected and collected at the other end. There are several types of liquid chromatography that are used, mainly differing in the type of stationary phase. For instance, ion-exchange chromatography (IEC) is based in the net charge of the molecule, hydrophobic interaction chromatography (HIC) in hydrophobicity, affinity chromatography (AC) in the biological function and specific interactions and size exclusion chromatography (SEC) relates with the size and shape of the biomolecule for separation [47].

1.4.1. Size exclusion chromatography (SEC)

The basic principle of this method is to use porous materials to separate macromolecules based on their differences in molecular sizes. In this chromatography technique, the matrix is constituted of inert molecules with defined size pores. The sample solution, containing molecules of different dimensions, is passed continuously with a constant flow rate through the column. Molecules larger than pores cannot permeate into gel particles, and they are not retained in the matrix. Thus, these larger molecules pass through spaces between porous particles and move rapidly through the column. On the other hand, the molecules smaller than the pores are diffused slowly into pores, and as molecules get smaller, they leave the column with proportionally longer retention times. Dextran, agarose, and polyacrylamide are some examples of materials used in SEC [46]. SEC has also accompanied the advances in RNA purification. With this technique it is possible to separate the components of a mixture of RNAs based on their molecular size and is the simplest type of chromatography for the purification of oligonucleotides. SEC has contributed as a polishing step in removing salts from short oligoribonucleotides by HPLC. This additional desalting step prevents cytotoxic effects from trace by-products of synthesis or trace solvents which may carry over from purification. However, SEC-based methods still require several time-consuming preparatory steps, such as phenol/chloroform extractions to remove proteins followed by desalting and sample concentration [38]. Puglisi and co-workers developed SEC-based purification schemes using fast performance liquid chromatography (FPLC) systems that allowed the efficient elimination of unreacted nucleotides, enzymes, short abortive transcripts, and the high molecular weight pDNA template from the desired RNA product. These approaches were performed under non-denaturing conditions, which allows the exclusive separation of

monomeric RNA from oligomerized RNA, and avoided harsh precipitation steps that may cause RNA aggregation and degradation [48]. In addition, this technique has been explored as an alternative method to preparative denaturing PAGE in purifying homogeneous-length RNA obtained by *in vitro* transcription [49].

1.4.2. Ion-exchange chromatography (IEC)

IEC is a technique based on electrostatic interactions between charged groups and an oppositely charged solid chromatographic material. Such interactions are reversible by changing the pH or ionic strength of the mobile phase. Positively charged matrices adsorb negatively charged compounds, being named anion-exchange matrices. On the other hand, matrices with negatively charged groups are known as cation-exchange matrices and adsorb positively charged molecules [46, 50]. Both cation and anion exchangers can be classified as either weak or strong exchangers. Weak cation exchangers are comprised of weak acid that gradually loses its charge as the pH decreases (e.g. carboxymethyl groups), while strong cation exchangers are comprised of a strong acid that is able to sustain its charge over a wide pH range (e.g. sulfopropyl groups) [50]. Anion-exchange chromatography (AEXC) is an extremely popular technique, especially in plasmid purification. It takes advantage of the polyanionic nature of DNA, due to the presence of phosphate groups on the nucleic acid backbone and is therefore, conveniently captured on a resin derivatized with positively charged functional groups [51]. Furthermore, anion exchange matrices are also described to have some inherent hydrophobicity which favors oligonucleotides separation. The AEXC selectivity for oligonucleotides was found to be based on molecular size and sequence and the success of separating double and single-stranded nucleic acids relies on the distinct hydrophobic behaviors of the G-C, and A-U base pairs. Additionally, it has the advantage of displaying simplicity on the operations, since it is performed in mild aqueous conditions without the use of hazard and high cost eluents, at low to moderate operating pressures [38].

1.4.3. Hydrophobic interaction chromatography (HIC)

Hydrophobic interaction chromatography (HIC) is a key bioseparation technique at laboratory, preparative and industrial scales. HIC takes advantage of biomolecular hydrophobicity and promotes separation based on hydrophobic interactions between immobilized hydrophobic ligands and non-polar regions of biomolecules. In an aqueous rich environment, retention of polar biomolecules is promoted at high salt concentrations, driven by the displacement of ordered water molecules around the biomolecules and the ligands, which then increases the entropy of the environment,

attracting the non-polar groups onto the stationary phase. In terms of elution, it occurs by decreasing the salt concentration of the mobile phase which weakens the hydrophobic interactions. Regarding nucleic acids, HIC takes advantage of the higher hydrophobicity of single stranded nucleic acids that show a high exposure of the hydrophobic aromatic bases when compared with double stranded nucleic acids. In this case, the hydrophobic bases are packed and shielded inside the helix and thus, interaction with the support is greatly reduced. On the other hand, the high content of single strands in RNA and gDNA impurities enables hydrophobic interactions to take place [52, 53]. This technique is often used because of the minimal requirement of solvents and the structural damage to the biomolecules as their activity is maintained, despite the complex mechanism involved in the interaction. However, the requirement of high concentrations of salt, is viewed as a disadvantage due to the associated costs and environmental impact, especially regarding the industrial application of this method. The temperature, pH, and salts can have a significant impact on HIC retention and selectivity, but by optimizing these conditions, separation can be efficient [50].

1.4.4. Reverse-phase chromatography (RPC)

Reversed-phase chromatography (RPC) is usually based on hydrophobic interactions between the non-polar solute and a non-polar stationary phase. Elution of solutes occurs in decreasing order of polarity or increasing hydrophobicity [54]. However, when the target molecules are polar, which is the case of nucleic acids, reverse-phase principles are achieved by adding amphiphilic organic ions to the mobile phase. These ions establish ionic interactions with the target molecules, forming hydrophobic non-polar ion pairs that bind to reverse-phase resins. RPC can be used to purify supercoiled plasmid DNA from crude cell lysates and from anion exchange chromatography eluates [55]. Low molecular weight RNA and linear and open circular plasmid forms are usually well separated from the late eluting sc plasmid DNA. Endotoxins and gDNA remain bound to the column and are removed after sanitizing the column with NaOH. The need to use organic solvents and mixtures to elute plasmid molecules in RPC constitutes a significant disadvantage [53, 55].

1.4.5. Affinity chromatography (AC)

Affinity chromatography (AC) is used for the purification of enzymes, hormones, antibodies, nucleic acids, and specific proteins [46]. Although the early definition was related to the interactions like that occurring in many biological systems, such as the binding of an enzyme with a substrate or of an antibody with an antigen, the meaning of

the affinity concept, in biomolecule separation context, has undergone evolutionary changes over the years, especially to answer to the challenges of purifying new biomolecules with clinical and therapeutic interest [51]. Since its discovery, affinity chromatography has grown to include a wide variety of ligands for analytical and preparative applications, however the biological origin of some of these ligands is viewed as a limitation, since these ligands tend to be fragile and associated with low binding capacity. For this reason, the design of synthetic ligands, which would combine the selectivity of natural ligands with high capacity and durability of synthetic systems, is an emerging area for improving affinity chromatography. The specific interactions occurring between ligand and target molecules can be the result of either electrostatic and/or hydrophobic interactions, van der Waals forces and hydrogen bonding. Because of this diversity of possible interactions, the elution step can be performed specifically, using a competitive ligand, or non-specifically, by changing the pH, ionic strength, or polarity, depending on the matrix and the chemical characteristics of biomolecules [51]. AC has been applied to RNA purification processes, based on the natural biorecognition occurring between RNA and amino acids. Histidine and arginine have been used as affinity amino acid ligands, and their ability to isolate different RNA species demonstrated multipurpose applicability in molecular biology analysis and RNA therapeutics preparation, highlighting the potential contribution of these methods to overcome the challenges of RNA purification [56]. However, one of the most used affinity chromatography strategies is the oligo(dT) for mRNA purification, based on the nucleotide base pairing specificity of adenine with thymine. This type of AC, allows mRNA purification directly from a biological sample or from previously isolated total RNA. Additionally, it is a simple and reliable method, and is, therefore, a frequent choice over other methods [38]. The specific nature of the underlying biomolecular interactions is a major advantage of affinity chromatography since it results in a high selectivity and high resolution. In general, in a single step, affinity chromatography can offer immense advantages over other less selective and time-consuming multi-step procedures [38, 56, 57].

1.5. Supported Ionic liquids (SILs)

1.5.1. Brief introduction of ILs

Ionic liquids (ILs) are organic salts composed of large and asymmetric organic cations and organic or inorganic anions, presenting a melting temperature below 100 °C. It is estimated that there are more than a thousand of possible formulations of ILs, in which hundreds of those are chemically and physically well-known and characterized. Due to their ionic nature, ILs present two outstanding properties: negligible volatility and non-flammability. Because of these two characteristics, ILs have received a "green" connotation. Besides these characteristics, there are other properties such as biodegradability and biocompatibility that should be additionally considered when ILs are used. Ionic liquids are employed as advanced fluids in separation, extraction and purification techniques for the isolation and analysis of various products [58, 59], and this enormous potential is often due to structural diversity and adjustable properties. Application of ionic liquids as an environmentally friendly method in extraction processes has been proposed, taking into consideration the importance of clean and green manufacturing. Ventura and co-workers have extensively reviewed the ionic liquid-based extraction processes of different bioactive compounds, including biomass active compounds, amino acids, fats, essential oils, vitamins, saponins, carotenoids and more complex molecules such as nucleic acids, proteins, enzymes, and antibodies [60].

ILs can be also used as forming phase component of aqueous biphasic systems (ABS). ABS are formed by the combination of two water-soluble compounds, such as two polymers, a polymer and a salt/IL, or two salts/ILs in aqueous media, and where above a given concentration, they become immiscible with the formation of two-phases [61]. ABS extraction is a more compatible and sustainable method compared with traditional liquid-liquid extraction in which volatile and hazardous organic solvents are employed. In IL-based ABS, hydrophobic or hydrophilic ionic liquids are used with another phase forming component for phase formation. Additionally, this method provides a technically simple, energy-efficient, easily scalable, mild, and non-toxic separation method for product recovery. There is a large availability of cations and anions that can be used to synthetize the ILs with the desired properties (e.g., unlike polymers and salts, the hydrophobicity of ILs can be tailored). ABS composed of carbohydrates and ILs constitute the most environmentally friendly class of ABS. The two-phase separation depends on the solubility of carbohydrate and do not require inorganic salts. In this system, the hydrophilic IL salt-out the less water-soluble carbohydrates and remains present as the lower phase.

IL-based ABS have been mostly studied for the extraction of various biomolecules including amino acids, peptides, DNA, enzymes, and proteins. Figure 4 shows the

partition of a biomolecule, such as RNA for instance, in two phases: an IL rich phase and a salt rich phase [58, 62-64].

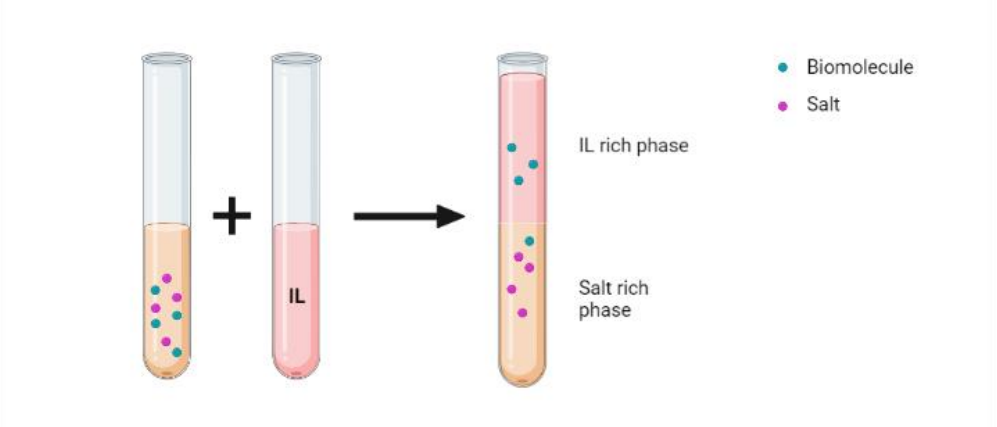


Figure 4 - Schematic representation of ionic liquid based aqueous biphasic system.

1.5.2. Biocompatibility of ILs

As mentioned above, biocompatibility and toxicity should be considered when it comes to biological applications, since the unstable nature of biomolecules makes them vulnerable to denaturation that causes subsequent loss in bioactivity. Thus, the toxicity of ILs have been investigated prior their use during the manufacturing and storage of biopharmaceuticals, such as peptides, cytokines, antibodies, vaccines and nucleic acids. It has been found that nontoxic ILs can be prepared by selecting biocompatible organic cations and inorganic anions [65]. It is widely accepted that the head group of the cation has a deciding role in the toxicity and longer side chains have the most negative effect on cells. Therefore, toxicity can be decreased by either reducing the hydrophobicity of the side chain by shortening long alkyl chains, preferentially four carbons or less, or by introducing polar functional groups into the IL cations. In addition to the strong side-chain effect, the chemical structure of the cationic group also has a significant effect on the toxicity. Usually, imidazolium, pyridinium and quinolinium head groups have shown stronger toxicological effects than morpholinium, which is the head group that has been found to be the least toxic. In the other hand, the toxicological effects of anions are not very significant, however hydrophobic, and most fluorinated species are quite appropriate for nontoxic ILs. Despite the discoveries made recently, it is important that the biological compatibility of ILs should be defined in the context of their specific target applications [65, 66]. Being this role particularly relevant for the stabilization of biopharmaceuticals. The stabilization of three-dimensional structures of biomolecules during manufacturing and storage is the priority for biopharmaceutical engineering since the structures determine their functions and biological activities [67]. ILs have

been proven to significantly enhance the thermal stability of proteins [68, 69], protect their three-dimensional structure [70], inhibit protein aggregation or unfolding [71], prevent proteins from contacting with denaturants [72] and maintain the activity of enzymes [73].

Additionally, ILs have been used as constituents of liquid mobile phases. There has been studies demonstrating that the use of ILs in separation techniques, particularly in HPLC as additives of mobile phases, with the goal of surpassing the negative effect of free silanol groups on the long retention time of analytes, can lead to a decrease in the number of organic solvents and additives used, as well as to decrease the energy consumption by increasing the speed of analysis without compromising the analytical performance or even improving it [74, 75].

1.5.3. SILs as Chromatographic Matrices

ILs have been also used as stationary phases, bound to silica for instance, and are named as supported ionic liquids (SILs). In 1985, Moreira and Gushikem used silica gel functionalized with 3(1-imidazolyl)propyl groups to adsorb and pre-concentrate metal ions from ethanol solutions, having been the first use of ILs combined with silica [76]. ILs were first applied as stationary phases in gas liquid chromatography and later in HPLC and Thin Layer Chromatography (TLC). The incorporation of ILs in stationary phases allows specific interactions with target compounds, which results in high extraction efficiencies and selectivity while reducing the amount of hazardous organic solvents traditionally used [59, 74, 77].

In general, stationary phases play a crucial role in separation science because they are responsible for retention and elution of analytes. There are characterized by several factors, for example, silanol activity, hydrophobicity, ion exchange capacity, presence of metal impurities and steric properties. Lately, there has been a demand for producing new stationary phases for liquid chromatography, to improve column efficiency, permeability, and stability. SILs are among the most important branches in the IL field. SILs can include both the immobilization of ILs onto silica materials and the attachment of ILs onto polymer matrices. Moreover, SILs can be immobilized in different ways, requiring interactions between the IL (anion and/or cation) and the support material or functional group, and, although it can be argued that ILs once bound to a solid support, no longer constitute a true IL, the unique properties of ILs still remain in SILs. Initially, most imidazolium-based ILs stationary phases were prepared via coating IL layers on the support surfaces without chemical bonds. However, the driving of a mobile phase can disarrange IL films when they flow through physically adsorbed ionic liquid membranes,

and then render them unstable. The concept of SILs covalently linked onto a stationary phase has propelled an advance in separation technologies. Such SILs have been used in solid-phase extraction (SPE), liquid chromatography (LC), gas chromatography (GC) and capillary electrochromatography (CEC), gas capture, and mass spectrometry [78, 79].

