Development of a method for the purification of minicircles

Cláudia Patrícia Almeida Alves

Thesis to obtain the Master of Science Degree in

Biological Engineering

Supervisors: Prof. Duarte Miguel de França Teixeira dos Prazeres

Prof. Gabriel António Amaro Monteiro

Examination Committee

Chairperson: Prof. Arsénio do Carmo Sales Mendes Fialho

Supervisor: Prof. Duarte Miguel de França Teixeira dos Prazeres

Member of the Committee: Prof. Maria Ângela Cabral Garcia Taipa Meneses de Oliveira

December 2014

“We used to think our future was in the stars, now we know our future is in our genes.”

James Watson

Acknowledgments

I would like to express my sincere gratitude to my supervisors, Professor Miguel Prazeres and Professor Gabriel Monteiro, for the valuable and constant support, motivation and guidance during the development of this thesis. The opportunity to work on a topic that so greatly interested me since the beginning and the knowledge transmitted by both my supervisors have contributed not only to the conclusion of this work, as well as to increase my interest in pursuing new ideas in this field of study.

I would also like to give a special acknowledgment to Gabriela Gomes for her valuable help and availability to introduce me in the beginning of this thesis to the ÄKTApurifier10 purification system and to Professor Ana Azevedo and Professor Ângela Taipa for their suggestions and help with the ÄKTApurifier10 quantification system and availability to discuss some doubts whenever I needed.

For all the support, help, friendship and for their contribution in keeping a pleasant work environment every day, I would like to give a special thanks to all my lab colleagues, namely Salomé Magalhães, Sofia Duarte, Sara S. Rosa, Sílvia Andrade, Marisa Salgueiro, Cátia Jorge, Diana Marques, Pedro Pereira, João Trabuco, Ana Rosa, Sara A. Rosa and Raquel Santos. I would also like to extend this acknowledgment to Ricardo Pereira and Rosa Gonçalves for their participation in the management of the lab material and help whenever needed.

I am also grateful to Liliana Brito who gave me the first introduction to the minicircle production strategy and the plasmid to initiate my work. Her support and friendship have been valuable during these months.

Last, but not least, I would like to dedicate this thesis to my parents and sister and appreciate their constant and unconditional support and comprehension during all these months. A very special thanks goes also to my friends, especially Tânia Dores, for the support and friendship.

Abstract

Gene therapy and DNA vaccination are fields of study in which researchers have a high interest, being registered an ongoing increase in reported clinical trials. Plasmid DNA as a non-viral delivery vector has been widely applied, having however the disadvantage of resulting in low transfection efficiency and transgene expression. The pursue of safer and more efficient delivery vectors resulted in the development of an *in vivo* recombination process in *E. coli* in which a vector comprising only an eukaryotic sequence is obtained by excision from a replicative and prokaryotic sequence containing plasmid – the minicircle DNA production system. This results in the formation, during bacterial cellular growth, of a therapeutically valuable DNA molecule – the minicircle – and of a miniplasmid comprised of the prokaryotic sequence. The high similarity in physicochemical properties of these molecules hinders the development of an efficient purification method for separation and recovery of minicircle molecules.

The principal focus of this work was the development of an efficient, simple, reproducible and scalable method for minicircle purification. With this purpose, differences in hydrophobicity were explored by modifying the structure of miniplasmids by resorting to a nicking enzyme. As a consequence, supercoiled miniplasmids were converted into their open circular forms and a single hydrophobic interaction chromatography method was successfully employed to isolate the therapeutically valuable supercoiled minicircle molecules from miniplasmid and RNA. Method robustness and reproducibility were verified. Future work must focus on process monitoring and product quality control.

Key-words: gene therapy, minicircle, enzymatic restriction, DNA-nicking enzyme, hydrophobic interaction chromatography

Resumo

A terapia genética e vacinação com DNA são áreas nas quais se tem observado um elevado interesse, sendo registado um elevado e contínuo número de ensaios clínicos. O uso de plasmídeos como vectores não virais de entrega de genes tem sido amplamente aplicado resultando, no entanto, em baixas eficiências de transfecção e expressão génica. A procura por vectores mais seguros e eficientes resultou no desenvolvimento de um processo de recombinação *in vivo* em *E. coli,* no qual um vector de sequência eucariótica é obtido por excisão a partir de um plasmídeo replicativo contendo sequências procarióticas – o sistema de produção de minicírculos de DNA. Durante o crescimento bacteriano, moléculas de interesse terapêutico – minicírculos – são formadas em conjunto com miniplasmídeos que contêm toda a informação procariótica. Um dos obstáculos ao desenvolvimento de processos eficientes para purificação de minicírculos reside na elevada semelhança físico-química entre minicírculos e miniplasmídeos.

O principal objectivo deste trabalho foi o desenvolvimento de um método eficiente, simples, reprodutível e com possível aumento de escala para a purificação de minicírculos. Com esta intenção, foi explorada uma diferença em hidrofobicidade não existente naturalmente entre estas moléculas, recorrendo a uma enzima capaz de cortar uma das cadeias do miniplasmídeo, relaxando-o. A separação do minicírculo com interesse terapêutico de moléculas de miniplasmídeo e RNA foi efectuada com sucesso num único passo de cromatografia de interacção hidrofóbica, sendo verificadas a robustez e reprodutibilidade do método. Monitorização do processo e controlo de qualidade do produto deverão ser objecto de trabalhos futuros.

Palavras-chave: terapia genética, minicírculos, restrição enzimática, DNA-nicking enzima, cromatografia de interacção hidrofóbica

Index

List of tables

List of figures

List of abbreviations

AAT – α1-antitrypsin
AEX – anion exchange chromatography
Arg – arginine
BGH – bovine growth hormone polyadenylation signal
bp – base pair
CMV – cytomegalovirus immediate early promoter
CV – column volume
DBC – dynamic binding capacity
DEAE – diethylaminoethyl
DCW – dry cell weight
DNA – deoxyribonucleic acid
E. coli – *Escherichia coli*
EDTA – ethylenediaminetetraacetic acid
FIX – factor IX
GDEPT – gene-directed enzyme prodrug therapy
gDNA – genomic DNA
GFP – green fluorescence protein
HIC – hydrophobic interaction chromatography
His – histidine
LB – Luria Bertani
LPS – lipopolysaccharides
Lys – lysine
kan^R – kanamycin resistance
mAU – mili absorbance unit
MC – minicircle
MIDGE – minimalistic, immunological defined gene expression
MiLV – micro-linear vector
MP – miniplasmid
MRS – multimer resolution site
MWCO – molecular weight cut-off
oc – open circular
OD – optical density
PAGE – polyacrylamide gel electrophoresis
PCL – packaging cell line
PCR – polymerase chain reaction
pDNA – plasmid DNA
PP – parental plasmid
RBPS – recombination based plasmid separation
RBS – ribosome binding site
Re – recombination efficiency
RNA – ribonucleic acid
sc – supercoiled
SDS – sodium dodecylsulfate
TIL – tumor-infiltrating lymphocytes
Tris – tris(hydroxymethyl)aminomethane
VEGF – vascular endothelial growth factor

1. Introduction

1.1 Gene therapy

The concept of gene therapy arose around 1970, along with the successful isolation of specific bacterial genes from DNA and complete synthesis of the gene for yeast alanine transfer RNA. The existing knowledge of genetic diseases and of the structure and function of DNA led scientists to consider the treatment of monogenic diseases as clinically possible by replacing the defective gene in the cells of the patients with foreign DNA [1]. The technological advances in molecular cloning and molecular biology in the years that followed have rendered possible the development of genetic therapeutic protocols, as well as allowed the broadening of applications for this novel technology [2].

Currently, gene therapy is defined as the introduction of nucleic acids into cells, by utilization of a delivery vector, with the aim to treat, cure or prevent a disorder by modification of endogenous gene expression. The administration of the therapeutic vector can be accomplished by *in vivo* or *ex vivo* gene delivery, with intention to insert, correct or knockdown a gene of the patient [3]. In an *in vivo* approach, the vector for gene transfer is directly injected into the patient, whereas in an *ex vivo* gene transfer protocol, the genetic modification is performed *in vitro*, using autologous cells of the patient, after which the modified cells are re-infused into the patient [2, 3].

The progression of gene therapy from a possibility to a real application was achieved in, approximately, 20 years, with the first protocol reaching phase I clinical trial in 1989 [4]. This pioneering work had as objective the optimization of the treatment of patients with advanced metastatic melanoma and was performed by a team of researchers led by Steven Rosenberg of the National Center Institute in Bethesda, USA [4]. In the study developed, the researchers used an *ex vivo* methodology with autologous cells obtained from tumors of patients. These cells were used to initiate cultures of tumor-infiltrating lymphocytes (TIL) that were later transducted with a retrovirus, derived from the Moloney murine leukemia retrovirus. The transduction was used in order to insert a gene coding for resistance to neomycin and thus allow identification of transduced cells after their infusion into patients. A major outcome of this research was the conclusion that retroviral-mediated gene insertion, with a retrovirus not capable of replication, could be a safe method for human gene therapy and that this type of vector could be useful for delivery of genes to treat inherited genetic diseases [5, 6].