In theory, there are five ways to immobilize ILs in stationary phases: with only one cation; one anion; zwitterionic, and co-immobilization of ions, as shown in the figure below [80].

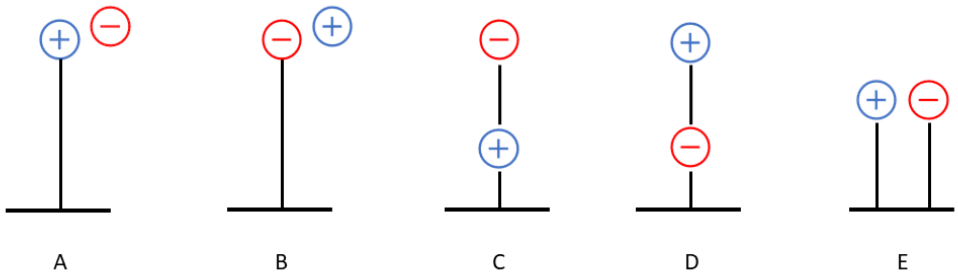


Figure 5 - Immobilized ILs in stationary phases: (A) and (B) Ionic liquids immobilized via their cations or anions; (C) and (D) zwitterionic immobilization; (E) co-immobilization of ions [78, 81].

Usually, ILs are immobilized onto supports via their cation or anion and the respective counterion is free, as shown in figure 5 A and B. Their counterions are liable to suffer an exchange when a mobile phase includes ion species, which can be an advantage because of the ability to change the surface properties but may be a disadvantage when it comes to the stability of the column during an assay [80, 81]. To avoid this issue of lose the counterion, zwitterionic immobilization could be a solution, given that in this case, both cations and anions are linked by covalent bonds, figure 5 C and D. Another way is to co-immobilize both anion and cation, figure 5 E, which can enhance stability during the use of different buffers in the mobile phase. This copolymerized ionic liquid phase facilitates multiple interactions with the analytes, originating from the multifunctional groups, including long alkyl chain, imidazolium, and benzene sulphonate moieties [82].

1.5.3.1. Silica Supported Ionic Liquids (SSILs)

Silica is one of the most used stationary phases in HPLC chromatography due to the large surface area, and high thermal and mechanical stability. A great advantage of using silica is that a wide range of selectivity can be achieved since it can be easily modified by the immobilization of different compounds to silanol groups. In contrast, one important disadvantage of silica-based materials occurs upon the analysis and separation of basic compounds. In this case, strong interactions can be established between these compounds and the free residual silanol groups at the surface of these supports, limiting the effective separation. This shows how crucial it is to perform the appropriate modification of the surface of the stationary phase. Hence, it is highly important to know all the possible chemical routes and identify the best options for the immobilization of a certain ligand. One of the most attractive properties of SSILs is their potential for multi-modal retention properties, being capable to interact with analytes through different mechanisms, such as hydrophobic, electrostatic, and hydrogen bonding, depending on the IL nature [75, 78].

Imidazolium IL-based stationary phases display good retention behavior due to the potential to interact with solutes through multiple interactions. Qiu and coworkers [83] were one of the first groups to show the application of N-methylimidazolium chloride in a silica stationary phase. It was used for the separation of a mixture of common inorganic anions by anion-exchange chromatography, and it was concluded that the stationary phase presented more than anion-exchange characteristics, for instance, it allows the establishment of reverse phase interactions, showing multi-modal behavior. Later, Yang and coworkers [84] studied the differences between 1-methylimidazolium chloride and 2-methylimidazolium chloride functionalized silicas for mixed-mode chromatography application. Nucleosides, nucleobases, water-soluble vitamins, sulfonamides, and saccharides were separated by hydrophilic chromatography, while inorganic anions were separated by anion exchange chromatography. For both chromatographic methods, the SSILs matrices presented good performance. The use of pure alkyl stationary phases has some disadvantages, such as low compatibility with highly aqueous eluent and insufficient selectivity towards more polar solutes. Moreover, C18 and C8 continue to be the most used support phases in reverse-phase liquid chromatography. Furthermore, Zhang and co-workers [85], synthesized novel stationary phases based on modified alkylimidazoliums to form polar-embedded phases, which have lower hydrophobicity compared to C18 and are notable for their stability in aqueous media, improved performance in separation of polar compounds, and higher selectivity.

1.5.3.2. Polymeric Supported Ionic Liquids (PSILs)

Regarding the polymeric supported ILs (PSILs), imidazole and pyridine groups have been the most used ILs in the functionalization process [77]. The first work in this field was published in 2009 by Fontanals and coworkers [86], where N-methylimidazolium trifluoroacetate was functionalized onto a VBC-DVB (vinylbenzyl chloride-divinylbenzene) co-polymer. This novel IL-polymer was used as an SPE (Solid-Phase Extraction) sorbent for the extraction of ten pharmaceutical products from water samples, which are emerging pollutants in the environmental field. Such polymer-based material have been studied increasingly, however most of them are still focusing on small molecules, metals separation, and gas separations for analytical methods [87]. Currently, it is yet essential to design novel PSIL-based supports for the separation and purification of high value biomolecules, such as proteins and nucleic acids [78]. Alves and co-workers [88] synthetized a new PSIL from Sepharose CL-6B containing a 3-(10-carboxydecyl)-2-methylbenzothiazol-3-ium bromide as a ligand for protein purification. The main goal of this work was to evaluate the individual contribution of the azolium functional group in an affinity or multi-modal relationship with standard proteins. This important work has shown the potential application of multi-modal IL-based ligands for the separation and purification of proteins. For the separation of three types of nucleic acids, namely ribosomal RNA, small RNA, and genomic DNA from a complex bacterial lysate in just one step, our research group used a macroporous-based chromatographic support (Toyopearl® AF-Epoxy-650M), and performed its functionalization with 1-methyl-3-propylimidazolium chloride, since this IL presented a multi-modal behavior [89].

1.6. Experimental design in chromatography

1.6.1. Quality by Design

Quality by design (QbD) is defined as a systematic approach that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management [90].

Quality has been given importance by all regulatory agencies of pharmaceutical products and it means customer satisfaction in terms of service, product, and process. Features like performance, trustworthiness, robustness, user-friendly, and service ability must be built into the product and such product should be free from imperfections (e.g. impurities) [91]. The principles of QbD have been used to advance product and process quality in every industry. Because of the need for effective pharmaceutical drugs with safety profiles, industries are currently investing a large quantity of their capital for the

drug discovery and process development, in order to design quality products, with high consistency in their manufacturing process to deliver the proposed products [92].

1.6.2. Experimental Design

Experimental design is a structured, organized method for determining the relationships between factors capable of directly affecting the process and the output of that process. Moreover, this is a method for achieving process knowledge through the establishment of mathematical relationships between process inputs and outputs [90]. It is used to optimize several operating conditions of various processes and improve chromatographic separation performance. Ideally, several factors have a simultaneous effect on a process. Nevertheless, identification and optimization of significant factors as a function of experimental design is most effective to achieve a competent result by fewer experimental trials, so one or several predetermined factors are deliberately manipulated to perceive their influence on the experimental outcome. Based on the experimental goal, the designs can be classified into two categories: screening designs or optimization designs, also known as response surface designs [90, 93, 94].

The most used screening designs are Plackett-Burman design (PDB) and Fractional Factorial design (FrFD), whereas Full Factorial design (FFD), Box-Behnken design (BBD), Central Composite design (CCD), and Mixture design (MD) are some examples of commonly used optimization designs [95-98]. In table 2, are listed some of the characteristics of this types of designs.

Table 2 - Applications, advantages and disadvantages of various screening and optimization models of experimental design (FFF: Full Factorial Design; FrFD: Fractional Factorial Design; PBD: Plackett-Burman Design; CCD: Central Composite Design; BBD: Box-Behnken Design; MD: Mixture Design).

Types	Application conditions	Advantages	Disadvantages	Ref.
FFD	Suitable for a small number of factors	Can estimate all effects and the interactions of each factor	Test numbers will dramatically increase with the growing of factors	[96]
FrFD	Interactions between factors can be negligible, or the factors are too high	Needs a small number of experiments	Only considers a small number of main effects and lower order interactions	[97]

PBD	Screening and robustness test	Needs a very small number of experiments	Only studies the main effects of factors	[91]
CCD	Optimization	Determines the optimal levels of factors by including central point	Considers extreme conditions	[104, 105]
BBD	Optimization	Alternative to CCD	Not sequential	[104, 105]
MD	The sum of the proportions of all the components of a mixture must always be equal to 100%	Can optimize the formulation components	Factors cannot change independently of each other	[98]

1.6.3. Factorial designs

This category includes FFD and FrFD. In FFD, all combinations at all levels of all factors are tested at least one time, because of that, exists the advantage of being capable of estimating all the main effects and the interactions of each factor. On the other hand, the total number of experiments will increase with the increased number of factors, so it is more suitable for the cases including few factors [93]. EL-Shorbagy research group used a two-level FFD to optimize a high-performance liquid chromatography method for the determination of two drugs used for the treatment of hepatitis C, sofosbuvir and ledipasvir [99]. In this case, the design was suitable due to the low number of factors. The other factorial design is the FrFD, which only requires a fraction of all combinations, so that it can be used when the number of factors is larger [100]. Its application is often associated to the implementation of a following optimization design, like CCD or BBD, to assure that factors being screened have a real significant contribution to these

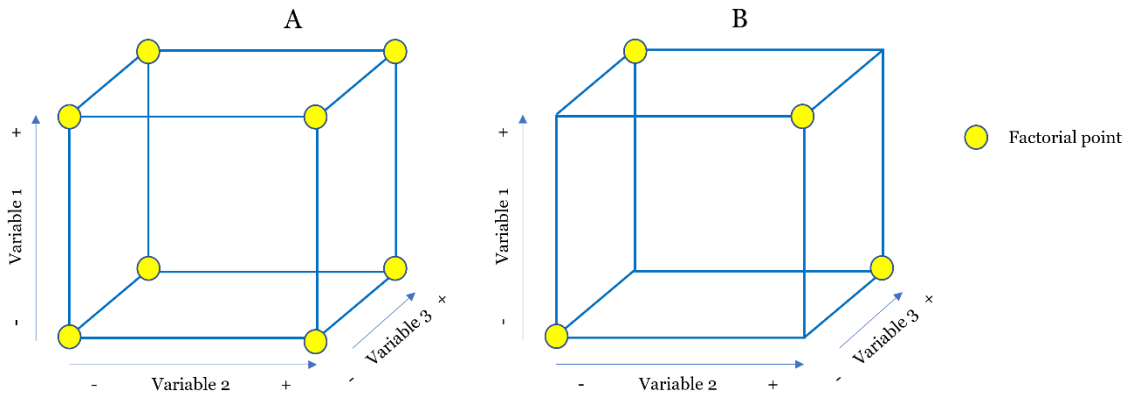


Figure 6 - Representative scheme of full factorial design (A) and fractional factorial design (B) (Adapted from [93]).

responses. For instance, Teglia and co-workers [100], developed a gas chromatography method using an FrFD and CCD for the determination of residual solvents in pharmaceuticals. In this work [100], three significant factors were selected from six factors, by an FrFD and were optimized using a CCD. The experiment runs were reduced from 64 to 24 compared with FFD. Thus, this reported method successfully separated the analytes in study with higher resolution in a shorter period of time.

1.6.4. Plackett-Burman design

PBD represents one of the most used screening designs, being employed in two different manners. First, as a screening design, it is used for the recognition of the most important factors among a high number of factors [101]. For instance, Liu and co-workers took advantage of this characteristic to select three key factors from five analysis parameters for BBD optimization, which have important effects on the signal area and resolution of seven chelating agents [102]. Second, PBD is used for ruggedness or robustness studies during method validation. Furthermore, the main advantage is the evaluation of many factors with the smallest number of experiments, nevertheless, it is only used to estimate the main effects of the factors since it does not take interactions among factors into account [102].

1.6.5. Central Composite design

CCD is a common and effective optimization design involving three design points: two-level FFD or FrFD points, central points, and star points [103]. This takes advantage of containing the two-level FFD or FrFD, which can be preliminary implemented to access the factors, however, it requires more central points, and it considers extreme conditions, which could lead to an increase of experiments or unsatisfactory results [104, 105].

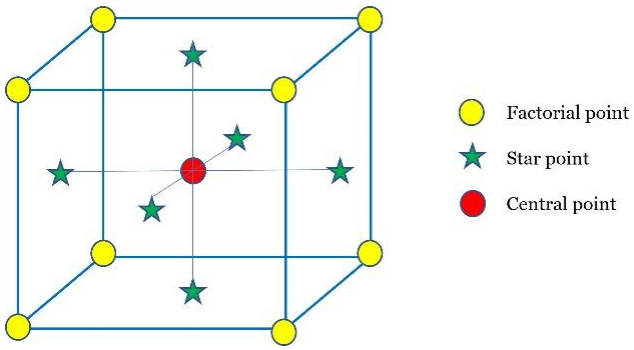


Figure 7 - Representation of the three design points of Central Composite Design (Adapted from [93]).

1.6.6. Box-Behnken design

BBD is another optimization design, but unlike CCD, the BBD does not examine the extreme factor combinations. Therefore, BBD shows its unique advantage of fewer experimental runs than CCD for the same number of factors. Nevertheless, BBD has the disadvantage that the design is not sequential, as a consequence, the experimental results of the previous batch are almost useless for the next batch of experiments, and each batch of experiments should be redone [104, 105]. The application of BBD should be contemplated for systems with more than two factors where the optimum value is known to stay in the middle of the factor ranges [101].

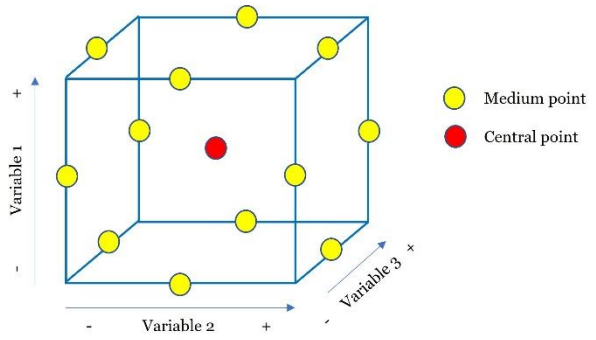


Figure 8 - Representative scheme of Box-Benheken Design (Adapted from [93]).

1.6.7. Mixture design

MD is utilized when a few factors are subject to conditions where the sum of the proportions of all components of a mixture is 100%. It differs from other optimization designs because the response produced is a function of the proportions of each component. The applications of MD mainly focus on the optimization of the organic solvent ratio for the extraction of organic compound and the mobile phase composition in chromatographic processes [106]. For instance, if there are a mixture of 3 components such as methanol, acetonitrile and water in a mobile phase it is possible to fix one the components when the other two are changed. Then, for this particular design, its plane area becomes a triangular plane in the three-dimensional factor space [101].

1.6.8. Liquid chromatography (LC) optimization

LC is an important branch of chromatography and takes liquid as mobile phase to separate different components in the stationary phase. Considering that the factors such as mobile phase pH, buffer concentration, flow rate, column temperature, loaded sample will have a great impact on separation, the application of DoE in LC has great advantages [93]. Narenderan and colleagues [107], used DoE for the validation and optimization of a stability-indicating liquid chromatography/mass spectrometry method to identify the major oxidative degradation products in pemetrexed, which is an antitumor drug, and it is crucial to monitor its concentration in human plasma. In this study, flow rate, percentage of acetonitrile in mobile phase, and buffer pH were investigated with BBD and was considered a credible and sensitive method with good results. In another study, Sharma and collaborators [108], wanted to determine 10 impurities in tenofovir alafenamide fumarate tablets, and for this, they developed an ultra-performance LC analytical method. Then, to identify the significant factors, including column temperature, the proportion of tetrahydrofuran in organic solvent, and the proportion of mobile phase B at the end of the first gradient, a PBD was applied. Additionally, to optimize the process, a BBD was used to get a good resolution and short retention time of tenofovir alafenamide fumarate peak. In our research group, [109], a DoE was successfully develop for the purification of pre-miRNA-29 by a new O-phospho-L-tyrosine affinity chromatography strategy. A BBD was used to identify optimum conditions in the separation process. The key factors selection was based on the three most critical factors that affect the chromatography process, being these the column temperature, binding and elution conditions and the responses were relative recovery and purity. The group obtained satisfactory results, being the final values of recovery and purity, 84.24% and 26.61%, respectively. Optimization is undoubtedly a necessity when it comes to the speed of implementation of the chromatographic method. Additionally, the ability to obtain results is increased and the error that exists when optimizing by trial and error is decreased.

Chapter 2 – Global Aims

Considering the therapeutic potential of microRNAs, it is in high demand the development of efficient purification processes. Thus, developing new strategies for microRNA production with high purity degree and biologically active is extremely required. One of the strategies might be the use of the recombinant production of these biomolecules using prokaryotic hosts, such as *Rhodovulum sulphidophilum*. The expression system *R. sulfidophilum* DSM 1374 allows the production of human pre-miR-29b with a straightforward recuperation process of low molecular weight RNA. In here, the goal is not to produce high quantities of RNA but rather to optimize its purification, while preserving the biological active form. Therefore, in this work it was used a novel purification strategy, based on multimodal chromatography, by using supported ionic liquids (SILs) to purify the pre-miR-29b, in order to assure that the prerequisite for high purity of biopharmaceuticals from regulatory agencies is met.

Hence, the main objective of this work is to develop an experimental design for the chromatographic process involved in the purification of pre-miRNA-29b, by taking advantage of the multimodal interactions that occur between this molecule and an ionic liquid-based ligand immobilized onto a silica support (SilPrImPrACl).

In order to fulfill this purpose, the following assignments have been made:

- Evaluation and understanding of binding and elution profile of pre-miRNA-29b at different levels of pH and ionic strength.
- Development of an experimental design, establishing the key factors and their range of values to optimize them, according to the initial screening.
- Perform a final assay, in order to assess the recovery and purity of pre-miRNA-29b.
- Evaluate the quality of the target molecule by quantifying the levels of impurities.

Chapter 3 – Materials and methods

3.1. Materials

For solid culture media preparation, the following reagents were used: yeast extract, D-glucose anhydrous ($C_6H_{12}O_6$), 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, sodium chloride, 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 ($CaCl_2.2H_2O$), and ammonium sulfate ($(NH_4)_2SO_4$) from Panreac (Barcelona, Spain), and kanamycin from Thermo Fischer Scientific Inc. (Waltham, USA). For the extraction of low molecular weight RNA, it was 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). Chromatographic assays were carried out by using phosphate buffers that are prepared with dipotassium hydrogen phosphate (K_2HPO_4) from Panreac (Barcelona, Spain), potassium dihydrogen phosphate (KH_2PO_4) from Sigma-Aldrich (St. Louis, Missouri, USA) and sodium chloride (NaCl), acquired from Panreac (Barcelona, Spain). Milli-Q water treated with diethyl pyrocarbonate (DEPC) with a concentration of 0.01%, was used to prepare these buffers. The chromatographic support used in this work was silica beads functionalized with the ionic liquid 1-(3-aminopropyl)imidazole chloride, previously synthetised by our group, in the scope of the PUREmiRSILs project. The urea-PAGE gel was prepared with urea from Acros (Waltham, USA), acrylamide, tetramethylethylenediamine (TEMED), ammonium persulfate (APS), Tris-Borate-EDTA (TBE) buffer and formamide for sample preparation. For protein quantification it was used a Dye Reagent Concentrate and bovine serum albumin (BSA), both purchased from Bio-Rad (California, EUA). For real-time quantitative PCR (qPCR) assays, performed to quantify the gDNA, a Maxima® SYBR Green/Fluorescein qPCR Master Mix Kit from Thermo Fisher Scientific Inc. (Waltham, EUA) was used. The primers for these assays were 5'ACACGGTCCAGAACTCCTACG-3′ (forward) and 5′-

CCGGTGCTTCTTCTGCGGGTAACGTCA-3′ (reverse), both acquired from Stabvida (Caparica, Portugal).

3.2. Methods

3.2.1. RNAs Production

RNA was obtained from a culture of *Rhodovulum sulfidophilum* (RdV) transformed with the pBHSR1-RM plasmid containing the sequence of human pre-miRNA-29b. RdV 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/O_2 ratio was established as 1:5. To start the pre-fermentation process, inoculum was transferred from the solid medium to the Erlenmeyer flask and left in an orbital shaker for approximately 27 h until it reaches 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 other Erlenmeyer flask where fermentation took place. This volume was calculated by the equation below:

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

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

The fermentation process was kept for approximately 72 h and then the medium was centrifuge at 4 °C and 3900 g, for 10 min to obtain cell pellets which were stored at -20 °C. Moreover, since RdV can produce the biomolecule extracellularly, the resulting liquid medium was also stored at -20 °C.

3.2.2. Low molecular weight RNA extraction

For RNA extraction, it was used the acid guanidinium thiocyanate-phenol-chloroform method. Firstly, cell pellets were thawed at room temperature and resuspended with 0.8% NaCl solution, followed by centrifugation at 6000 g, 4 °C, for 10 min. The supernatant was discarded and 5 mL of D solution (4 M guanidinium thiocyanate, 0.025 M sodium citrate pH 7, 0.5% sodium N-lauroylsarcosinate and 0.1 M β-mercaptoethanol) was added. After an ice incubation for 10 min, 500 µL of 2 M sodium acetate pH 4 and 5 mL of phenol solution were added to the lysis tubes, with careful homogenization at each step. The lysate consists of RNA, DNA, proteins, and cellular

debris. Afterwards, 1 mL of chloroform/isoamyl (49:1) acid mixture was added, followed by vigorous shaking. Then, it was incubated for 15 min in ice, followed by a centrifugation at 10000 g, at 4 °C, for 20 min. Phase separation of the sample occurred when phenol/chloroform was added, due to the different densities. In addition, the pH 4 of the added buffer, allowed an effective separation between the nucleic acids. This allows the formation of 3 phases: the upper aqueous phase rich in RNA, the interphase rich in DNA and the bottom organic phase enriched in proteins, lipids, and cell debris. Therefore, the following step consisted in transferring the maximum volume of the upper phase into a new tube, very carefully to avoid DNA contamination. Into this new tube, it was added 5 mL of isopropanol to the RNA enriched phase for its precipitation. After homogenization and centrifugation at 10000 g, 4 °C, for 20 min, supernatant was discarded, and 1.5 mL of D solution and 1.5 mL of isopropanol were added. Once again, a centrifugation was performed at 10000 g, 4 °C, for 10 min and the supernatant was discarded. Afterwards, 2.5 mL of 75% ethanol in DEPC treated water was added and incubation at room temperature for 10-15 min took place followed by centrifugation at 10000 g, 4 °C, for 5 min. After a final discard of the supernatant, the pellet was dried at room temperature for 5-10 min before being resuspended with 1 mL of DEPC treated water and incubated for 10-15 min at room temperature. The integrity of RNA samples was evaluated by agarose gel electrophoresis and concentration was measured in the Nano Photometer (IMPLEN, United Kingdom).

3.2.3. Chromatography assays

Chromatographic assays to study the separation of pre-miR-29 from other RNAs were performed in the equipment AKTA Pure with the software UNICORN™ 6.3 (GE Healthcare Biosciences, Uppsala, Sweden). The support (SIL) under evaluation was spherical silica functionalized with the IL 1-(3-aminopropyl)imidazole chloride, the SilPrImPrACl. This support was packed in a HiScale 16 column (Cytiva, Sweden) with 16 mm diameter x 10 mm of height. For this, the support was dispersed in Milli-Q water and then added to de column until it reached about 1 cm of height, being extremely important to avoid drying the matrix during this process.

For each chromatographic assay, the column was equilibrated with 10 mM phosphate buffer (pH 6-8), previously filtered and sonicated, using a flow rate of 1 mL/min. After equilibration, samples of RNA extracted from RdV, with concentrations from 600 μg/mL to 1000 μg/mL, were independently injected using a 100 μL loop. Completed the binding step, several steps of increasing salt concentration from 1 M up to 2 M of NaCl in 10 mM phosphate buffer were applied in order to analyse different retention patterns and eventual RNA species separation. Different binding and elution conditions were

screened in these assays, namely the working pH values that ranged from 6 to 8, the amount (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 lastly analysed by urea-PAGE gel electrophoresis and using other quantification methods.

3.2.4. Column regeneration

Column regeneration is crucial to ensure the reproducibility between assays and maintaining high separation performance of these supports. For that, 3 column volumes of a solution of 0.1 M of sodium hydroxide (NaOH) were added to the support to remove any residues of bound RNA samples, followed by 3 column volumes of a solution of 0.5 M of HCl for maintaining the counterion (Cl^-) in the matrix. After the solutions were added, the column was washed with DEPC treated water in large volumes.

3.2.5. Urea-PAGE analysis

Urea-PAGE was performed in order to verify if the separation conditions used in the chromatographic assays were able to properly separate the pre-miR-29 from the rest of the RNA extract. Gels were prepared by firstly mixing 7.2 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 ethylenediaminetetraacetic acid disodium salt 2-hydrate (EDTA)), followed by heating at 40 °C. Afterwards, 5 mL of acrylamide was added to the mixture, followed by the addition of 25 μL of TEMED and 100 μL of PSA to initiate the polymerization process. Once the polymerization of the gels has 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, 15 μL of each sample was injected into the wells. 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 ultraviolet (UV) light exposure in the Bio-Rad equipment using the ChemiDoc software.

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

The Central Composite Design (CCD) was used as model for the optimization of pre-miRNA-29b purification. The factors studied were the loaded mass of RNA, NaCl

concentration in the binding step and NaCl concentration in the elution step, based on the pH screening performed (Table 3). Each variable was coded by letters A, B and C, and each one had 3 levels: -1, 0 and 1. The software used was Design Expert (version 13), and it generated 20 experiments to be performed, and finally, two responses were established: the percentage of relative recovery (R1) and the purification factor of pre-miRNA-29b (R2). The percentage of relative recovery was obtained by measuring the intensity of the band of the pre-miRNA-29b of the peak of interest relatively to the intensity of the same band in the initial sample. The purification factor is a ratio between the purity of the pre-miRNA-29b in the peak of interest and in the initial sample. The purity was calculated by measuring the intensity of the band of pre-miRNA-29b relatively to all the bands present in the analyzed lane. To measure the intensity of the bands, it was used the Image Lab 5.2.1 software (Bio-Rad, Hercules, California, USA).

Table 3 – Factors and respective low, medium, and high values for the optimization of pre-miRNA-29b purification.

Code	**Factor**	**Levels**		
		Low	**Medium**	**High**
		-1	0	1
A	Loaded mass of RNA (µg)	60	80	100
B	[NaCl] (M) in 10 mM phosphate buffer pH 7.5	1.1	1.3	1.5
C	[NaCl] (M) in 10 mM phosphate buffer pH 7.5	1.6	1.8	2

3.2.7. Protein quantification

Bradford Protein Assay was used to quantify total protein content in both RNA initial samples and in the purified fractions. A calibration curve was constructed, using BSA as standard protein with the following concentrations: 0, 0.2, 0.5, 1, 2, 3, 4, 5 and 6 µg/µL. Thus, each sample was quantified in triplicate using 159 µL of Milli-Q water, 1 µL of sample and 40 µL of Bio-Rad Protein Assay Dye Reagent Concentrate. Then, the absorbance of the plate was read at 595 nm, thereby allowing the calculation of the amount of protein in the samples using the calibration curve below.

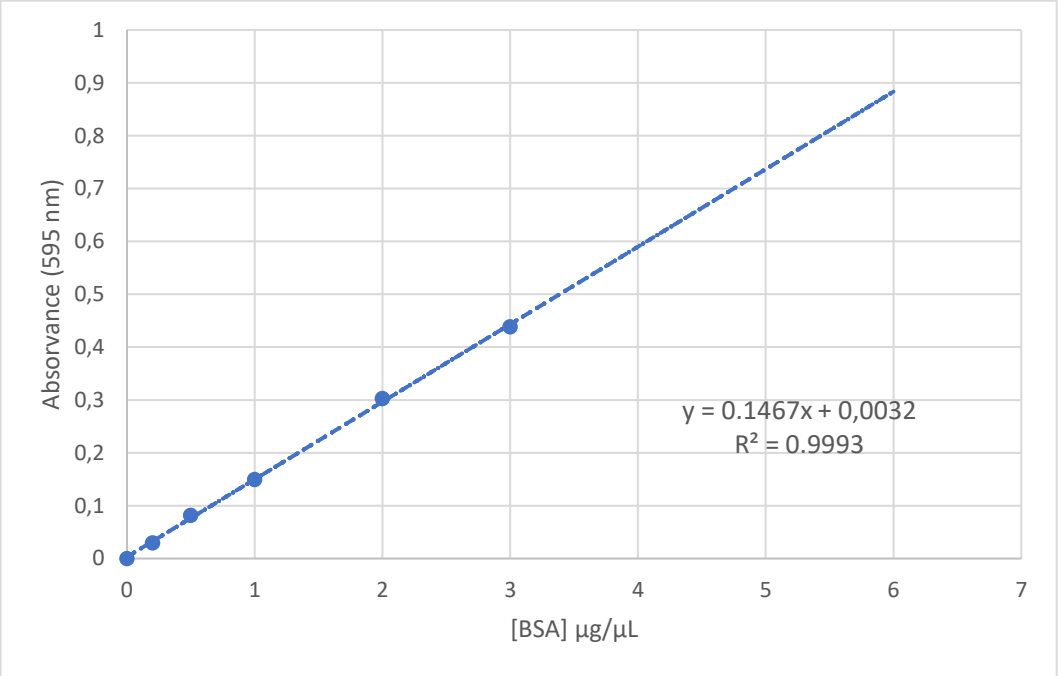


Figure 9 - Calibration curve obtained for the protein quantification, using BSA standards.

3.2.8. gDNA quantification

For quantitative analysis of gDNA present in both RNA initial samples and in the purified fractions, qPCR was performed in 96-well optical plates, using the Maxima® SYBR Green/Fluorescein qPCR Master Mix (2X) (Thermo Fisher Scientific Inc.) in a CFX Connect™ Real-Time PCR Detection System (BioRad). To quantify gDNA, a calibration curve was designed with gDNA concentrations varying from 0.005 to 50 ng/µL. The calibration curve was obtained by a correlation between the quantitation cycle (Cq) and the logarithmic of gDNA concentration.