Regarding the use of gene therapy for therapeutic purposes, the first clinical trial accepted for phase I, was registered in 1990 and was related with the treatment of a severe combined immunodeficiency disease (SCID) caused by a deficiency in adenosine deaminase (ADA) [4]. The protocol applied in this study involved the use of an *ex vivo* expansion of autologous T lymphocytes and retrovirus as delivery vector of the gene coding a normal copy of ADA [7]. In the beginning of the same decade, the first reports of the use of a protein-coding gene for genetic immunization were published [8, 9], demonstrating that RNA and DNA vectors directly injected into mouse tissues generated the

expression of antigens not normally produced by host cells. These findings brought to attention that DNA vectors could be used, as delivery vectors *in vivo*, for the generation of humoral and cellular immune responses against pathogens or tumor antigens, resulting in the emergence of DNA vaccines as a new field of study [10].

Despite some drawbacks and concern about the safe application of human gene therapy, this field has raised an ongoing interest among the scientific community [6], as can be assessed from the number of clinical trials reported since 1989 on the *Journal of Gene Medicine* [4]. The last reports indicate that 44 clinical trials were approved worldwide during the first six months of 2014 [4].

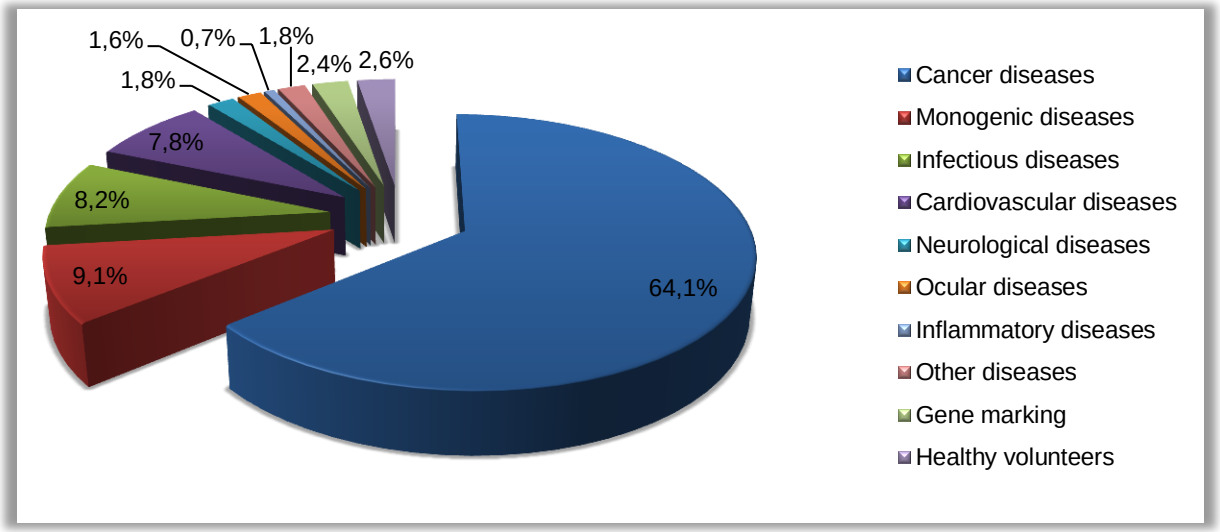

Figure 1.1: Indications addressed by gene therapy clinical trials, as reported in The Journal of Gene Medicine. Percentages refer to a total of 2076 clinical trials approved worldwide between 1989 and 2014 [4].

From the distribution of clinical trials according to their target disease (Figure 1.1), it is evident that the majority of studies is related with cancer (n= 1331), followed by monogenic (n=188), infectious (n=170) and cardiovascular diseases (n=162) [4]. The number of clinical trials focusing on cancer can be explained by the large number of patients affected worldwide and by the fact that this disease is a major cause for mortality. Along the years, a diversified number of types of cancers have been subject to attempts of gene therapy by insertion of tumor-repressing genes, immunotherapy or gene-directed enzyme prodrug therapy (GDEPT – a strategy based on the delivery of a gene that encodes an enzyme which, although being non-toxic, is able to convert a prodrug into a potent cytotoxin). A few examples of cancers targeted by gene therapy include lung, skin, haematological and nervous system tumors [11].

The strategies relying on the insertion of tumor suppressor genes account for 162 clinical trials up to date [4] and involve the delivery of genes coding for proteins (as the gene coding the tumor protein p53 that has been the gene most used in these clinical trials) that are able to halt tumor growth and promote tumor cells apoptosis [6].

In cancer immunotherapy, several approaches have been applied to control or eradicate tumors by the increase of the patient weak immune response against tumor antigens, including the intra-tumoral injection of vectors encoding cytokines or major histocompatibility molecules. The development of

vaccines has also been attempted with naked DNA, tumor cells engineered to express immunostimulatory molecules, recombinant viral vectors encoding tumor antigens and also with dendritic cells expressing tumor antigens or tumor-derived RNA [6].

1.1.1 Delivery vectors

In gene therapy, the introduction of genetic material into cells in order to achieve the desired therapeutic effect relies on the safe and efficient utilization of a delivery vector. The vectors developed with this objective fall mainly into two categories – viral and non-viral vectors – and should be able to safely deliver the gene of interest into cells and guarantee its transcription, while preserving the gene from degradation [12].

As shown in Figure 1.2, viral vectors have been the most used in gene therapy clinical trials, with 897 reports of adenovirus and retrovirus application. In recent years, with the improvement of transfection efficiency of non-viral delivery systems, the number of clinical trials that resort to this strategy has been steadily increasing, reaching currently a number almost three times higher (n=379) than reported until 2004 (n=132) [4, 6]. The interest in non-viral vectors can be explained by concerns regarding the safety of viral vectors and by the fact that they are relatively simple to produce, exhibit low toxicity and immunogenicity in transfected hosts, allowing repeated applications, and theoretically present no limit in the size of the gene inserted [12].

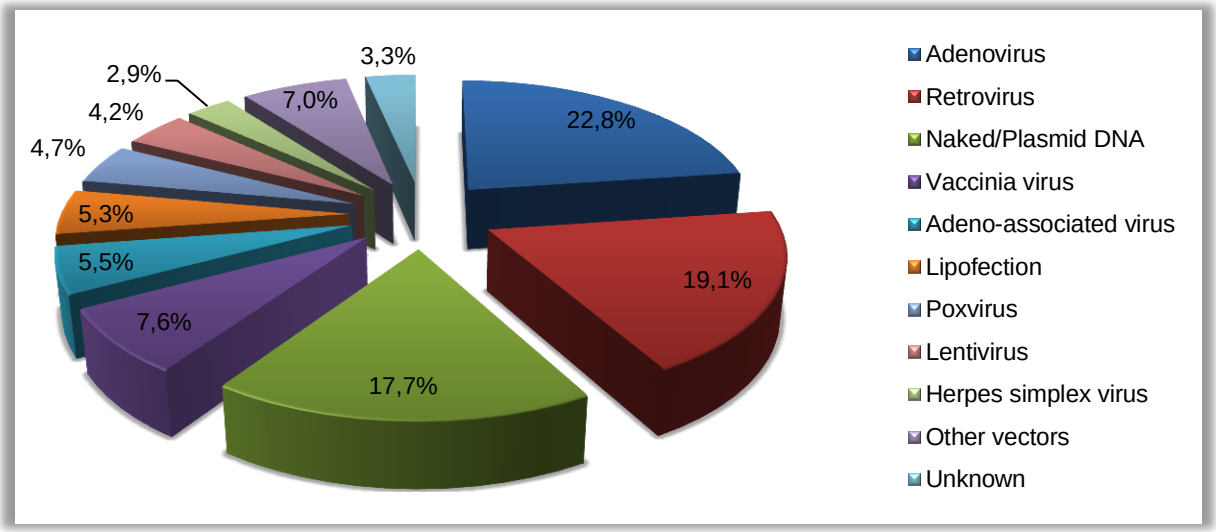

Figure 1.2: Vectors used in gene therapy clinical trials, as reported in The Journal of Gene Medicine. Percentages refer to a total of 2076 clinical trials approved worldwide between 1989 and 2014 [4].

1.1.1.1 Viral vectors

Typically, viral vectors used in gene therapy are made replicate-deficient by deletion of sequences related to virus replication, rendering them safe but dependent on a packaging cell line (PCL) for the production of recombinant viruses. Other genes responsible for pathogenicity and expression of immunogenic viral antigens can also be deleted from the virus genome. The packaging cell line is

engineered in order to express the genes deleted from the virus and thus enable the synthesis of viral components and production of the recombinant virus [12].

Viral vectors are used in gene therapy due to the high efficiency of viruses in introducing DNA into cells. This characteristic is advantageous for gene therapy since it allows efficient transduction of the gene of interest by recombinant viruses. A wide range of virus, such as adenoviruses, adeno-associated viruses, retroviruses, lentiviruses and herpesviruses, among others, have been used in this context. The selection of the virus for the desired application is dependent on the nature of the virus itself that can be divided in two classes – non-lytic and lytic – according to their survival and replication strategies [12]. The class nomenclature is related with maintenance of cell integrity after infection. In the case of infection with a non-lytic virus, such as retroviruses and lentiviruses, after cell infection the virus produces virions from the cellular membrane of the infected cell without compromising its integrity. This behavior is not observed with lytic viruses (e.g. adenovirus and herpes simplex virus) that promote cell lysis after viral replication and virion production. Other properties that vary among viruses, and must be taken into account for selection, include the size permitted for gene insert, duration of gene expression after transduction and ability of the vector to infect the cells and integrate into its genome [12, 13].

Clinical trials using these vectors are currently being conducted in all four trial phases. Up to date, the only vector reported on the Journal of Gene Medicine [4] to reach phase IV of clinical trial – the phase related with post-marketing of a treatment approved by FDA [14] – was Gendicine, an adenovirus coding for tumor protein p53 that is being applied as a tumor suppressor in thyroid cancer. This product was developed by SiBiono GeneTech Co. of Shenzhen, China, and reached the market in 2004 [4, 11]. Another gene therapy product currently authorized by EMA for marketing purposes in Europe, since October 2012, is Glybera [15]. This product was developed by uniQure (Amsterdam, the Netherlands) and consists of an engineered copy of the human LPL gene with improved enzyme activity packaged with a tissue-specific promoter in a non-replicating adeno-associated type 1 virus vector (AAV1 vector). This gene-based medicine was developed to treat adult patients diagnosed with lipoprotein lipase deficiency, a severe life-threatening complication characterized by attacks of pancreatitis [15, 16]. uniQure is also currently involved in the development of other gene based therapies, resorting to adeno-associated viruses, for diseases such as heamofilia B, Acute Intermittent Porphyria (AIP), Sanfilippo B Syndrome, Parkinson's disease and congestive heart failure [16].

1.1.1.2 Non-viral vectors

Comparatively to their viral counterparts, non-viral vectors present several advantages, especially in safety aspects. The safest and simplest method of non-viral transfection is the direct injection of naked plasmid DNA into tissues (e.g. liver, skeletal and heart muscles) or to a vessel of the systemic circuit [17]. A report by Wolff and co-workers in 1990 of uptake and extended *in vivo* expression, for at least 60 days, of transgenes inserted on a plasmid and directly injected into the leg muscle of mice generated high interest and suggested the potential therapeutic application of plasmid DNA in gene therapy [8, 18]. Since this demonstration, intramuscular injection of plasmid DNA for gene therapy and design of DNA vaccines

has been a growing field of research [18]. The transfection efficiency of this type of delivery is, however, conditioned by the degradation of the DNA vector by nucleases present in the physiological fluids and extracellular space, as well as by recognition of these vectors by cells of the immune system. In order to improve efficiency, several physical methods (e.g. electroporation, gene gun, etc.) and synthetic delivery vectors have been developed [17, 19]. Typically, from these two approaches, the most used in human gene therapy is the entrapment of the non-viral vector on a synthetic delivery vector such as inorganic particles, liposomes and polymeric complexes [19].

To be used in a therapeutic protocol, the plasmid DNA vector has to be carefully designed, avoiding sequences that can induce a toxic effect or difficult the production process, while complying to established prerequisites. These prerequisites indicate that the plasmid backbone should present the appropriate promoter and a transcription terminator, elements for optimized mRNA processing and translation (e.g. Kozak consensus and polyadenylation signal), prokaryotic origin of replication to enable propagation on bacteria host and a selection marker (usually genes coding for antibiotic resistance) for vector amplification in bacteria. Regarding the selection marker, further impositions are enforced by regulatory agencies, mainly because these genes may lead to inflammatory reactions *in vivo* and can be transmitted to endogenous enteric bacteria [20, 21]. The only selection marker allowed by regulatory agencies in therapeutic DNA products is the gene for resistance to kanamycin (KanR) [20]. However, to improve vector safety, the deletion of antibiotic resistance genes is advised [20, 22].

The therapeutic success of the vector is dependent on a high transfection efficiency. The size of the vector is directly related with the accomplishment of this objective and should be kept at a minimum. Typically, without taking into account the size of the therapeutic gene of interest, the size of plasmid vectors vary in a range from 4 to 12 kb [20]. Studies have been conducted to evaluate the effect of plasmid size on transfection efficiency and resulted in the realization that transgene expression is inversely proportional to vector size [23]. Transient transfection studies conducted by Yin *et al.* [24] have demonstrated that larger plasmid vectors can negatively influence promoter/enhancer activity. These experiments involved the generation of several constructs by insertion of stuffer DNA fragments ranging from 0.65 to 5 kb. An inverse relation between size and promoter/enhancer activity was obtained that was attributed to probably be due to the lower transfection efficiency of larger vectors [24]. This observation can be explained by a better ability of smaller vectors to overpass the barriers to enter the cells and reach the nucleus [20].

Deletion of non-essential sequences is also important to increase vector stability and minimize integration events as the clinical use of vectors presenting prokaryotic origin of replication and resistance markers can lead to the dissemination of prokaryotic recombinant DNA in patients. Sequences with un-methylated CpG residues are commonly found in prokaryotic DNA and can be useful for vaccination purposes by activation of the immune system due to their recognition by Toll-like receptor TLR9. The injection of plasmid DNA containing selected genes from pathogens has been demonstrated to elicit protective immune responses and has showed possible application against malaria and AIDS leading to the study of DNA vaccination for other infectious diseases such as hepatitis B and C and tuberculosis [18].

However, when the vector is intended for gene therapy purposes, the stimulation of the immune system is not beneficial and thus deletion of prokaryotic sequences would result in a safer vector [20, 21, 25]. Despite the fact that the vector will not integrate into the genome of the cells subjected to gene therapy, the plasmid can be maintained episomally and transcribed for an extended period of time, until it will eventually be eliminated. This renders plasmid DNA based therapies very interesting when transient transgene expression is desired. For example, the use of plasmids encoding the VEGF (vascular endothelial growth factor) gene has shown positive effects on patients, with revascularization of ischemic limbs and hearts of subjects receiving the treatment [18].

Structural conformation is another important characteristic of DNA vectors that has influence on transfection efficiency [20]. Three different plasmid DNA (pDNA) conformations have been studied to determine the most physiologically active conformation for gene delivery – linearized plasmids, relaxed open circular (oc) plasmids and compact supercoiled circular (sc) plasmids. Remaut *et al*. [26] have compared these conformations in transfection studies, concluding that there is no difference in transfection efficiency between these forms when they are delivered directly to the nucleus by microinjection at similar intranuclear concentrations. However, this observation was not reported when using pDNA-containing lipoplexes or microinjection of naked pDNA in the cytosol. In both cases, the sc pDNA resulted in higher transfection efficiency thus leading to the conclusion that sc/lipoplexes could enter cells more easily or present a better release of pDNA from the lipoplexes into the cytoplasm and that naked sc pDNA form is the one presenting a better ability to reach and enter the cell nucleus. The results obtained with naked sc pDNA in the cytosol could be explained by a better stability of this form as well as by a faster diffusion in the cytosol and easier entry into the cell nucleus [26]. The fact that sc DNA is the most physiologically active form is also recognized by regulatory agencies. Guidance documents from FDA indicate that therapeutic pDNA vectors to be used in phase I clinical studies should present a minimum homogeneity of 80% sc DNA [27].