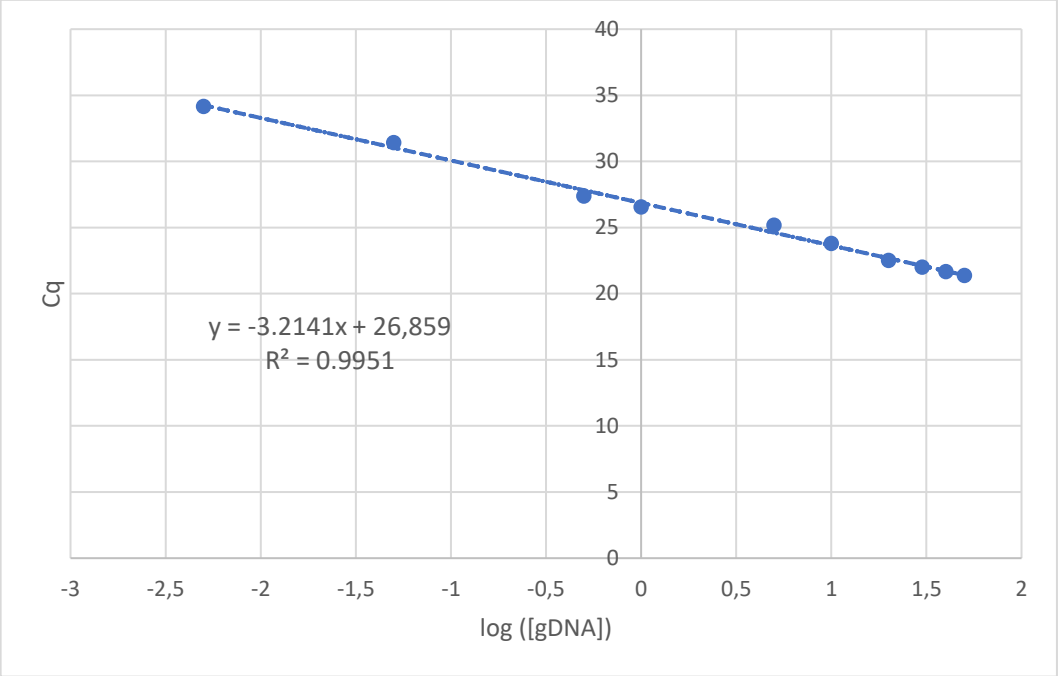


Figure 10 - Calibration curve obtained from the analysis of gDNA standards in a range of concentrations from 0.005 to 50 ng/µL.

qPCR reaction was prepared to contain: 1 µL of the sample, 1.2 µL of each specific primer (forward primer - 5'ACACGGTCCAGAACTCCTACG-3'; and reverse primer - 5'-CCGGTGCTTCTTCTGCGGGTAACGTCA-3') used to amplify a 181-bp fragment of the 16S rRNA gene, 10 µL of Maxima® SYBR Green/Fluorescein qPCR Master Mix (2X) and 6.6 µL of nuclease-free water, to a final volume of 20 µL per reaction. The reaction conditions were 95 ˚C for 10 min for initial denaturation, followed by 39 cycles of 95 ˚C for 10 sec, 60 ˚C for 30 sec, and 72 ˚C for 15 sec. In the end, the samples were incubated at 65 ˚C for 5 sec with an increment of 0.5 ˚C until 95 ˚C for the melting curves. All reactions were completed in triplicate, and Cq values were averaged from the triplicate.

Chapter 4 – Results and Discussion

4.1. Initial screening of optimal conditions for pre-miRNA-29b purification

The intermediate support (SiPrCl) functionalized with 1-(3-Aminopropyl)imidazole (SilPrImPrACl) was the studied support. In order to understand the binding and elution pattern of pre-miRNA-29b in the SilPrImPrACl support (Figure 11), several conditions were studied, namely by adjusting the ionic strength and pH values of the studied buffers, in order to determine the electrostatic interactions occurring in this process. These interactions are mainly promoted by the charged group in the cation, since RNA molecules have negative charge. However, it is important to notice that this type of support is multimodal, and can promote other interactions, for this reason it is very important to test a range of pH values and salt concentrations to analyze the differences between them. From a previous work performed in the group, it was inferred that by using this chromatographic support, an excellent separation between DNA and RNA was accomplished, when a 10 mM K_2HPO_4/KH_2PO_4 buffer solution at pH 8.0 was used for the binding step and a 1.5 M NaCl in 10 mM K_2HPO_4/KH_2PO_4 at pH 8.0 was used for the elution of RNA. Not only the ionic strength could influence the binding and elution profiles of RNA, but also the pH value of the buffers could be a key factor for establishing such interactions. Thus, a first screening of pH values was performed. The SilPrImPrACl support was packed in a chromatographic column and tested on AKTA Avant equipment to evaluate the effect of pH on the retention of RNA.

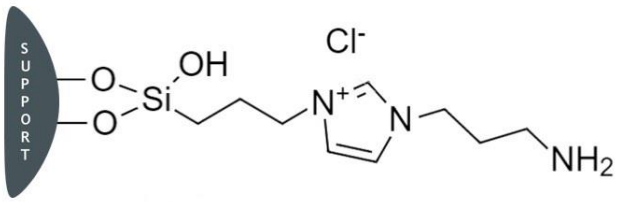
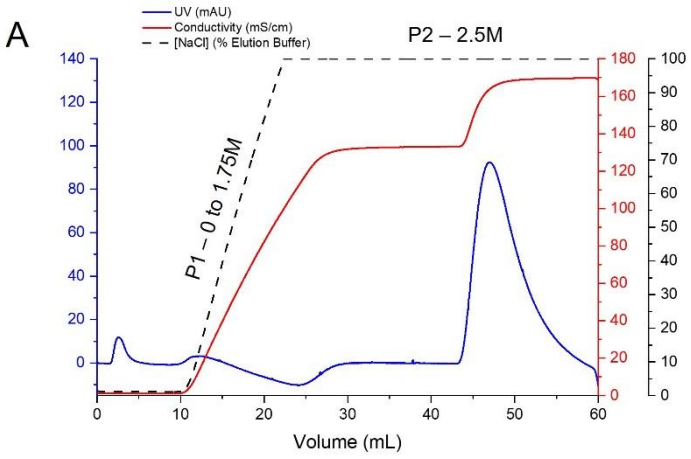


Figure 11 – Representation of the chemical structure of SilPrImPrACl support.

4.1.1. Screening at pH 6

The chromatographic assays performed at pH 6 are represented in Figure 12. Since it was observed that the binding was very effective at pH 6 the elution of the desired RNA was

carried out with a higher ionic strength than was initially foreseen. By observation of the chromatograms and respective gel of Figure 12, it is noticeable that RNA elution only occurs when high concentrations of NaCl are used. From the chromatograms represented in Figure 12 it is possible to observe the presence of two peaks, but only the peak P2 is relevant in terms of RNA concentration values. However, the purity of pre-miRNA-29b is not significant, as confirmed by the electrophoresis. Thus, by using these conditions, at pH 6, no selectivity is achieved for the pre-miRNA-29b, regarding the separation from other RNA species, since most of RNA is only eluted when 2 M or higher concentrations of NaCl are used. With this, it can be concluded that this is not the best condition to develop a purification process for the target molecule, the pre-miRNA-29b, since higher salt concentrations in the process are not advantageous since the time spent later to remove it from the samples is longer, and if not removed correctly it can influence further analysis, and the purity of the final sample would be compromised by the quantity of impurities present.



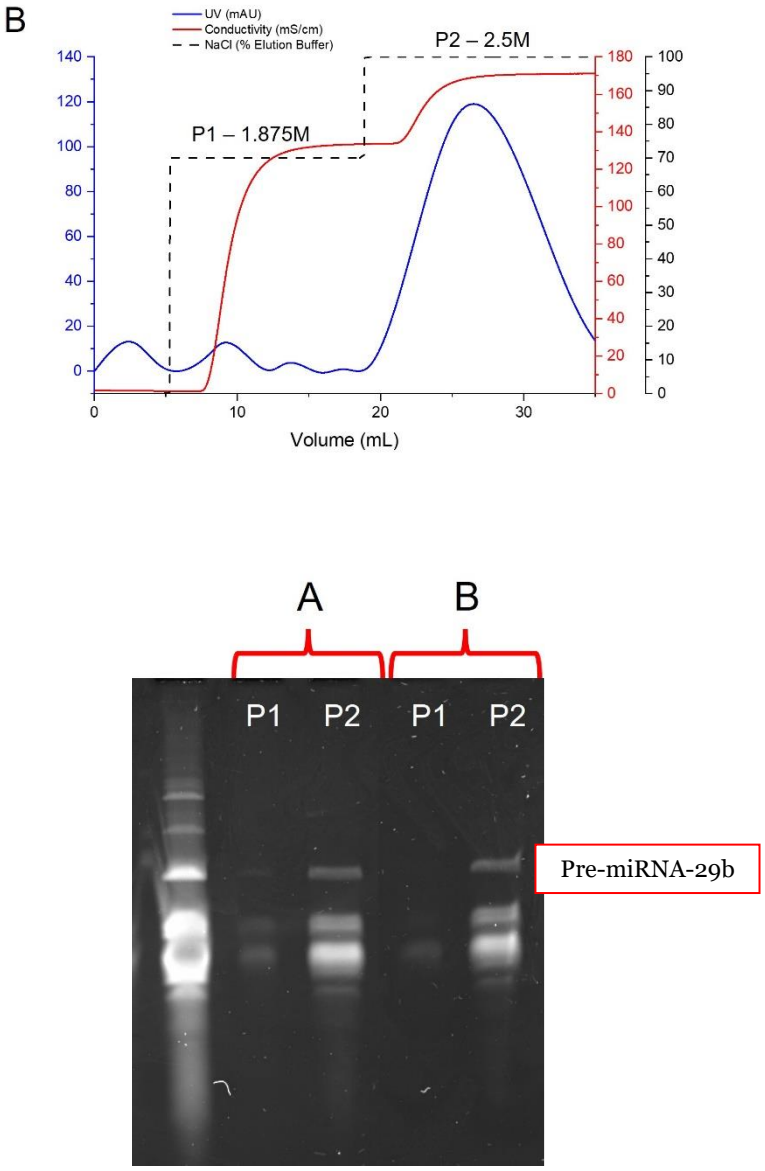


Figure 12 - Chromatograms representing the binding and elution profile at pH 6 and respective UREA-PAGE electrophoresis (Chromatogram A: P1 – Gradient from 0 to 1.75 M of NaCl and then a step performed at 2.5 M of NaCl; Chromatogram B: Stepwise gradient at 1.875 M of NaCl and 2.5M of NaCl).

4.1.2. Screening at pH 7

To improve the selectivity for pre-miRNA-29b with the SilPrImPrACl ligand, additional screening assays were performed at pH 7, and the resulting chromatograms are presented in Figure 13. Unlike the experiment at pH 6, it was not verified the need for increasing the ionic strength to values higher than 2 M to achieve RNA elution. However, when using these conditions, higher number of steps to separate the target molecule from the impurities were used, what is not desirable, because although it is not detectable in the electrophoresis gel, it might occur the loss of part of the pre-miRNA-29b during the process, which leads to compromised recovery and purity. From chromatograms A and B (Figure 13), it can be observed that at concentrations of 1.4 M and 1.5 M, respectively, the UV peaks at 260 nm are not defined and have negative values, which may represent a nonspecific interaction of the sample with the matrix and can be minimized by changing the concentration of NaCl in this step or reducing the number of steps. Nevertheless, this salt gradient and pH 7 resulted in an increased retention of pre-miRNA-29b, proving that electrostatic interactions, and other interactions occurring between the ligand and the target molecule are favored and are stronger between 1.1 M and 1.7 M, also reinforcing the multimodal behavior of the ligand. For this reason, and by comparing with the results at pH 6, it can be noticed that the pH and the concentration of NaCl have a direct influence in the retention of the target molecule, and even the impurities (mainly RNA species). Further studies at pH 8 are detailed in the next topic, following the screening study, in order to draw final conclusions about the retention behavior of the target molecule with different pH values.

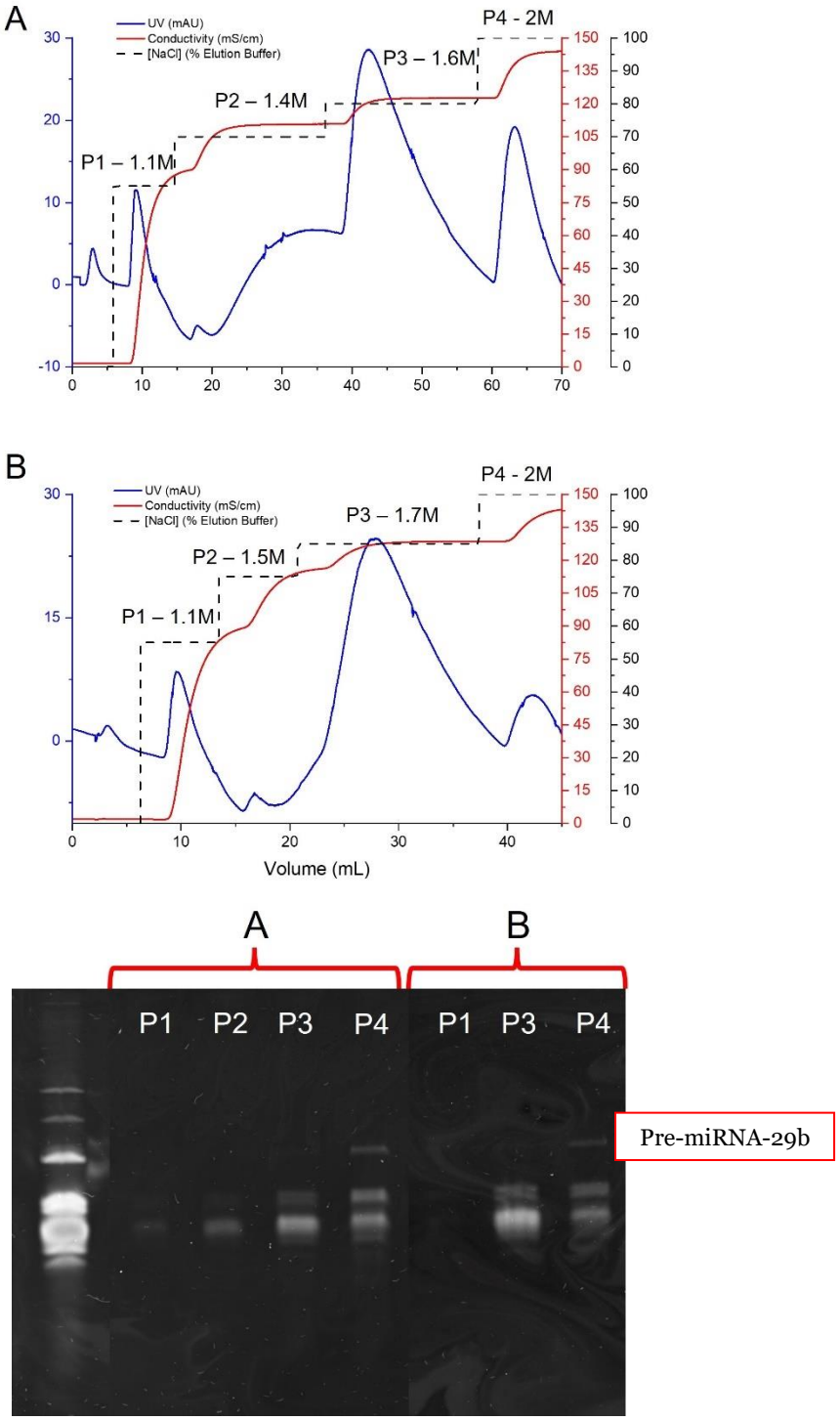
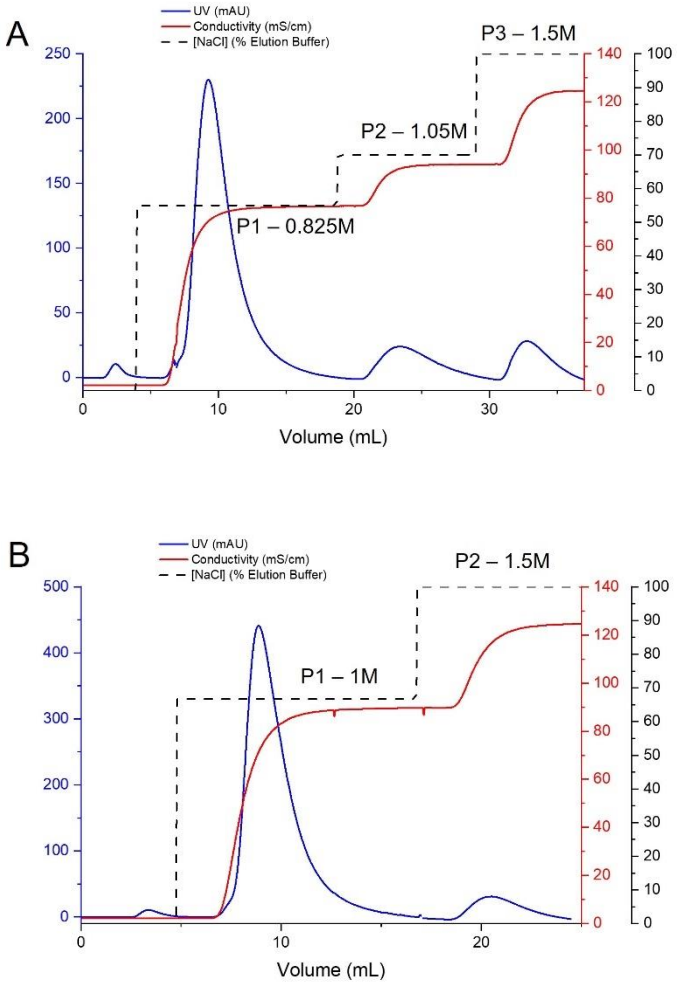