As a consequence of the concerns regarding the safety of non-viral vectors and interest in improving their transfection efficiency, a new generation of non-viral vectors has been developed. These vectors are composed of sequences partially or totally devoid of prokaryotic elements. In vectors that do not present any prokaryotic sequence, both the prokaryotic origin of replication and antibiotic resistance genes have been eliminated. A few mini DNA vectors bearing these characteristic have been designed, including minicircles, MiLV (micro-linear vectors) and MIDGE (minimalistic, immunological defined gene expression) vectors. These latter two systems are minimal-size linear vectors composed of an eukaryotic expression cassette and polyadenylation signal, flanked by two short hairpin oligonucleotide (ODN) sequences, which differ only in the method by which they are obtained [25]. In the case of MiLV vectors, the production is conducted *in vitro* resorting to digestion of the starting plasmid, ligation of the fragment of interest with ODN hairpins and amplification of the vector by polymerase chain reaction (PCR). These vectors were reported by Wang *et al*. [28] as an attempt to overcome the costly and time consuming production of MIDGE vectors [28]. For production of MIDGE vectors, the fragment of interest is obtained by *in vitro* digestion of starting plasmids, containing the expression cassette, with a restriction

endonuclease and by ligation of the ends of the fragment with ODN hairpins [25]. A recent phase I clinical trial has been approved by FDA in 2013, using a MIDGE vector harboring a tumor necrosis factor for gene transfer in skin metastases of melanoma [4].

A new minimalistic vector – DNA ministring – has recently been proposed by Nafissi *et al*. [29]. This vector is, similarly to MiLV and MIDGE, a mini linear covalently closed DNA vector but is intended to be produced in one step *in vivo*, relying in a temperature-regulated recombinase system for linearization of parental plasmid [29, 30].

1.2 Minicircles

A minicircle (MC) is a non-viral vector and is defined as a supercoiled double-stranded DNA molecule, composed only by the therapeutic eukaryotic expression cassette [21].

The studies that led to the formulation of the minicircle concept were performed in the early 1990s and were related with an attempt to understand the molecular mechanisms behind the conversion of single-stranded recombinant adeno-associated viral (rAAV) vector genomes to double-stranded DNA (dsDNA) once inside tissue cells. The fact that these vectors were converted to double-stranded monomeric and concatemeric episomes, which presented a more stable life-long transgene expression than the one obtained with plasmid DNA hosting the same eukaryotic expression cassette, was the result that contributed to further studies. Based on this realization, Kay and co-workers investigated the conversion of purified linear dsDNA containing only the eukaryotic expression cassette and rAAV proviral vectors containing the same expression cassette, concluding that both were converted to circular and concatemeric episomes in tissues and presented a similar robust and persistent transgene expression [31].

To overcome the need of isolation of DNA fragments from gels and *in vitro* ligation for the construction of minicircles, a new strategy was developed that resorted to the utilization of a prokaryotic inducible recombinase system. Although the production of minicircles was still more labor intensive than plasmid production, the amount produced allowed the study of these new vectors in small and large animals from which it was possible to realize their usefulness leading to future attempts in improving the production [31].

Even though both the prokaryotic origin of replication and selection marker are still needed for vector propagation on bacterial host, these sequences are then excised during the recombination event. Due to the absence of prokaryotic sequences and, consequently, smaller size when compared to conventional plasmid vectors, minicircles present higher biosafety, transfection efficiency and bioavailability. These characteristics render minicircle vectors as a safe and promising alternative to plasmid DNA in gene or cell therapy applications and in DNA vaccination [20, 21, 32].

1.2.1 Minicircle production

Typically, the production of minicircles relies in the production *in vivo* of a parental plasmid on a bacteria host in which a recombination event is induced. Although an *in vitro* recombination strategy could also be possible by using a recombinase or a restriction digestion followed by ligation, these methods would be difficult to scale-up for industrial application [21].

In prokaryotic organisms the recombination events occur either by general recombination between homologous DNA sequences or by specialized recombination events that can be grouped as transposition processes or conservative site-specific recombination. In either case, the formation of a complex between a recombinase and, at least, two DNA sites is required. These mechanisms are responsible for the rearrangement of genetic information and can result in the insertion, deletion, duplication, inversion or translocation of DNA sequences. The occurrence of these rearrangements can be related to specific physiologic functions or programmed processes, but can also be a spontaneous process and contribute to the genetic diversity of a population [33].

The production of minicircles relies in a site-specific recombination strategy [21]. In this type of recombination, a breakage of four DNA strands occurs at specific positions of two separate recombination sites. These four strands are then exchanged and resealed, without involving DNA synthesis or degradation, resulting in either one or two DNA molecules, depending in the intervening number of molecules and relative orientation of the recombination sites. If the recombination process occurs between two specific sites existing in two DNA molecules, the recombination event is termed as intermolecular and results in the integration of both DNA sequences in one molecule (as exemplified by arrow 3 in Figure 1.3). On the other hand, if the specific sites are on the same DNA molecule, the recombination is termed intramolecular and can result in an inversion or excision of the DNA segments depending on the relative orientation of the recombination sites. A molecule exhibiting specific recombination sites in inverted orientations will lead to the formation of one molecule where one of the sequences between the specific sites is in an inverted orientation relatively to the original molecule (as exemplified by arrows 1 in Figure 1.3). If on the contrary, the specific sites of the molecule that suffers recombination are on the same direction, the recombination event will result in the excision of a DNA segment and formation of two DNA molecules, each consisting in a specific site and the DNA sequence that was flanked by the two specific sites in the original molecule (as exemplified by arrow 2 in Figure 1.3) [33].

Figure 1.3: Outcomes from site-specific recombination. Recombination sites on DNA sequence are in dark blue and its orientation is indicated with green arrows. Position of distinct genetic markers around recombination loci is indicated by letters A and B (adapted from Hallet *et al*. [33]).

The last type of recombination described above is the one applied for the production of minicircles (MC). These vectors can thus be described as double-stranded DNA molecules, comprised only by an eukaryotic expression cassette, that are produced *in vivo* on a bacteria host, upon induction of intramolecular site-specific recombination of a parental plasmid (PP) that will lead also to the formation of a molecule containing all the prokaryotic backbone – the miniplasmid (MP). Comparatively to an *in vitro* production, this strategy is simpler requiring only the induction of the production of the recombinase [21]. Several recombinase systems, summarized in Table 1.1, have been used to induce the production of minicircles, being, in all cases, derived from two enzyme families – tyrosine and serine recombinases [34].

Table 1.1 – Characteristics of the recombination systems used for the production of minicircles (adapted from Schleff *et al*. [34]).

Recombinase family	Recombination system	System origin	Recognition sequence	Type of recombination	Inducer	Ref.
Tyrosine	λ-integrase	Bacteriophage λ	*att*B and *att*P sites	unidirectional	heat	[21]
	Cre-recombinase	Bacteriophage P1	lox sites	bidirectional and fully reversible	L-arabinose	[35]
	FLP-recombinase	Yeast plasmid 2 µm circle	FRT sites	bidirectional and fully reversible	heat	[36]
Serine	ϕC31-integrase	*Streptomyces* bacteriophage ϕC31	*att*B and *att*P sites	unidirectional	L-arabinose	[37]
	ParA resolvase	Plasmid RK2 or RP4	res sites	unidirectional	L-arabinose	[32, 38, 39]

The pioneer work on minicircle development and production was performed in the late 1990s by Darquet and co-workers [21]. In this study, researchers reported the use of an *in vivo* minicircle production strategy relying in a site-specific recombination, mediated by the λ-integrase, on D1210HP *E. coli* strain. With this purpose, the parental plasmid (7.4 kb) was constructed with two recognition sites for the recombinase enzyme flanking the eukaryotic cassette. These recognition sequences were the *att*P (530 bp), from the phage λ genome, and *att*B (31 bp), from the *E. coli* genome. These sites were placed in the same direction on the parental plasmid in order to, as described above, promote the excision of DNA from this molecule and thus give origin to two DNA molecules – a minicircle comprised by the eukaryotic cassette with 3.4 kb and a miniplasmid with the 4 kb bacterial backbone. With this system, for the recombination event to occur, besides the integrase the presence of two host proteins (FIS and IHF) and a thermal shift to 42°C are also needed. Upon thermal induction of the bacterial strain harboring both a λ thermosensitive lysogen (defective for lethal and lytic functions) and the parental plasmid, the λ-integrase is produced and mediates the recombination. The lysogen used was Xis^-, meaning that the Xis proteins, required for recombination between *att*L and *att*R sites in the products obtained from the recombination event, were not present after induction [21, 40]. This strategy was successful for minicircle production

(≈0.13 mg/L of bacterial growth) and the transfection efficiency studies performed with minicircles showed a 2 to 10-fold increase when compared to parental plasmid. Nevertheless, the product obtained after induction presented 40% of un-recombined parental plasmid and about 30% of the minicircle produced were in a multimeric form, meaning that an improvement of the recombination was needed [21, 41].

An attempt to improve this recombination system was reported just a year after, by Kreiss and co-workers [41]. In order to resolve the minicircle multimeric forms, this team of researchers resorted to the introduction of a slightly modified *parABCDE* locus from plasmid RK2 into the minicircle sequence, between the *att*P and *att*B sites, successfully reducing the multimeric minicircles to less than 2 % of all minicircle produced [41]. The optimization of the culture conditions of the same bacterial strain used by Darquet *et al.* [21] (D1210HP), namely the study of the kinetics of integrase-mediated recombination, led to an improvement in the minicircle production to 1.5 mg/L of bacterial growth [41].