Figure 13 - Chromatograms representing the binding and elution profile at pH 7 and respective UREA-PAGE Gel electrophoresis. (Chromatogram A: P1 – First peak performed at 1.1M; P2 – Second peak performed at 1.4M; P3 – Third peak performed at 1.6M and P4 – Fourth peak performed at 2M of NaCl; Chromatogram B: P1 – First peak performed at 1.1M; P2 – Second peak performed at 1.5M; P3 – Third peak performed at 1.7M and P4 – Fourth peak performed at 2M of NaCl).

4.1.3. Screening at pH 8

To infer if there are additional differences between binding and elution profiles at different pH values, the following chromatographic assays were studied at pH 8. Theoretically, by using a higher pH value than those studied in the previous experiments, the interaction of pre-miRNA-29b with the SIL would behave differently, in this case it could be anticipated a weaker interaction with the ligand, which would lead to a lower retention time and a complete elution of RNA species. As a result, it was studied a maximum NaCl concentration of 1.5 M, as it can be seen in figure 14.



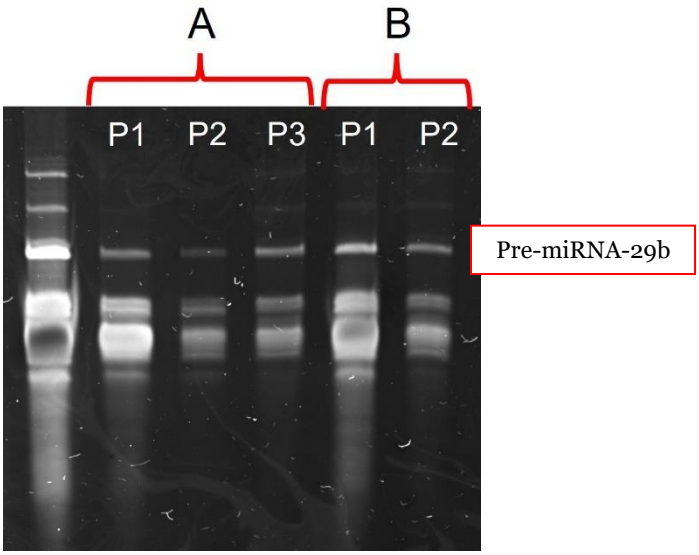


Figure 14 - Chromatograms representing the binding and elution profile at pH 8 and respective UREA-Page Gel electrophoresis. (Chromatogram A: P1 – First peak performed at 0.825M; P2 – Second peak performed at 1.05M and P3 – Third peak performed at 1.5M; Chromatogram B: P1 – First peak performed at 1M; P2 – Second peak performed at 1.5M).

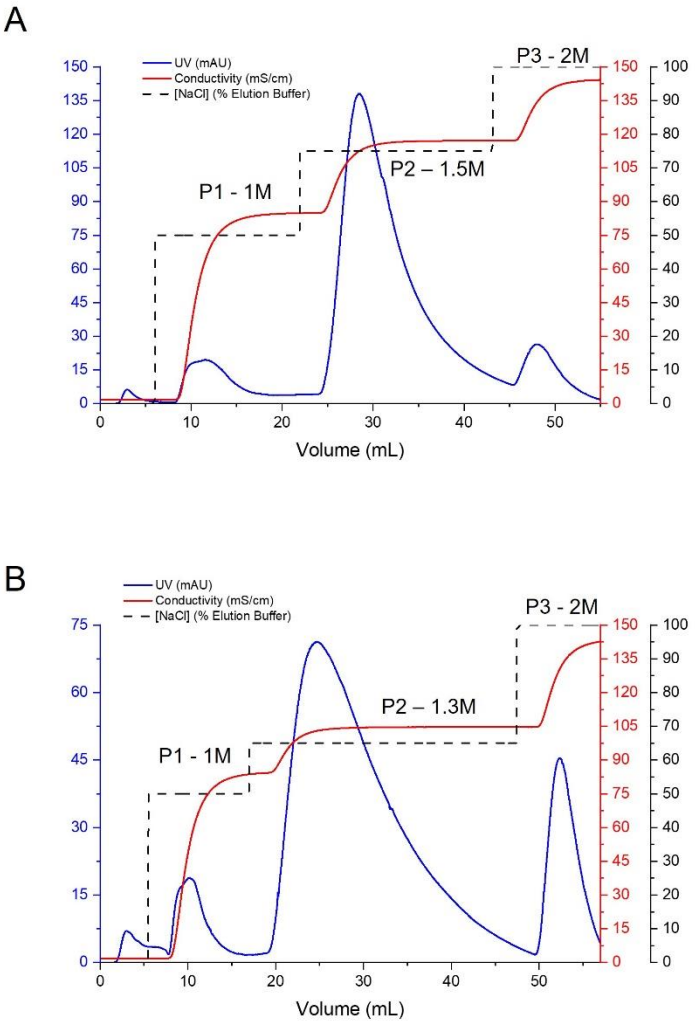
By observing the Figure 14, the initial thesis is, in fact, confirmed. Two assays were performed, where in the first (A) three elution steps were tested and in the second (B), only two steps. Despite this, there are no significative differences in the purification of pre-miRNA-29b and its separation from the other RNA species. Thus, these binding and elution conditions do not allow the desired degree of purification. This is additionally confirmed by the electrophoresis gel, where it can be observed that, on one hand, a significant amount of pre-miRNA-29b is lost throughout the process, as it is not bound onto the column in the first step (low recovery yield), and second hand, the purification of the pre-miR-29b was not achieved, as visualized in P3 (chromatogram A) and P2 (chromatogram B) electrophoretic analysis, where different bands of other RNAs are still present. Furthermore, it is clear that the pH has a great influence on the retention since an increase in the pH of the buffer from 6.0 to 8.0 significantly decreased the retention of the pre-miRNA-29b. At pH 8, there are no retention visible of this molecule which ends eluting with the other impurities, showing no selectivity when compared with previous studies in our research group. For example, using a carbonyldiimidazole monolith modified with agmatine, it was possible to successfully separate the pre-miRNA-29 using a stepwise gradient, being verified the elution of the majority of RNAs at 1.75 M NaCl and the pre-miRNA-29 at 2.5 M at pH 9.5. This pH value offers a higher retention of pre-miRNA-29 in this support, showing the differences in interactions between the agmatine and the SIL used in this work [110]. Additionally, the group was also able to separate pre-miRNA-29b from other RNAs with an increasing NaCl

concentration at pH 8, from 280 mM (elution of RNA species) to 360 mM (elution of pre-miRNA-29) using the L-arginine-Sepharose 4B support.

Overall, working in this range of pH from 6 to 8, showed that pH has a strong influence in pre-miRNA-29b binding and elution behavior due to the different interactions that can occur between the molecule and the SIL, proving the multimodal characteristic of these types of support. It is noticeable that at a lower pH the interactions are stronger, which suggests favored electrostatic interactions but also the occurrence of other non-covalent interactions. On the opposite, at pH 8, the lack of retention capacity demonstrate that such electrostatic interactions are much weaker, and others occur less often.

4.1.4. Final pH Screening

As it was disclosed so far, the binding and elution conditions are crucial for an optimum separation and purification of pre-miRNA-29b. Since the use of high salt concentrations can be a disadvantage for preserving the RNA integrity and stability, it was thus decided to use a maximum of 2 M NaCl for RNA elution. Regarding the pH value, and taking into account the previous screening results, a value between 7 and 8 would be potentially advantageous, considering the retention profiles and nonspecific interactions. Hence, the chromatographic conditions used herein were a pH value of 7.5 and a concentration of 2 M NaCl for the elution (figure 15).



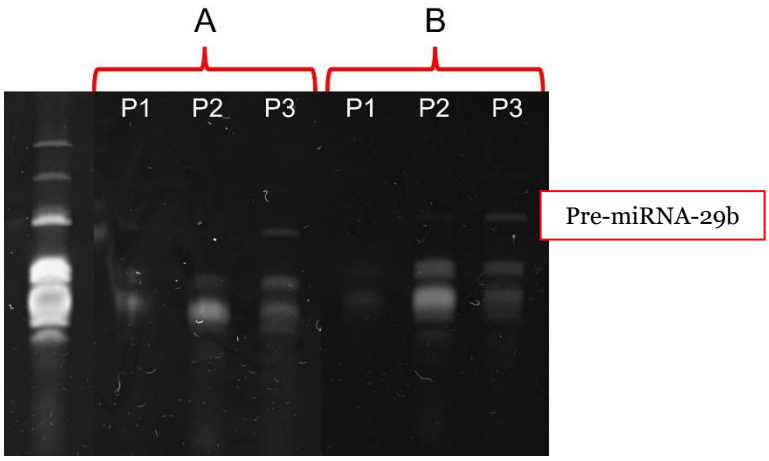


Figure 15 - Chromatograms representing the binding and elution profile at pH 7.5 and respective UREA-PAGE Gel electrophoresis. (Chromatogram A: P1 – First peak performed at 1M; P2 – Second peak performed at 1.5M and P3 – Third peak performed at 2M; Chromatogram B: P1 – First peak performed at 1M; P2 – Second peak performed at 1.3M and P3 – Third peak performed at 2M).

By analyzing the chromatograms and the respective gel in figure 15, it was verified that the results were as expected. It can be noticed that upon using a NaCl concentration of 1.5 M for the first elution step (Peak 2, chromatogram A) or 1.3 M (Peak 2, chromatogram B) the nonspecific interactions that occurred at pH 7 were in this case reversed and the pre-miRNA-29b is more retained, being eluted in the last peak, which is the desirable. Nevertheless, the profile for the separation of the target molecule from the other RNA species is quite similar to the assays performed at pH 7. Additionally, it is visible in the assays performed at pH 7.5, and also at pH 7.0, that when higher concentration of NaCl was used in the binding step (P1) it is possible to remove a higher portion of the impurities, what can be important for the purification of the target molecule.

Afterwards, with this screening results, it is proposed an experimental design with three factors to optimize purification of pre-miRNA-29b.

4.2. Design of Experiments (DoE)

After performing the initial screening of pH and NaCl concentration in the binding and elution steps, an experimental design was then proposed for the optimization of the pre-miRNA-29b purification process, maintaining the pH value at 7.5. The analysis of the Central Composite Design (CCD) model resulted in 20 experiments, in which it was obtained as results recovery values from 21% and 100%, and purification factors from 0.39 to 6.66 (Table 4). There are some recovery values that exceed the maximum theoretical value (100%), which could be due to sample contamination.

Table 4 – Central Composite Design runs and obtained responses (A: loaded mass of RNA, B: [NaCl] in binding, C: [NaCl] in elution, R1: percentage of relative recovery, R2: purification factor).

Run	A	B	C	R1	R2
1	100	1.1	1.6	25	0.39
2	80	1.3	1.8	37	4.5
3	80	1.3	1.8	155	3.67
4	60	1.1	1.6	38	1.57
5	60	1.5	2	23	5.61
6	100	1.5	1.6	0	0
7	80	1.3	1.6	87	3.54
8	80	1.3	1.8	140	3.58
9	80	1.3	1.8	67	2.25
10	80	1.3	2	49	3.49
11	100	1.5	2	74	3.63
12	60	1.5	1.6	0	0
13	80	1.3	1.8	69	3.04
14	60	1.1	2	100	0.8
15	60	1.3	1.8	52	2.78
16	80	1.3	1.8	29	1.24
17	100	1.1	2	58	0.84
18	100	1.3	1.8	35	1.33
19	80	1.1	1.8	62	1.46
20	80	1.5	1.8	21	6.66

The equations and the three-dimensional response surface plots Figure 16 (Recovery) and Figure 17 (Purification Factor)), resulting from the 20 runs performed show the

responses to the percentage of relative recovery (R1) and purification factor (R2) as a function of the different factors:

$$R1 = 56 - 2.10\ (A) - 16.50\ (B) + 15.40\ (C)$$
$$R2 = 3.34 - 0.4570\ (A) + 1.08\ (B) + 0.8870\ (C) + 1.19\ (BC) - 1.65\ (A^2)$$

By looking at the equations and the three-dimensional response surface plots for the percentage of relative recovery (R1, figure 16) and the purification factor response (R2, figure 17), it is possible to evaluate what are the main effects. The positive or the negative value designates a positive or a negative effect on the response. In the case of the percentage of relative recovery (R1), the loaded mass of RNA and the concentration of NaCl on the binding step have a negative effect which means that higher values implicate lower response, representing lower RNA recovery. The concentration of NaCl on the elution step have a positive effect, which means that a higher response comes with a higher concentration of this salt. On the other hand, only the loaded mass of RNA has a negative impact on the purification factor response (R2). Furthermore, for recovery, the model is linear, which shows that the interactions between factors are not significant, that is, each factor acts independently, which was expected and advantageous for the optimization process. If the factors acted in a dependent manner, for instance the concentration of NaCl used for binding and elution steps, the information given, in terms of recovery yield, would not be useful since the main goal of recovery is to collect the maximum of pre-miRNA-29b at the peak of interest regardless of its purity. On the other hand, for the purification factor, the model is quadratic, so it considers the interactions between the concentration of NaCl in the binding and elution step. For instance, the strongest interaction in the model is for NaCl concentrations in the binding and elution step, positively influencing the purification factor. As expected, the model indicates that, besides pH, NaCl concentration in binding and elution steps plays a crucial role in the purification process, due to the fact that the interactions are more or less specific depending on these factors. Changing the pH and salt concentration can influence the surface charge of the ligand and the interactions with various types of RNA will be altered, which in terms of selectivity and consequent purification of the pre-miRNA-29b is very important. The loaded mass of RNA alone, when at an intermediate value, boosts the purification factor to higher values, which in this case is the optimum value to work with. This can be related to the specific amount of ligand present in the support and the geometry of the column, for this reason it is important to consider that this optimal value of loaded mass of RNA found herein, could not be the appropriate value in case of using another column with a different shape or/and amount of chromatographic support. The

maximum column capacity relative to the pre-miRNA-29b is thus a dependent factor by the column geometry and matrix characteristics, unlike when changing the NaCl concentration of elution buffers, which remains independent. However, by altering properties relatively to the interaction of the target molecule with the support, such as the pH of the buffer, both the dependent and independent variables will be affected, and additional optimization studies would be required. To better understand these equations and the information given about the adjustment of the model to the two responses, it is necessary to perform the ANOVA analysis and the fitting conditions. Analyzing table ANOVA (Table 5) for the two responses (R1 and R2), the most important aspects to evaluate are F-value, p-value, lack of fit, R^2, adjusted and predicted R^2, and adequate precision. We notice that the Central Composite Design (CCD) fits in two different ways to each of these responses.