Although some recombinases of the tyrosine family will naturally lead to a bidirectional and reversible recombination (Table 1.1), mutations on the recognition sequences have been attempted that can avoid these phenomena, leading to occurrence of the desired unidirectional recombination and thus avoiding the loss of minicircle after formation. An example of this strategy is the mutation on lox sites, recognized by cre-recombinase, to induce a shift in the reaction equilibrium of *cre*/*loxP* reaction toward increased production of minicircle. This attempt was performed by Bigger and co-workers [35] and was executed by modification of the terminal 5 nucleotides on lox sites resulting in an improvement of the generation of monomeric minicircle. However, high amount of concatamerization was still observed [34]. The bacterial strain used was engineered to express the *cre*-recombinase under the tight control of the *araC* regulon. The rationale behind this modification was related with the need to avoid a premature Cre recombination that would result in a loss of minicircle during cellular growth due to its non-replicative attribute. During recombination two molecules are formed from each parental plasmid – a minicircle and a miniplasmid. The miniplasmid, presenting all bacterial backbone, is the only replicative and antibiotic resistant form (besides residual parental plasmid that has not undergone recombination) and, thus will have a competitive advantage over the minicircle molecule. The *cre*-expressing bacterial strain developed showed high stability and was easily controlled by alteration of the carbon source available for bacterial metabolism. The better results were obtained with a control based on the use of 0.5% glucose on an overnight growth at 37°C, to inhibit plasmid recombination, followed by washing and addition of 0.5% arabinose to induce the recombination event during a cellular growth performed for 4-6 h, also at 37°C. From the results obtained for the smaller minicircle studied (3.4 kb), these researchers concluded that the engineered strain produced a higher amount of minicircle (reaching yields of 300 µg/L of culture) when compared to the original strain that produced only over 200 µg of minicircle per litter of culture. Transfection assays were also performed that demonstrated higher transfection efficiency for the minicircles studied (3.4 kb) in comparison to their plasmid counterparts (6.4 kb). In fact, when transfected on a mole:mole basis (resorting to the addition of stuffer DNA in the minicircles), the minicircle vectors demonstrated a 4.5-fold increase on expression of reporter gene. The values reported were even higher when using a weight:weight basis (8.8-fold increase). When using a mole:mole basis, without the addition

of stuffer DNA on the minicircles, an increase of 152-fold on reporter gene expression was observed, highlighting the reduced cytotoxic effect of these molecules as well as their higher transfection efficiency [35].

The problems related to the production of minicircle multimers were, in great measure, solved by the use of recombinase enzymes from the serine family, as the ϕC31-integrase and ParA resolvase [20, 34]. The ϕC31-integrase, derived from *Streptomyces* bacteriophage ϕC31, was used by Chen and co-workers [37] to mediate the intramolecular recombination of parental plasmids, with *att*B (34 bp) and *att*P (39 bp) sites, produced by the *E. coli* Top10 strain. The ϕC31-integrase is a recombinase from the resolvase/invertase family and the production of minicircles resorting to its recombination system is, in theory, more efficient when compared to other recombinases such as *Cre* recombinase, due to its involvement in the mediation of a unidirectional recombination. This recombinase was placed on the plasmid under the control of the L-arabinose inducible pBAD/araC system [37]. In this system, the *araC* gene has a regulatory function, activating or repressing the transcription from BAD promoter. In the presence of arabinose, the araC protein will initiate the transcription from the BAD promoter while in the absence of arabinose this transcription is repressed, even though still occurring at very low levels. The leaky expression in the absence of arabinose can be further repressed by the presence of glucose in the medium that will reduce the levels of 3'-5'-cyclic AMP, thus lowering expression of the BAD promoter [42, 43].

In their studies, Chen and co-workers [37] constructed two plasmids with different eukaryotic cassettes – one coding for α1-antitrypsin (human AAT) and other coding for human factor IX (human FIX) – and realized that, from the growth and recombination strategies applied, the one which led to higher recombination efficiency consisted in an overnight bacterial growth at 32°C in absence of L-arabinose, followed by resuspension of the cells in fresh medium supplemented with L-arabinose and incubation at the same temperature for 60 to 120 minutes. Employing this strategy, the yield of minicircle obtained after purification was 150-200 µg/L of bacterial growth with recombination efficiency over 97% [37]. The transgene expression of the minicircles and the un-recombined plasmids was also compared. From the values obtained, the researchers concluded that the two minicircles studied presented a 45 and 560-fold more transgene expression in mice liver than their parental plasmid counterparts, depending if the vector presented the gene for human AAT or human FIX, respectively [37].

Comparatively to the recombination strategies presented above that were only able to achieve recombination efficiencies lower than 97%, the ParA resolvase system was the only to successfully achieve a complete recombination of the parental plasmid [38, 39]. This resolvase is also from the serine family and is expressed under the transcription control of the pBAD/araC system independently of host factors besides topoisomerases, which are ubiquitary enzymes for decatenation [34, 38].

In 2004, Jechlinger and co-workers [38] successfully produced minicircles using this recombination system, reporting a 100% recombination efficiency after only 30 minutes of induction of recombinase expression, when the ParA resolvase was encoded on the plasmid harboring the resolution sites and thus worked independently of the plasmid copy number. Further studies regarding the ParA

placement on a helper plasmid with a lower copy number than the plasmid containing the resolution sites, led to a recombination efficiency of approximately 50%, indicating that a high ratio of ParA resolvase to plasmid is needed to achieve complete recombination [38].

This recombination system was also used by Mayrhofer and co-workers [39] for the development of a scalable minicircle production system, resorting to the addition of glucose to the medium to prevent premature expression of the ParA resolvase. The strategy relied in a pre-culture in LB-medium at 28°C for 15 hours, under selective conditions and in the presence of glucose. This culture was then used to seed a bioreactor with 5 L of working volume, with controlled operational variables at 37°C, in which the recombination was induced by addition of L-arabinose at OD_{600nm} between 3.5 and 5.0. A recombination efficiency of 99.57% was obtained after 1 hour of induction [39].

The tight control of repression and expression of the site-specific recombinase in *E. coli* is crucial to achieve an efficient recombination. To accomplish this goal, the expression system should present an efficient silencing of gene expression before the induction of recombination, in order to avoid a premature recombination and thus allow the obtainment of a maximum minicircle yield. From the recombinases studied, the ParA resolvase has shown the best recombination efficiency, being capable of completely converting the parental plasmid into the two desired products – miniplasmid and minicircle – without the formation of unwanted multimeric forms [34, 39].

1.2.2 Minicircle applications

Minicircle DNA vectors are considered safer and more efficient delivery vectors for gene therapy applications and DNA vaccination when compared to plasmid DNA, being associated with higher transfection efficiency [21]. Several studies have been performed in order to evaluate the efficacy of this new technology for the treatment of several diseases by using minicircles encoding different proteins, such as human endothelial growth factor for stimulation of proliferation of endothelial cells in the treatment of peripheral arterial disease [44], ornithine trascarbamylase modified for mitochondrial translation [35], human factor IX and α1-antitrypsin [37] for treatment of haemophilia and deficiencies on this enzyme, respectively. Another interesting possible application of these vectors is for generation of induced pluripotent stem cells with reprogramming minicircles containing a defined set of transcription factors [45].

1.2.3 Minicircle purification

After successful recombination, the subsequent step is the isolation of the form of interest - the therapeutically valuable minicircle – from the DNA mixture obtained during the bacterial growth and other unwanted side products. The DNA mixture consists in minicircle, miniplasmid and, eventually, residual parental plasmid that did not undergo recombination [34, 46]. The presence of these forms hinders the purification process due to the high similarity in terms of physicochemical properties like charge or hydrophobicity. The isolation of minicircles from miniplasmids is further hampered by their almost equal size. Thus, the inherent characteristics of these three DNA molecules render the application of

conventional chromatographic methods such as anion exchange, hydrophobic interaction or size exclusion, highly inefficient for minicircle purification purposes [46].

The first attempts at minicircle purification relied in the use of restriction enzymes *in vitro*, selected to linearize the miniplasmid and parental plasmid, while having no enzymatic activity over the minicircle. The circular minicircle was then isolated by caesium chloride density gradient centrifugation, using the intercalating agent ethidium bromide (CsCl-EtBr) [21, 35, 36] or by agarose gel electrophoresis [36]. The fact that these methods are neither suitable for the purification of clinical grade material or large-scale production and are labor intensive and expensive techniques that result in the achievement of low yields of purified minicircle highlighted the need for the development of better purification strategies [21, 39, 46].