Table 5 – ANOVA table for Central Composite Design model (R1: percentage of relative recovery response; R2: purification factor response).

Response	F-value	p-value	Lack of Fit	R^2	Adjusted R^2	Predicted R^2	Adequate precision
R1	1.02	0.4078	0.8937	0.1612	0.0039	-0.1577	3.7196
R2	7.24	0.0015	0.5574	0.7211	0.6215	0.4631	8.7944

For recovery, we noticed that this model is not significant due to the low F-value of 1.02, and a p-value of 0.4078, which means that there is a 40.78% probability that the F-value may be influenced by noise. The lack of fit is not significant, which is positive considering that the desirable is that the values fit the model, in the case of the relative percentage of recovery (R1), a linear model. As for the fit parameters, R^2 describes the proportion of the total variability explained by the model, in this case, 0.1612 (16.12%), the adjusted R^2 represents the adjustment of the statistic to the size of the model, which means, unlike R^2, it considers only the factors and interactions that influence the response. In this case, and because the model is not significant, the adjusted R^2 assumes a low value, of 0.0039. Finally, the predicted R^2 evaluates the capacity of the model in the prediction of new data, in this case, the value is -0.1577, which reveals the inability of the model to predict this response (R1) adequately. The adequate precision is the ratio between signal and noise, and a value greater than 4 is acceptable and desirable. In this case, it is 3.7196, which represents an inappropriate signal. On the other hand, the analysis of the statistical data in relation to the purification factor (R2) shows that the model is significant with an F-

value of 7.24 and a p-value of 0.0015, that is, there is a probability of 0.15% of this F-value is influenced by noise. The lack of fit assumes the value of 0.5574 and is not significant, which, as previously mentioned, is desirable and proves that the quadratic model fits into R2. The fit parameters show improvements in relation to the response related to the relative percentage of recovery (R1), with an R^2 of 0.7211, an adjusted R^2 of 0.6215, which is in accordance with the significance of the model for this response, and a predicted R^2 of 0.4631. Finally, the signal-to-noise ratio is 8.7944, i.e., greater than 4, which indicates an appropriate signal and proves that the model can be used to navigate the design space. Overall, the model shows a distinguished fit for the two responses, one being not significant (R1) and the other significant (R2). Therefore, in this study the model is used, but it is not a reliable source of optimization for the relative percentage of recovery and the purification factor.

In general, it is important to understand how to improve the design, namely by performing a new screening using appropriate software, such as Design Expert, where we could use another model more suitable for this screening such as Plackett-Burman Design (PBD). In this case, since PBD only estimates the main effects of the factors and does not take interactions among factors into account, it would be interesting to make an overview of which factors can actually influence the chromatographic process, i.e., pH and NaCl concentration used, the loaded mass of RNA, column pressure, column size, retention time, flow rate and temperature. With such model, it could be then understood which factors have the most influence on the process, and also, the ideal range of values to work with these same factors. Then we could reapply the Central Composite Design (CCD) with the new factors and analyze the results and consequently be able to validate the model. Moreover, CCD model requires more center points, and it considers extreme conditions, which could lead to an increase of performed experiments and/or unsatisfactory results, which seems to be the case of this study. Box-Behnken Design (BBD) is another possible model to be also considered in further studies, since unlike CCD, the BBD does not examine the extreme factor combinations and has the additional advantage of needing fewer experimental runs than CCD for the same number of factors. Hence, the application of BBD should be contemplated for systems with more than two factors where the optimum value is known to stay in the middle of the factor ranges [101].

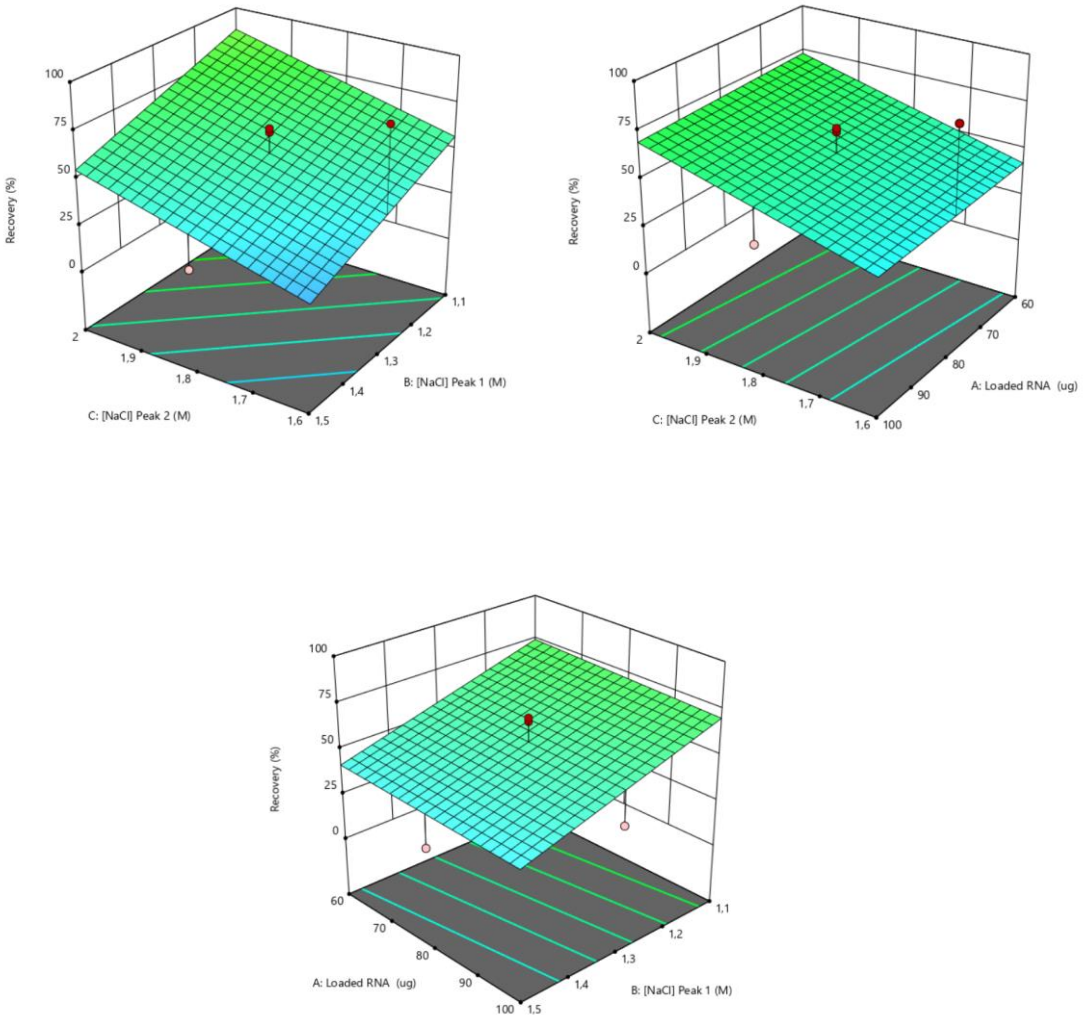


Figure 16 - Three-dimensional response surface plot of interactions of variables binding, elution and loaded mass of RNA and its effect on the percentage of relative recovery response. One parameter for each graph is at a hold value, corresponding to the central point of the absent variable.

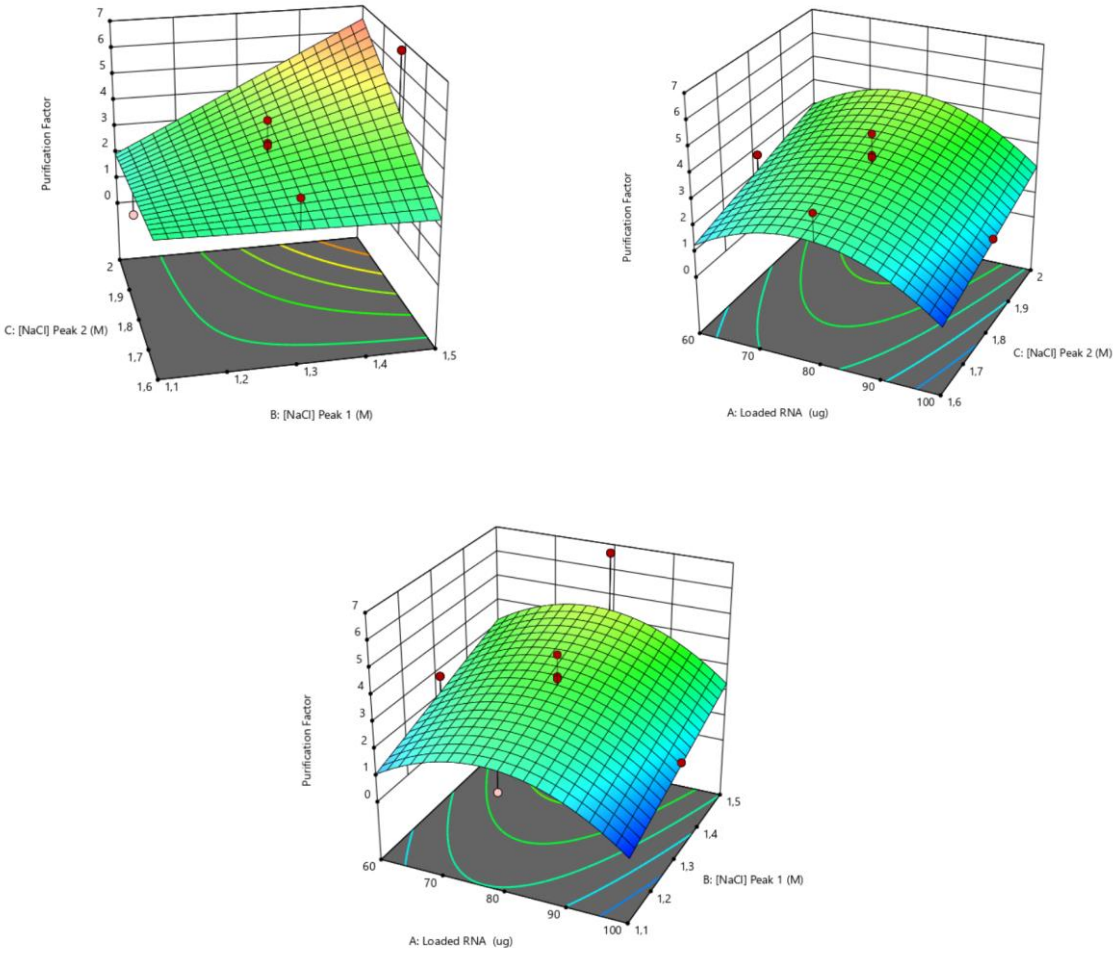


Figure 17 - Three-dimensional response surface plot of interactions of variables binding, elution and loaded mass of RNA and its effect on the purification factor response. One parameter for each graph is at a hold value, corresponding to the central point of the absent variable.

4.2.1. Model validation and reproducibility

In order to validate the method and verify the reproducibility of the column, the chromatographic profiles of the six central points of the experimental design were used since the conditions of equilibrium, binding, elution and loaded mass of RNA were maintained. Therefore, all six chromatographic profiles showed similarity as can be seen in Figure 18. It is important to notice that a regeneration step of the columns is performed every 4 assays, being this crucial to ensure reproducibility between trials and to maintain high separation performance of these supports. There are some variabilities in the height of the peaks of chromatograms A and B, what can be explained by the amount of sample that effectively eluted. This can possibly be associated with the manipulation of the sample during its homogenization, due to a poorer regeneration or due to discrepancies of mass transfer effects, for instance if there are small temperature changes between these assays. Furthermore, it is possible to observe in Figure 18 that the time of each assay is very similar, since the retention time does not fluctuate as much, being this advantageous for further evaluating the accuracy of results. This proves the reproducibility of the column and the potential of this method for pre-miRNA-29b purification.

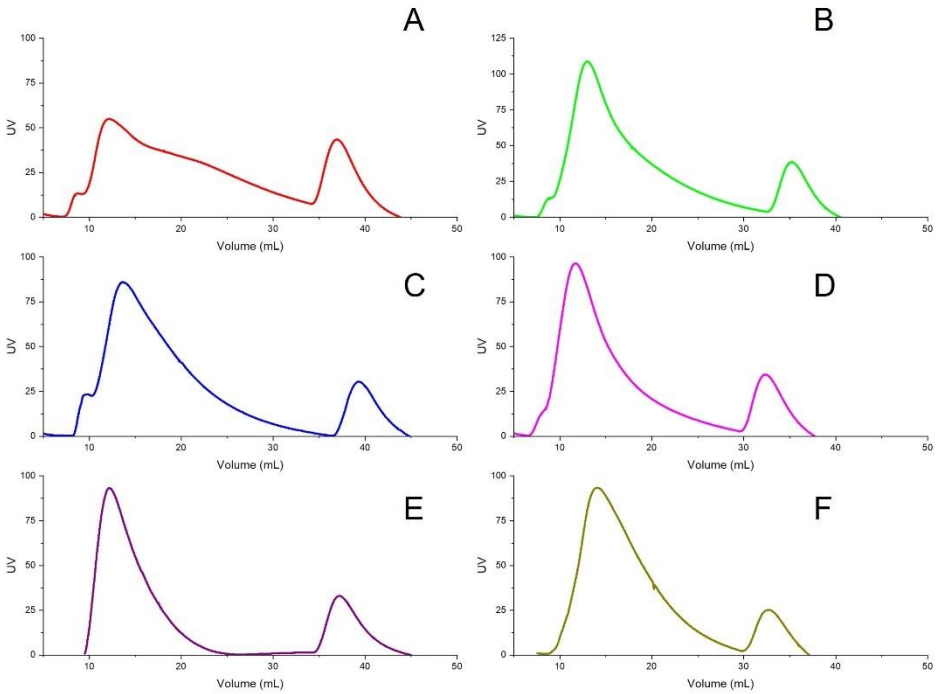


Figure 18 - Representative chromatograms for the validation of the method. Conditions were maintained in the six assays: loaded mass of RNA: 80 µg; [NaCl] in binding: 1.3 M; [NaCl] in elution: 1.8 M.

4.3. Final optimization

Although the DoE model used herein is not very suitable for predicting the ideal range of values that increases the response factors evaluated: the purification factor and specially the percentage of relative recovery, because it was not significative, it was possible to identify a range of values for each studied factors that were better for obtaining optimal purification factors. This can be inferred by the analysis of the three-dimensional response surface plots of the assays performed in the initial screening (Figures 16 and 17), where it was revealed that the interaction of NaCl concentration in the binding and elution steps is potentially significant for the purification process. It was already evident that a higher concentration of this salt in the first step is important to separate the pre-miRNA-29b from the other RNA species, hence the conditions chosen in this final assay were: 75 µg of loaded RNA, 1.5 M of NaCl in the binding step, and 2 M of NaCl in the elution step. The representative chromatogram and electrophoresis can be seen bellow, in figure 19.

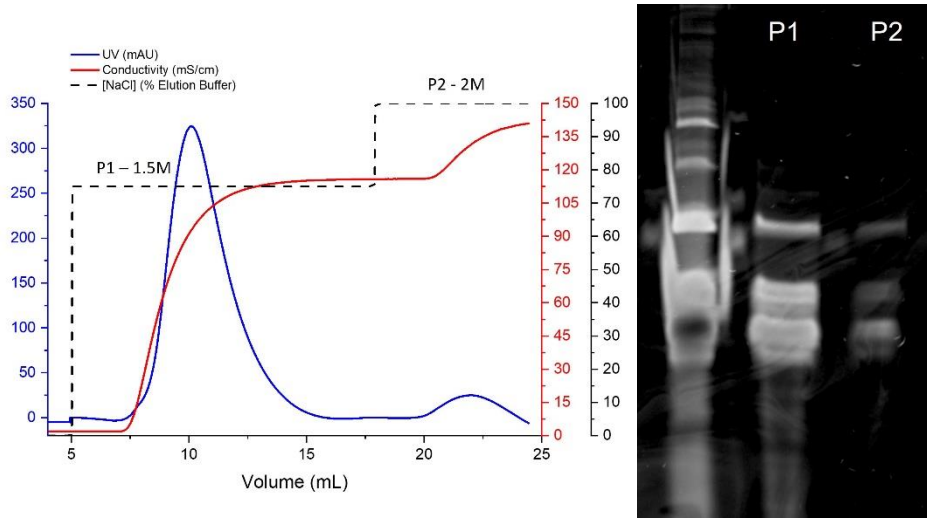


Figure 19 - Representative chromatogram of the final assay for the purification of pre-miRNA-29b and respective UREA-PAGE Gel electrophoresis (P1 -First peak performed at 1.5M; P2- Second peak performed at 2M).

As seen before, these conditions provided the removal of the majority of contaminants and some separation of the target molecule, however, the amount of pre-miRNA-29b in the first peak is still high, which resulted in 35.1% recovery, despite the high purification factor of 1.55. Furthermore, it was important to determine the quality of the final sample, by performing the quantification of protein and gDNA.

Table 6 - Protein and gDNA quantification in the initial sample, binding step, and elution step.

Samples	[Proteins] µg/µL	[gDNA] ng/µL
Initial	0.087	6.34
Impurities fraction (Binding Step)	0.057	3.13
Pre-miRNA-29b fraction (Elution Step)	0.019	1.95

These results indicate that the conditions used herein were able to effectively reduce most of the impurities throughout the process. By the analysis of Table 6, it is possible to conclude that in the final sample of purified pre-miRNA-29b was possible to remove 78% of proteins and 70% of gDNA. It is noticeable that most of these contaminants are eluted in the first gradient step, which is desirable and proves the selectivity of the column relatively to the pre-miRNA-29b. Notice that each sample had a different volume, therefore, this percentage values were calculated based on the mass of proteins and gDNA in each sample. Additionally, the fact that the column is able to separate contaminants from the pre-miRNA-29b and thus reduce them in the final sample, is a requirement for future therapeutic applications, and consequently a good indicator of the method for the purification of biomolecules with therapeutic potential. In a therapeutic point of view, this is very important since these contaminants can induce immunogenic responses in patients, and in the case of gDNA, it is a concern that it can be integrated into the genome of the patient [111]. Therefore, quantification and reduction of possible contaminants in this type of samples is very important regarding the therapeutic application, being thus proved that this SIL is highly efficient upon removing the impurities from the pre-miRNA-29b. Regarding the low recovery yield obtained herein, it is well known that chromatographic processes will sometimes give lower recovery values specially if high purification factors are achieved, as in the case shown in this last assay. Moreover, when high concentrations of NaCl concentrations are used for the elution in the chromatographic steps it is more likely for losses of the target molecule due to degradation and destabilization of native structure or the formation of inactive aggregates and unfolded species that cannot be recovered from the column [112, 113].

Chapter 5 – Conclusions and future perspectives

In recent years, studies on biomolecules with therapeutic potential have been increasing, however, efficient, and economically feasible processes to obtain these biomolecules purified from impurities while still being active, are currently in high demand. Therefore, in this work, we proposed a method using a novel ionic liquid (IL)-based chromatographic support for the purification of pre-miRNA-29b. Primarily, a binding/elution screening was done by changing the pH values from 6 to 8, and the RNA species were eluted upon increasing concentrations of NaCl. It was verified that at pH 6 the binding of the RNA was quite strong towards the chromatographic support (SilPrlmPrACl), and thus the elution was only possible by using a high salt concentration and at pH 8. It was then concluded that a pH value between 7 and 8 would be ideal. With this initial screening, an optimal pH value of 7.5 was chosen for further purification studies and an experimental design was proposed with the recovery and purification factor as responses. The selected factors were studied at three levels as: loaded RNA mass (60 μg, 80 μg, 100μg), [NaCl] in 10 mM phosphate buffer at binding step (1.1 M, 1.3 M, 1.5 M), and [NaCl] in 10 mM phosphate buffer at elution stage (1.6 M, 1.8 M, 2 M). The 20 assays performed resulted in recovery values from 21% to 100%, and purification factors from 0.39 to 6.66. The analysis of the results showed that the model is not robust enough to make a reliable prediction of an optimal point for the optimization of the purification process. Moreover, the choice of a particular chromatographic condition and mode of operation should be conducted to be compatible with the overall process purification as well as it should give an appropriate balance of recovery yield, final purity, and product quality [114]. Therefore, it was chosen, within the range of values of the factors studied, those that would increase the purification factor, namely by using a loaded mass of 75 μg and [NaCl] in 10 mM phosphate buffer at binding and elution of 1.5 M and 2 M, respectively. A final assay was thus performed under these conditions and recovery values of 35.1% and a purification factor of 1.55 were obtained. In addition, the quantification of contaminants showed that the process removed about 78% of the proteins present in the sample and 70% of the gDNA, what is a great result considering the complexity of the initial sample and that only a purification step was used. It was also possible, from six assays performed under the same equilibrium, binding, and elution conditions, to prove the reproducibility of the column.

In summary, the supported ionic liquid in study, SilPrlmPrACl, showed to be suitable for the purification of pre-miRNA-29b, however, in the future, it is necessary to understand how to optimize the process with the help of a different experimental design model.

Starting with screening, it would be important to perform a more quantitative rather than qualitative screening, i.e., to study the influence of pH and NaCl concentration and other factors that may influence both purification and recovery, such as loaded RNA mass, column pressure, column size, flow rate, and temperature using a statistical model such as Plackett-Burman Design. This way we would be able to see which factors really influence the purification process, and a more reliable range of values. Afterwards, an optimization model, such as Central Composite Design (CCD), could again be reapplied, which would potentially be more robust for predicting an optimal working zone for the purification of pre-miRNA-29b.

Chapter 6 – Bibliography

1. Zhao Y, Zuo X, Li Q, Chen F, Chen YR, Deng J, Han D, Hao C, Huang F, Huang Y *et al*: **Nucleic Acids Analysis**. *Sci China Chem* 2021, **64**(2):171-203.
2. Lehninger AL, Nelson DL, Cox MM, Cox MM: **Lehninger principles of biochemistry**: Macmillan; 2005.
3. Travers A, Muskhelishvili G: **DNA structure and function**. *FEBS J* 2015, **282**(12):2279-2295.
4. Talap J, Zhao J, Shen M, Song Z, Zhou H, Kang Y, Sun L, Yu L, Zeng S, Cai S: **Recent advances in therapeutic nucleic acids and their analytical methods**. *J Pharm Biomed Anal* 2021, **206**:114368.
5. Pierce BA: **Genetics: A conceptual approach**, vol. 1, Fourth Edition edn: Macmillan; 2010.
6. Mountain A: **Gene therapy: the first decade**. *Trends in biotechnology* 2000, **18**(3):119-128.
7. Wirth T, Parker N, Ylä-Herttuala S: **History of gene therapy**. *Gene* 2013, **525**(2):162-169.
8. Kaufmann KB, Büning H, Galy A, Schambach A, Grez M: **Gene therapy on the move**. *EMBO molecular medicine* 2013, **5**(11):1642-1661.
9. Maslova O, Koliada A, Vaiserman A: **Gene Therapy**. In: *Encyclopedia of Biomedical Gerontology*. 2019: 136-146.
10. Alhakamy NA, Curiel DT, Berkland CJ: **The era of gene therapy: From preclinical development to clinical application**. *Drug Discovery Today* 2021, **26**(7):1602-1619.
11. Baptista B, Carapito R, Laroui N, Pichon C, Sousa F: **mRNA, a Revolution in Biomedicine**. *Pharmaceutics* 2021, **13**(12):2090.
12. Guil S, Esteller M: **RNA–RNA interactions in gene regulation: the coding and noncoding players**. *Trends in biochemical sciences* 2015, **40**(5):248-256.
13. Baptista B, Riscado M, Queiroz JA, Pichon C, Sousa F: **Non-coding RNAs: Emerging from the discovery to therapeutic applications**. *Biochem Pharmacol* 2021, **189**:114469.
14. Zhong Z, Mc Cafferty S, Combes F, Huysmans H, De Temmerman J, Gitsels A, Vanrompay D, Portela Catani J, Sanders NN: **mRNA therapeutics deliver a hopeful message**. *Nano Today* 2018, **23**:16-39.
15. Verbeke R, Lentacker I, De Smedt SC, Dewitte H: **The dawn of mRNA vaccines: The COVID-19 case**. *J Control Release* 2021, **333**:511-520.

16. Hombach S, Kretz M: **Non-coding RNAs: classification, biology and functioning**. *Non-coding RNAs in colorectal cancer* 2016:3-17.
17. Brandenburger T, Somoza AS, Devaux Y, Lorenzen JM: **Noncoding RNAs in acute kidney injury**. *Kidney International* 2018, **94**(5):870-881.
18. Wang J, Samuels DC, Zhao S, Xiang Y, Zhao YY, Guo Y: **Current Research on Non-Coding Ribonucleic Acid (RNA)**. *Genes (Basel)* 2017, **8**(12).
19. Yu AM, Batra N, Tu MJ, Sweeney C: **Novel approaches for efficient in vivo fermentation production of noncoding RNAs**. *Appl Microbiol Biotechnol* 2020, **104**(5):1927-1937.
20. Kirchner S, Ignatova Z: **Emerging roles of tRNA in adaptive translation, signalling dynamics and disease**. *Nat Rev Genet* 2015, **16**(2):98-112.
21. Zhang P, Wu W, Chen Q, Chen M: **Non-Coding RNAs and their Integrated Networks**. *J Integr Bioinform* 2019, **16**(3).
22. Green J, Ansari M, Ball H, Haqqi T: **tRNA-derived fragments (tRFs) regulate post-transcriptional gene expression via AGO-dependent mechanism in IL-1β stimulated chondrocytes**. *Osteoarthritis and cartilage* 2020, **28**(8):1102-1110.
23. Iwasaki YW, Siomi MC, Siomi H: **PIWI-Interacting RNA: Its Biogenesis and Functions**. *Annu Rev Biochem* 2015, **84**:405-433.
24. Ponnusamy M, Yan KW, Liu CY, Li PF, Wang K: **PIWI family emerging as a decisive factor of cell fate: An overview**. *Eur J Cell Biol* 2017, **96**(8):746-757.
25. Gallas A, Alexander C, Davies MC, Puri S, Allen S: **Chemistry and formulations for siRNA therapeutics**. *Chem Soc Rev* 2013, **42**(20):7983-7997.
26. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, Sun W, Dou K, Li H: **Circular RNA: A new star of noncoding RNAs**. *Cancer Lett* 2015, **365**(2):141-148.
27. Zhang SF, Gao J, Liu CM: **The Role of Non-Coding RNAs in Neurodevelopmental Disorders**. *Front Genet* 2019, **10**:1033.
28. Bridges MC, Daulagala AC, Kourtidis A: **LNCcation: lncRNA localization and function**. *J Cell Biol* 2021, **220**(2).
29. Tang XJ, Wang W, Hann SS: **Interactions among lncRNAs, miRNAs and mRNA in colorectal cancer**. *Biochimie* 2019, **163**:58-72.
30. Pereira P, Queiroz JA, Figueiras A, Sousa F: **Current progress on microRNAs-based therapeutics in neurodegenerative diseases**. *Wiley Interdiscip Rev RNA* 2017, **8**(3).

31. Calore F, Lovat F, Garofalo M: **Non-coding RNAs and cancer**. *Int J Mol Sci* 2013, **14**(8):17085-17110.
32. Kyzar EJ, Bohnsack JP, Pandey SC: **Current and Future Perspectives of Noncoding RNAs in Brain Function and Neuropsychiatric Disease**. *Biol Psychiatry* 2022, **91**(2):183-193.
33. Pereira P, Pedro AQ, Queiroz JA, Figueiras AR, Sousa F: **New insights for therapeutic recombinant human miRNAs heterologous production: Rhodovolum sulfidophilum vs Escherichia coli**. *Bioengineered* 2017, **8**(5):670-677.
34. Caruthers MH: **A brief review of DNA and RNA chemical synthesis**. *Biochemical Society Transactions* 2011, **39**(2):575-580.
35. Baronti L, Karlsson H, Marušič M, Petzold K: **A guide to large-scale RNA sample preparation**. *Analytical and Bioanalytical Chemistry* 2018, **410**(14):3239-3252.
36. Mu X, Hur S: **Immunogenicity of In Vitro-Transcribed RNA**. *Accounts of Chemical Research* 2021, **54**(21):4012-4023.
37. Ryczek M, Pluta M, Błaszczyk L, Kiliszek A: **Overview of Methods for Large-Scale RNA Synthesis**. *Applied Sciences* 2022, **12**(3):1543.
38. Martins R, Queiroz JA, Sousa F: **Ribonucleic acid purification**. *J Chromatogr A* 2014, **1355**:1-14.
39. Pereira P, Pedro AQ, Tomás J, Maia CJ, Queiroz JA, Figueiras A, Sousa F: **Advances in time course extracellular production of human pre-miR-29b from Rhodovulum sulfidophilum**. *Applied microbiology and biotechnology* 2016, **100**(8):3723-3734.
40. Umekage S, Kikuchi Y: **In vitro and in vivo production and purification of circular RNA aptamer**. *Journal of biotechnology* 2009, **139**(4):265-272.
41. Ponchon L, Dardel F: **Large scale expression and purification of recombinant RNA in Escherichia coli**. *Methods* 2011, **54**(2):267-273.
42. Li M-M, Wang W-P, Wu W-J, Huang M, Yu A-M: **Rapid production of novel pre-microRNA agent hsa-mir-27b in Escherichia coli using recombinant RNA technology for functional studies in mammalian cells**. *Drug Metabolism and Disposition* 2014, **42**(11):1791-1795.
43. Kikuchi Y, Umekage S: **Extracellular nucleic acids of the marine bacterium Rhodovulum sulfidophilum and recombinant RNA production technology using bacteria**. *FEMS microbiology letters* 2018, **365**(3):fnx268.