Several approaches have been proposed to obtain pure minicircles, namely *in vivo* restriction and affinity based chromatography [46]. For therapeutic use, the purified minicircles should comply with specific guidelines. The pharmaceutical application of DNA products relies on analysis and quality control of the product which are regulated by several regulatory recommendations (Table 1.2). These tests should be performed, before the initiation of Phase I clinical trials, to analyze several parameters such as general safety, purity and identity of the DNA products [27].

Table 1.2 – Recommended assays and specifications for control and release testing of plasmid DNA products (adapted from [27, 47, 48]).

	Parameter	Method	Suggested specifications
Impurities	**Purity**	A260/A280=1.8-1.95	Absorbance at 260 nm and 280 nm
	Proteins	BCA (bicinchoninic acid) assay, SDS-PAGE (SDS-polyacrylamide gel electrophoresis)	<1% [27]
	RNA	Agarose gel electrophoresis, Northern blot	<1% [27]
	gDNA	Hybridization blots, real-time PCR, fluorescence	<1% [27]
	Endotoxins	LAL (*Limulus* Ameobocyte lysate) assay	40 EU/mg plasmid [27]
	Microorganisms	Bioburden test and/or sterility test	<1 CFU [47]
Plasmid	**Identity**	Restriction enzyme mapping	Conformance to plasmid DNA map
	Homogeneity	Agarose gel electrophoresis, anion exchange-HPLC	>80% supercoiled form [27]
	Potency	Cell transfection	According to application [48]

1.2.3.1 *In vivo* restriction

The use of *in vivo* restriction as a strategy for minicircle purification was first proposed by Chen and co-workers [46, 49]. With this purpose, the parental plasmid constructed harbored the means for the subsequent purification of the minicircle DNA – a homing nuclease, namely I-SecI, and the respective recognition sequence localized in the bacterial backbone (outside the *att*B and *attP* sites) so that, after recombination, this recognition sequence would not be present in the minicircle. The plasmid was constructed by placement of the genes for the ϕC31 recombinase and I-SecI under the control of the

P_{BAB}/*araC* arabinose promoter. The selection of I-SecI was based on the fact that the *E.* coli genome does not present the 18 bp sequence necessary for recognition by this enzyme. The principle of the strategy was that the addition of L-arabinose would induce both the recombination of the parental plasmid and expression of the homing nuclease I-SecI. This would lead to formation of the minicircle and miniplasmid as recombination products and to expression of I-SecI, thus ensuring that the only circular DNA in the bacteria would be the minicircle, while the parental plasmid and miniplasmid were linearized by the I-SecI expressed and, subsequently, degraded by bacterial exonucleases, allowing the isolation of the minicircle by affinity chromatography. Although the formation of minicircle and almost complete destruction of the plasmid backbone were observed, this methodology involved the simultaneous production of I-Sec and ϕC31 recombinase and thus could result in a decrease in the yield of minicircle due to a premature destruction of parental plasmid. To overcome this problem, several alternatives to the first protocol were tested: (1) insertion of an additional copy of the recombinase gene to accelerate the recombination process, (2) incubation of the bacteria at the optimal temperature for ϕC31 recombinase expression (32°C) followed by an increase to 37°C to promote the endonuclease activity only after the recombination had occurred and (3) initial cultivation of the bacteria in the presence of L-arabinose at 32°C followed by addition of fresh LB, supplemented with 1% L-arabinose, at pH 8 and increase of incubation temperature to 37°C to favor I-SecI activity. This third approach was found to be the one that minimized loss of un-recombined plasmid, leading to a yield of 1-1.8 mg of minicircle per litter of overnight bacterial growth with a purity superior to 95% [49].

The concept behind this strategy was further studied by Kay and co-workers [50]. This work involved the use of a different bacterial strain (*E. coli* BW27783) with the constitutive promoter cp8 that was engineered to not express endonuclease A (in order avoid plasmid DNA degradation) by knockout of the gene and to express a second L-arabinose transporter. Relatively to the work developed by Chen *et al*. [49], this method presented also a difference in the location of the genes coding for ϕC31-integrase and I-SecI enzyme that were both relocated from the parental plasmid to the bacterial genome. The engineered strain was further improved, resulting in the strain ZYCY10P3S2T, whose genome includes ten copies of the cassette composed by the BAD promoter and ϕC31-integrase gene and three copies of the same promoter and I-SecI gene. The number of recognition sites for I-SecI on parental plasmid was also changed to 32 tandem copies of the I-SecI site in the plasmid backbone. After induction of recombination mediated by the addition of L-arabinose, the production of minicircle using this strategy resulted in a higher yield of minicircle produced (3.4-4.8 mg/L of overnight culture) with very low content of contaminant un-recombined plasmid and miniplasmid, ranging from 0.4% to 1.5% [50]. The robustness presented by this system led to the development of the MC-Easy™ Minicircle Production Kit, a commercial purification kit by SystemBiosciences, that includes medium for cell growth and recombination induction and optionally can include the ZYCY10P3S2T *E. coli* minicircle producer strain which can afterwards be transformed with the desired parental plasmid harboring the minicircle sequence [51].

1.2.3.2 Affinity-based chromatography

Although the application of an affinity based chromatography for the purification of minicircles has been hindered by inefficient recombination and the lack of a high performance affinity chromatography system for the purification of DNA, this is considered to be the method of choice for large-scale minicircle purification, as long as the sample to be purified does not contain un-recombined parental plasmid [46]. The absence of this form is required due to the fact that, in this type of separation, a short recognition sequence is placed on the sequence of the parental plasmid that will later be excised to give origin to the minicircle. Therefore, if the recombination achieved is not complete, the un-recombined parental plasmid will also harbor the recognition sequence with affinity for the separation matrix and will co-purify with the minicircle [39, 46].

The first affinity-based chromatography method for minicircle purification – the 'Recombination Based Plasmid Separation Technology' (RBPS-Technology) – was reported by Mayrhofer and co-workers, in 2008 [39]. This purification strategy is a two-stage process, relying in the high recombination efficiency of the ParA resolvase system to obtain a mixture composed mostly of recombination products. Minicircle in this mixture is purified later using a method based on a protein/DNA interaction (Figure 1.4) that exploits the affinity between *lacOs* sites (direct repeat of modified lactose operator sites) placed on the minicircle and the LacI (repressor of the lactose operon) immobilized on a chromatography matrix [39].

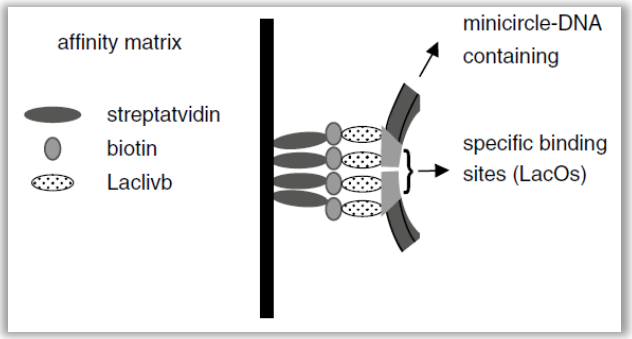

Figure 1.4: Representation of the RBPS-Chromatography matrix consisting in the interaction between a tandem repeat of a symmetric version of the lactose operator sites (LacOs) on the minicircle DNA and the repressor protein of the lactose operon (LacI) attached to the chromatography matrix. An *in vivo* biotinylation (ivb) sequence is C-terminally fused to the LacI protein (LacIivb). The LacIivb is immobilized on the matrix via biotin/streptavidin interaction (adapted from [39, 46]).

For the development of the chromatography matrix of the RBPS-Technology, a biotinylated version of the lactose repressor (LacIivb) was constructed by *in vivo* C-terminal fusion of a biotinylation sequence (ivb) to the LacI protein. The LacIivb fusion protein was expressed in cells cultured in the presence of biotin leading to a post translational modification reaction in which the biotin was covalently attached to LacIivb. After purification, the protein was applied to Streptavidin Sepharose High Performance (a resin carrying covalently bound tetrameric streptavidin) resulting in the immobilization of LacIivb via biotin/streptavidin interaction. The large scale purification of minicircles, harboring the LacOs sites, was carried out using 5 mL of this matrix. The recombination products were applied in a buffer

containing 400 mM NaCl. The salt concentration selected to be used in buffer for sample application was related with the capture of unspecific pDNA without LacOs sites at low salt concentrations. The use of a buffer with 400 mM NaCl avoided unspecific binding of pDNA lacking the lacOs sites without interfering with the capture of the minicircle DNA. The captured DNA was then eluted by a buffer containing 500 mM NaCl and IPTG (that also occupies the DNA binding sites and thus reduces the matrix capacity for specific binders). Q-PCR analysis revealed that resulting purified samples presented 98.8% minicircle DNA, 1.0% miniplasmid and 0.2% parental plasmid. Further quality control assays confirmed also that the purified minicircles complied with the recommendations of regulatory agencies. The capacity of the RBPS affinity matrix was also tested for different matrix batches, being obtained an average capacity of 184±42 µg of minicircle per mL of matrix. Tests conducted with samples of different concentrations led to the conclusion that the matrix binding capacity was directly related with this variable, being higher for more concentrated samples [39].