44. Suzuki H, Ando T, Umekage S, Tanaka T, Kikuchi Y: **Extracellular production of an RNA aptamer by ribonuclease-free marine bacteria harboring engineered plasmids: a proposal for industrial RNA drug production**. *Applied and environmental microbiology* 2010, **76**(3):786-793.
45. Ando T, Suzuki H, Nishimura S, Tanaka T, Hiraishi A, Kikuchi Y: **Characterization of extracellular RNAs produced by the marine photosynthetic bacterium Rhodovulum sulfidophilum**. *Journal of biochemistry* 2006, **139**(4):805-811.
46. Coskun O: **Separation techniques: Chromatography**. *North Clin Istanb* 2016, **3**(2):156-160.
47. Ismail B, Nielsen SS: **Basic Principles of Chromatography**. In: *Food Analysis*. 2010: 473-498.
48. McKenna SA, Kim I, Puglisi EV, Lindhout DA, Aitken CE, Marshall RA, Puglisi JD: **Purification and characterization of transcribed RNAs using gel filtration chromatography**. *Nature protocols* 2007, **2**(12):3270-3277.
49. Lukavsky PJ, Puglisi JD: **Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides**. *Rna* 2004, **10**(5):889-893.
50. Fekete S, Beck A, Veuthey JL, Guillarme D: **Ion-exchange chromatography for the characterization of biopharmaceuticals**. *J Pharm Biomed Anal* 2015, **113**:43-55.
51. Sousa F, Passarinha L, Queiroz JA: **Biomedical application of plasmid DNA in gene therapy: A new challenge for chromatography**. *Biotechnology and Genetic Engineering Reviews* 2009, **26**(1):83-116.
52. Mahn A, Asenjo JA: **Prediction of protein retention in hydrophobic interaction chromatography**. *Biotechnol Adv* 2005, **23**(5):359-368.
53. Prazeres D, Ferreira G: **Design of flowsheets for the recovery and purification of plasmids for gene therapy and DNA vaccination**. *Chemical Engineering and Processing: Process Intensification* 2004, **43**(5):609-624.
54. Stadler J, Lemmens R, Nyhammar T: **Plasmid DNA purification**. *The Journal of Gene Medicine: A cross-disciplinary journal for research on the science of gene transfer and its clinical applications* 2004, **6**(S1):S54-S66.
55. Diogo M, Queiroz J, Prazeres D: **Chromatography of plasmid DNA**. *Journal of Chromatography A* 2005, **1069**(1):3-22.
56. Pereira P, Sousa Â, Queiroz J, Correia I, Figueiras A, Sousa F: **Purification of pre-miR-29 by arginine-affinity chromatography**. *Journal of Chromatography B* 2014, **951**:16-23.

57. Rodriguez EL, Poddar S, Iftekhar S, Suh K, Woolfork AG, Ovbude S, Pekarek A, Walters M, Lott S, Hage DS: **Affinity chromatography: A review of trends and developments over the past 50 years**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2020, **1157**:122332.
58. Nasirpour N, Mohammadpourfard M, Zeinali Heris S: **Ionic liquids: Promising compounds for sustainable chemical processes and applications**. *Chemical Engineering Research and Design* 2020, **160**:264-300.
59. Buszewska-Forajta M, Markuszewski MJ, Kaliszan R: **Free silanols and ionic liquids as their suppressors in liquid chromatography**. *J Chromatogr A* 2018, **1559**:17-43.
60. Ventura SPM, Coutinho JAP: **Toward the Recovery and Reuse of the ABS Phase-Forming Components**. In: *Ionic-Liquid-Based Aqueous Biphasic Systems*. 2016: 285-315.
61. Albertsson PA: **Partitioning of Cell Particles and Macromolecules**, 3rd edn. New York; 1986.
62. Quental MV, Pedro AQ, Pereira P, Sharma M, Queiroz JA, Coutinho JAP, Sousa F, Freire MG: **Integrated Extraction-Preservation Strategies for RNA Using Biobased Ionic Liquids**. *ACS Sustainable Chemistry & Engineering* 2019, **7**(10):9439-9448.
63. Shukla SK, Pandey S, Pandey S: **Applications of ionic liquids in biphasic separation: Aqueous biphasic systems and liquid-liquid equilibria**. *J Chromatogr A* 2018, **1559**:44-61.
64. Hatti-Kaul R: **Aqueous two-phase systems**. *Molecular biotechnology* 2001, **19**(3):269-277.
65. Sivapragasam M, Moniruzzaman M, Goto M: **Recent advances in exploiting ionic liquids for biomolecules: Solubility, stability and applications**. *Biotechnology Journal* 2016, **11**(8):1000-1013.
66. Morrissey S, Pegot B, Coleman D, Garcia MT, Ferguson D, Quilty B, Gathergood N: **Biodegradable, non-bactericidal oxygen-functionalised imidazolium esters: a step towards 'greener'ionic liquids**. *Green Chemistry* 2009, **11**(4):475-483.
67. Vrikkis RM, Fraser KJ, Fujita K, MacFarlane DR, Elliott GD: **Biocompatible ionic liquids: a new approach for stabilizing proteins in liquid formulation**. 2009.

68. Rodrigues JV, Prosinecki V, Marrucho I, Rebelo LPN, Gomes CM: **Protein stability in an ionic liquid milieu: on the use of differential scanning fluorimetry**. *Physical Chemistry Chemical Physics* 2011, **13**(30):13614-13616.
69. Baker SN, McCleskey TM, Pandey S, Baker GA: **Fluorescence studies of protein thermostability in ionic liquids**. *Chemical communications* 2004(8):940-941.
70. Geng F, Zheng L, Liu J, Yu L, Tung C: **Interactions between a surface active imidazolium ionic liquid and BSA**. *Colloid and Polymer Science* 2009, **287**(11):1253-1259.
71. Kumar A, Venkatesu P: **Prevention of insulin self-aggregation by a protic ionic liquid**. *RSC advances* 2013, **3**(2):362-367.
72. Attri P, Venkatesu P, Kumar A, Byrne N: **A protic ionic liquid attenuates the deleterious actions of urea on α-chymotrypsin**. *Physical Chemistry Chemical Physics* 2011, **13**(38):17023-17026.
73. Lin X, Su Z, Yang Y, Zhang S: **The potential of ionic liquids in biopharmaceutical engineering**. *Chinese Journal of Chemical Engineering* 2021, **30**:236-243.
74. Soares B, Passos H, Freire CSR, Coutinho JAP, Silvestre AJD, Freire MG: **Ionic liquids in chromatographic and electrophoretic techniques: toward additional improvements in the separation of natural compounds**. *Green Chem* 2016, **18**(17):4582-4604.
75. Pino V, Afonso AM: **Surface-bonded ionic liquid stationary phases in high-performance liquid chromatography—a review**. *Analytica chimica acta* 2012, **714**:20-37.
76. Moreira JC, Gushikem Y: **Preconcentration of metal ions on silica gel modified with 3 (1-imidazolyl) propyl groups**. *Analytica chimica acta* 1985, **176**:263-267.
77. Vidal L, Riekkola M-L, Canals A: **Ionic liquid-modified materials for solid-phase extraction and separation: a review**. *Analytica chimica acta* 2012, **715**:19-41.
78. Bernardo SC, Carapito R, Neves MC, Freire MG, Sousa F: **Supported Ionic Liquids Used as Chromatographic Matrices in Bioseparation—An Overview**. *Molecules* 2022, **27**(5):1618.
79. Xin B, Hao J: **Imidazolium-based ionic liquids grafted on solid surfaces**. *Chemical Society Reviews* 2014, **43**(20):7171-7187.

80. Shi X, Qiao L, Xu G: **Recent development of ionic liquid stationary phases for liquid chromatography**. *Journal of chromatography A* 2015, **1420**:1-15.
81. Qiu H, Mallik AK, Sawada T, Takafuji M, Ihara H: **New surface-confined ionic liquid stationary phases with enhanced chromatographic selectivity and stability by co-immobilization of polymerizable anion and cation pairs**. *Chemical communications* 2012, **48**(9):1299-1301.
82. Qiu H, Zhang M, Chen J, Gu T, Takafuji M, Ihara H: **Anionic and cationic copolymerized ionic liquid-grafted silica as a multifunctional stationary phase for reversed-phase chromatography**. *Analytical Methods* 2014, **6**(2):469-475.
83. Qiu H, Jiang S, Liu X: **N-Methylimidazolium anion-exchange stationary phase for high-performance liquid chromatography**. *Journal of Chromatography a* 2006, **1103**(2):265-270.
84. Yang B, Liu H, Chen J, Guan M, Qiu H: **Preparation and evaluation of 2-methylimidazolium-functionalized silica as a mixed-mode stationary phase for hydrophilic interaction and anion-exchange chromatography**. *Journal of Chromatography A* 2016, **1468**:79-85.
85. Zhang M, Chen J, Gu T, Qiu H, Jiang S: **Novel imidazolium-embedded and imidazolium-spaced octadecyl stationary phases for reversed phase liquid chromatography**. *Talanta* 2014, **126**:177-184.
86. Fontanals N, Ronka S, Borrull F, Trochimczuk AW, Marcé RM: **Supported imidazolium ionic liquid phases: A new material for solid-phase extraction**. *Talanta* 2009, **80**(1):250-256.
87. Selvam T, Machoke A, Schwieger W: **Supported ionic liquids on non-porous and porous inorganic materials—A topical review**. *Applied Catalysis A: General* 2012, **445**:92-101.
88. Alves L, Ramos S, Sousa F, Boto R, Almeida P: **A Benzothiazolium Salt as Chromatography Ligand for Protein Purification**. *Chromatographia* 2014, **77**(23):1597-1605.
89. Neves MC, Pereira P, Pedro AQ, Martins JC, Trindade T, Queiroz JA, Freire MG, Sousa F: **Improved ionic-liquid-functionalized macroporous supports able to purify nucleic acids in one step**. *Mater Today Bio* 2020, **8**:100086.
90. Politis S, Colombo P, Colombo G, D MR: **Design of experiments (DoE) in pharmaceutical development**. *Drug Dev Ind Pharm* 2017, **43**(6):889-901.
91. Deidda R, Orlandini S, Hubert P, Hubert C: **Risk-based approach for method development in pharmaceutical quality control context: A**

critical review. *Journal of pharmaceutical and biomedical analysis* 2018, **161**:110-121.

92. Sangshetti JN, Deshpande M, Zaheer Z, Shinde DB, Arote R: **Quality by design approach: Regulatory need**. *Arabian Journal of Chemistry* 2017, **10**:S3412-S3425.
93. Sahu PK, Ramisetti NR, Cecchi T, Swain S, Patro CS, Panda J: **An overview of experimental designs in HPLC method development and validation**. *J Pharm Biomed Anal* 2018, **147**:590-611.
94. Dejaegher B, Vander Heyden Y: **The use of experimental design in separation science**. *Acta Chromatographica* 2009, **21**(2):161-201.
95. Peng L, Gao X, Wang L, Zhu A, Cai X, Li P, Li W: **Design of experiment techniques for the optimization of chromatographic analysis conditions: A review**. *Electrophoresis* 2022.
96. Gujral G, Kapoor D, Jaimini M: **An updated review on design of experiment (DOE) in pharmaceuticals**. *Journal of Drug Delivery and Therapeutics* 2018, **8**(3):147-152.
97. Fukuda IM, Pinto CFF, Moreira CdS, Saviano AM, Lourenço FR: **Design of experiments (DoE) applied to pharmaceutical and analytical quality by design (QbD)**. *Brazilian Journal of Pharmaceutical Sciences* 2018, **54**.
98. Lee R: **Statistical design of experiments for screening and optimization**. *Chemie ingenieur technik* 2019, **91**(3):191-200.
99. EL-Shorbagy HI, Elsebaei F, Hammad SF, El-Brashy AM: **Optimization and modeling of a green dual detected RP-HPLC method by UV and fluorescence detectors using two level full factorial design for simultaneous determination of sofosbuvir and ledipasvir: Application to average content and uniformity of dosage unit testing**. *Microchemical Journal* 2019, **147**:374-392.
100. Teglia CM, Montemurro M, De Zan MM, Cámara MS: **Multiple responses optimization in the development of a headspace gas chromatography method for the determination of residual solvents in pharmaceuticals**. *Journal of pharmaceutical analysis* 2015, **5**(5):296-306.
101. Hibbert DB: **Experimental design in chromatography: a tutorial review**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2012, **910**:2-13.
102. Liu Z, Chen Y, Hu Y: **Simultaneous separation and determination of seven chelating agents using high-performance liquid chromatography based on statistics design**. *Journal of separation science* 2020, **43**(4):719-726.

103. Ferreira SL, Junior MMS, Felix CS, da Silva DL, Santos AS, Neto JHS, de Souza CT, Junior RAC, Souza AS: **Multivariate optimization techniques in food analysis–A review**. *Food chemistry* 2019, **273**:3-8.
104. Ebrahimi-Najafabadi H, Leardi R, Jalali-Heravi M: **Experimental design in analytical chemistry—part I: theory**. *Journal of AOAC International* 2014, **97**(1):3-11.
105. Ebrahimi-Najafabadi H, Leardi R, Jalali-Heravi M: **Experimental design in analytical chemistry--part II: applications**. *Journal of AOAC International* 2014, **97**(1).
106. Bezerra MA, Lemos VA, Novaes CG, de Jesus RM, Souza Filho HR, Araújo SA, Alves JPS: **Application of mixture design in analytical chemistry**. *Microchemical Journal* 2020, **152**:104336.
107. Narenderan S, Ramesh J, Babu B, Meyyanathan S: **A stability-indicating LC–MS/MS method optimization for Pemetrexed through design of experiments: Identification and characterization of major oxidative degradation product**. *Journal of Pharmaceutical and Biomedical Analysis* 2020, **183**:113150.
108. Sharma G, Thakur K, Raza K, Katare O: **Stability kinetics of fusidic acid: Development and validation of stability indicating analytical method by employing Analytical Quality by Design approach in medicinal product (s)**. *Journal of Chromatography B* 2019, **1120**:113-124.
109. Afonso A, Pereira P, Queiroz JA, Sousa Â, Sousa F: **Purification of pre-miR-29 by a new O-phospho-l-tyrosine affinity chromatographic strategy optimized using design of experiments**. *Journal of Chromatography A* 2014, **1343**:119-127.
110. Pereira P, Sousa Â, Queiroz JA, Figueiras A, Sousa F: **Pharmaceutical-grade pre-miR-29 purification using an agmatine monolithic support**. *Journal of Chromatography A* 2014, **1368**:173-182.
111. Mamnoon B, Farivar TN, Kamyab AR, Ilghari D, Khamesipour A, Arzenani MK: **Quality control of widely used therapeutic recombinant proteins by a novel real-time PCR approach**. *Iranian biomedical journal* 2016, **20**(1):56.
112. Guo J, Carta G: **Unfolding and aggregation of monoclonal antibodies on cation exchange columns: effects of resin type, load buffer, and protein stability**. *Journal of Chromatography A* 2015, **1388**:184-194.
113. Gospodarek AM, Hiser DE, O'Connell JP, Fernandez EJ: **Unfolding of a model protein on ion exchange and mixed mode chromatography surfaces**. *Journal of Chromatography A* 2014, **1355**:238-252.

114. Vázquez-Rey M, Lang DA: **Aggregates in monoclonal antibody manufacturing processes**. *Biotechnology and bioengineering* 2011, **108**(7):1494-1508.