1.2.3.3 Monolith anion exchange chromatography

Very recently, in 2014, a method for minicircle purification relying in anion exchange chromatography (AEC) was proposed and patented [52].

In anion-exchange chromatography the purification of pDNA from gDNA and RNA relies in the interaction between these species and positively charged ligands such as tertiary or quaternary amines in the stationary phase. In this context, column loading is typically conducted at an ionic strength high enough to avoid binding of impurities such as proteins, oligonucleotides and low-molecular-weight RNA that have a lower charge density when compared with pDNA. Elution of bounded species is then conducted by increase of ionic strength of the mobile phase to selectively promote the decrease on electrostatic interaction between pDNA, high-molecular-weight RNA and gDNA and the stationary phase. The strength of electrostatic interaction between nucleic acids and an anion exchanger matrix is proportional to the number of phosphate groups in the molecule, and thus to the molecule overall net charge. This means that the binding of larger molecules is stronger and their elution time from the column will be higher than the one presented by nucleic acids of smaller sizes [48, 52]. The spatial conformation presented by pDNA is another aspect affecting the interaction. Although supercoiling can conceal parts of the pDNA molecule, leading to a decrease in the number of phosphate groups available to interact with the stationary fact, this same event can at the same time decrease the molecule size increasing its charge density. This is possibly one explanation for the separation of pDNA isoforms using this type of chromatography, being noticed that sc pDNA presents a stronger binding to the AEC support than open circular isoforms [48].

The matrix used for AEC presented diethylaminoethyl (DEAE) ligands, which are weak anion exchanger ligands that separate nucleic acids according to their overall net charge [52]. The monolith matrix is a highly porous solid material with large-sized pores, ranging between 600-750 nm, which enhances mass transport by convection and allows high flow-rates, when compared to conventional bead-based chromatographic supports where the mass transfer occurs mainly by slow diffusion. As a

consequence of the high flow rates and low backpressures, separations resorting to monoliths are extremely fast. Typically, these supports are characterized by high binding capacity, recovery and resolution power, independently of the flow-rate, and are of easy scale-up [52].

The proposed method relies on efficient recombination, mediated by ParA resolvase, of a parental plasmid containing several genetically-engineered restriction sites placed on the bacterial backbone. These restriction sites are directly related with the purification process, being used to linearize the miniplasmid *in vitro*, after induction of recombination during the bacterial growth, cell lysis, conditioning and pre-purification, reducing their size and, consequently, their charge and strength of interaction with the anion exchanger matrix. Pre-purified samples were thus subjected to enzymatic restriction, to linearize the miniplasmid and un-recombined parental plasmid, and applied into a commercial 0.34 mL anion exchanger monolith disk in order to separate the supercoiled minicircle from the contaminant species. The best separation was obtained when performing a step elution with buffers at pH 9 due to the fact that, at this pH, fewer groups on the chromatographic matrix were charged and thus less ligand sites were involved in the binding of the negatively charged DNA molecules that, in these conditions, could be eluted at lower salt concentrations (230-330 mM NaCl, depending on the samples to be purified). The application of this methodology to the separation of minicircles from linear fragments was found to be robust and of possible scale-up, thus being applicable to the production of minicircles in the order of grams to kilograms required for pre-clinical and clinical trials [52].

1.3 Purification strategies

The growing interest in plasmid DNA revealed by the fields of DNA vaccination and gene therapy and the possible application of these technologies for therapeutic purposes has been translated in the requirement of improvement fermentation, processing, purification and formulation strategies that should allow a generic and cost-effective production of plasmid DNA at large scale [53]. In order to be used for therapeutic purposes pDNA has to underdo purification processes that efficiently remove the contaminants content and allow the obtainment of a product complying with the quality specifications recommended by regulatory agencies (addressed above in Table 1.2), such as Food and Drug Administration (FDA) and European Medicines Agency (EMA) [54].

1.3.1 Primary purification and intermediate recovery

Minicircles are molecules that are produced *in vivo* upon induction of recombination of a parental plasmid on *E. coli* [46]. For plasmid DNA recovery, the cells have to be disrupted in order to break the three-layered envelope (plasmid membrane, peptidoglycan cell wall and outer membrane) of the gram-negative *E. coli*. In this process is obtained a solution comprised of cell debris, host proteins, endotoxins (such as lipopolysaccharides), genomic DNA (gDNA), RNA and pDNA, among others. Although the exact composition of the solution obtained at this point is dependent on the bacterial strain, growth conditions, growth phase and the presence of a plasmid imposes a metabolic burden on the cells,

the average plasmid content is expected to account for about 0.5-5% of dry cell weight (DCW) [48]. In spite of the fact that several techniques can be applied for cell disruption, alkaline solutions of a detergent, as NaOH/SDS solutions or its variations, are the most widely used methods. Besides disrupting and solubilizing the physical barriers of the cells, these methods will also result in the irreversible denaturation of gDNA and proteins. The main concern while performing cell disruption relying in an alkaline environment is the pH of the solution upon addition of the buffer used to perform cell lysis. A rapid and efficient homogenization is of extreme importance to ensure uniform distribution of pH across the full solution volume [48]. This process is however hindered by the increase on solution viscosity that results from denaturation of gDNA after cell disruption and can lead not only to incomplete lysis of the cells and/or undesired fragmentation of gDNA, as well as can have negative effects on pDNA. This negative influence on pDNA is due to the possible existence of regions with a pH higher than 12.2-12.4. When subject to alkaline conditions, the pDNA molecule will be partially denatured by separation of the complementary strands of the double-stranded pDNA. However, the “partially” aspect of this event is of maximum importance as only this will enable the maintenance of anchor base pairs formed by a certain number of nucleotides that remain paired and will in turn make possible the renaturation of the pDNA molecule to its supercoiled structure upon addition of the neutralization solution in the subsequent step. If the pDNA is subject to a pH higher than 12.5, the anchor base pairs may be lost and the neutralization will result in an irreversible denatured covalently closed DNA molecule [48].

After the lysis step, the solution is typically conditioned with chilled potassium acetate, to promote neutralization of the solution and precipitation of cell debris and gDNA, after which a unit operation (usually centrifugation at laboratorial scale and depth filtration at industrial scale) that allows the clarification of the solution by removal of the precipitate from the lysate is performed in a careful manner to reduce the release of proteins, LPS (lipopolysaccharides) and gDNA fragments from cell debris and precipitates [48]. The reduction in LPS content is of high importance due to their negative influence on transfection efficiency and toxic effect derived from their immunostimulatory properties [48, 54].

The aforementioned steps constitute the primary purification and are followed by the processing of the clarified lysate in order to further concentrate and purify the plasmid product – the intermediate recovery. In this step special attention is given to the reduction of the content in RNA and proteins, which together account for over 90% of the total mass of solutes [48]. Although attempt of chromatographic methods have been performed immediately after cell lysis, the best process-wise mode is to reduce the impurity load and process volume before subjecting samples to chromatography [54]. Intermediate recovery steps should allow a significant removal in contaminants, ideally to obtain a solution with a content of pDNA over 50% of all solutes in order to reduce the burden on the final purification stage where a unit operation of high cost and performance, typically a chromatographic operation, will be applied. To achieve this objective several unit operations, such as tangential flow filtration, precipitation, aqueous two-phase systems or adsorption, can be combined [48].

1.3.2 Final purification

The final purification of samples of containing plasmid DNA has as main goal the removal of open circular and denatured pDNA molecules from the supercoiled forms in order to obtain a product enriched in the most biologically active form – the supercoiled form. This step should also ensure the removal of contaminants that were not completely removed during the previous purification steps, such as residual gDNA, RNA and LPSs, and thus result in a product complying with the recommendations of regulatory agencies summarily described in Table 1.2. Although pDNA chromatography presents several disadvantages, such as limitations on mass transfer and binding capacity, this method is still the most used to perform the final purification of plasmid biopharmaceutical products [48]. The chromatographic process chosen for pDNA purification depends on the differences in properties that will be explored and can be [54]:

- size-related – size-exclusion chromatography (SEC), slalom chromatography (SLC)
- charge-related – anion exchange chromatography (AEC), hydroxyapatite chromatography (HAC)
- hydrophobicity-related – reversed-phase liquid or ion pair chromatography (respectively RPLC or RPIPC), hydrophobic interaction chromatography (HIC), thiophilic adsorption chromatography (TAC)
- affinity-related – triple-helix affinity chromatography (THAC), protein-DNA affinity chromatography, immobilized metal affinity chromatography (IMAC), boronate affinity chromatography, polymyxin B affinity chromatography.

Several chromatographic strategies have been employed for the manufacturing of pDNA therapeutic products as a single step or in combined modes. The selection of the chromatographic techniques to apply for pDNA purification is dependent not only on the nature and distribution of the residual impurities and contaminants but also by the quantity of pDNA that will be subject to purification. The advantages, disadvantages and applications of some of the aforementioned methods for pDNA purification are presented in Table 1.3 [54].

Besides their use on the manufacturing steps, the chromatographic methods mentioned are also useful as analytical tools to monitor and control pDNA quality during process and in the final formulations, giving information about if the product meets the desired specifications [54].

However, chromatographic separation of pDNA is hindered by the structural nature of available stationary phases, poor selectivity and co-elution of pDNA and impurities due to physicochemical similarities. Also low diffusion coefficients of pDNA and high viscosity of concentrated pDNA and impurities samples which is translated into high-pressure drops are constrains observed when chromatographic methods are applied for pDNA purification. High-pressure drops condition also the flow-rates at which chromatography can be conducted, being a burden possible to circumvent by dilution

of column feed or avoid of use of chromatographic particles with small diameter. An additional solution involves the application of expanded bed mode when performing the chromatographic method [54].

Table 1.3 – Summary of advantages, disadvantages and applications of chromatographic modes applied in pDNA separation and purification (adapted from [54]).

Type	Advantages	Disadvantages	Applications
SEC	Separates endotoxins, gDNA, RNA, oligonucleotides, proteins; Fractionates pDNA isoforms.	Low feed volumes and concentrations; Long chromatographic runs for high resolution media; Product dilution; Limited scale-up at manufacturing scale.	Final polishing step; Group separation; Endotoxin removal; Isoform fractionation.
AEC	Concentration of pDNA; Rapid and simple; Efficient removal of low charge density impurities; Some stationary phases separate pDNA isoforms.	Difficult resolution of sc pDNA from gDNA, other isoforms and multimers; Endotoxins and high M_r RNA co-elute with pDNA; Low capacity; Elution of pDNA in high salt.	pDNA capture; pDNA concentration quantitation.
HIC	Separates endotoxins and single stranded nucleic acids; Some stationary phases separate pDNA isoforms; Mild mobile phases; Rapid chromatographic runs.	In some cases pDNA is eluted in high salt; In some cases pDNA elutes in the flow through and is diluted.	pDNA or impurities capture; pDNA concentration quantitation

As the method applied in the present work relies in the separation of supercoiled DNA molecules of minicircle from open circular plasmid DNA, in the following subchapters will be addressed chromatographic methods that have been reported as successful for the separation of open circular and supercoiled pDNA molecules.

1.3.2.1 Amino acid/DNA affinity chromatography

One of the main problems appointed for the purification of pDNA by chromatographic methods has been the lack of stationary phases displaying good selectivity for separation between pDNA and impurities due to their similar binding affinities what, in turn, renders the desired purification a difficult task to achieve in a single chromatographic step [55]. Affinity chromatography is a powerful technique due to the use of a specific binding agent to analyze or purify molecules on the basis of their biological function or individual chemical structure [56].

Several studies on the interaction between proteins and DNA molecules have highlighted that this interaction is mediated by amino acid residues on the protein and nucleotides on the DNA chain. The occurrence of this type of interaction displays a wide dispersion in nature, playing a crucial role in several cellular processes such as formation and stabilization of nucleosomes in eukaryotes and control of DNA

transcription. Research at atomic level has demonstrated that specific interactions with nucleic acid bases are preferentially mediated by particular amino acids and are dependent on hydrogen bonding, electrostatic interactions, van der Waals and hydrophobic contacts, water-mediated bonds, cation-π and stacking interactions. From these, hydrogen bonds display the strongest specificity, being fundamental for specific recognition [56].

The topology of the DNA is also another important aspect, being indicated that torsional strain associated with supercoiled DNA promotes higher exposition of the bases on DNA chain that are thus more available to interact with amino acids [55-59].

The development of pseudo-affinity chromatography exploiting the interactions between three chromatographic supports harboring histidine, lysine or arginine as ligands and pDNA isoforms were attempted by Sousa *et al*. [55, 57-60] due to the demonstration of interactions at atomic level in several protein-DNA structures [56]. The objective of the proposed methods is the final purification of sc pDNA from open circular isoforms relying in a stationary phase matrix that presents a pseudo-affinity for sc pDNA conferred by its amino acid ligand. For all three supports studied, specific interactions between amino acid ligands with the bases of sc pDNA were explained by deformations on pDNA caused by torsional strain. These deformations are translated into a more compact molecule with higher exposure of bases that will interact with the amino acid ligand [55, 57-60].

Analyzing the separations proposed, the conditions that promote interactions between the amino acids referred and species in the sample subjected to purification, the strength by which the molecules interact with the matrix and the application for which the specific amino acid matrices can be used are summarily indicated in Table 1.4.

Table 1.4 – Characterization of amino acid-nucleic acids interaction and general application in affinity chromatography (adapted from Sousa *et al.* [56]).

Amino acid ligand	Global interaction conditions	Order of favored interaction	General application	Ref.
Histidine	High salt concentration	RNA-sc pDNA-gDNA-oc pDNA	Isolation of pDNA isoforms, purification of sc pDNA from lysate	[57, 60]
Arginine	Mild conditions	sc pDNA-RNA-gDNA-oc pDNA	Isolation of pDNA isoforms, purification of sc pDNA from lysate	[55, 58]
Lysine	Mild conditions	sc pDNA-oc pDNA	Isolation of pDNA isoforms	[59]

Another aspect influencing the purification is the temperature at which chromatographic runs are performed. In this context, retention time of species in the sample to purify varies and lower temperatures (5°C) result in better separations and isolation of sc pDNA molecules [57, 55, 61]. For chromatography using a support harboring histidine as an amino acid ligand, it was also noticed that dynamic binding capacity (DBC) could be improved by increasing the concentration of pDNA in the column feed [62].

Regarding the quality of sc pDNA recovered, chromatographic runs, performed with supports having histidine, lysine or arginine as amino acid ligands, resulted in purified sc pDNA molecules complying with the recommendations of regulatory agencies for impurity content and homogeneity [57, 55, 61].

1.3.2.2 Hydrophobic interaction chromatography

Hydrophobic interaction chromatography was a method originally developed for protein purification and thus the available matrices present a poor capacity for pDNA purification [54]. However, this type of chromatography has been successfully applied with this purpose, not only for pDNA separation from impurities [48, 54, 63] as well to perform the separation of pDNA isoforms [48, 54, 64, 65].

This type of chromatography is characterized by the exploitation of differences on hydrophobic characteristics of biomolecules on the samples to be purified and is typically conducted at 1.5 M ammonium sulfate when separation of pDNA molecules from higher hydrophobic impurities (RNA, denatured gDNA, oligonucleotides and endotoxins) is desired [54]. The higher hydrophobic character of these impurities when compared to pDNA is due to the single stranded nature of RNA, denatured gDNA and oligonucleotides which results in a higher exposition of hydrophobic bases that will interact with the hydrophobic matrix. Regarding endotoxins, its hydrophobicity is explained by the presence of lipid A [48].

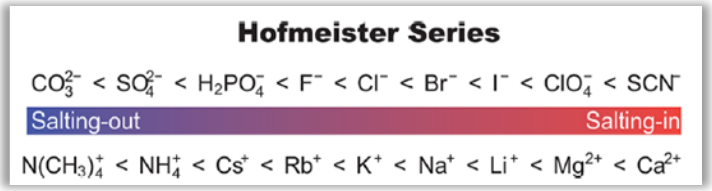

Figure 1.5: Representation of the Hofmeister series with anions arranged in terms of their water affinity and according to their effects on the solubility of macromolecules in aqueous solutions (adapted from [66]).

Although pDNA is naturally a hydrophilic molecule due to shielding of hydrophobic bases inside the double helix and exposition of sugar-phosphate chains that are available to establish hydrogen bonds with surrounding water molecules, this character can be to some extent modified by an increase on ionic strength of a kosmotropic salt on the mobile phase [48]. These salts are strongly hydrated (represented on the Hofmeister series ordered from strongly to weakly hydrated, Figure 1.5 (B)) and present stabilizing and salting-out effect on proteins and macromolecules [67].

Strategies relying in high kosmotropic salt content have been explored in several studies for pDNA isoforms separation as, for example, has been reported by Bo *et al.* [64] and Freitas *et al.* [67]. In both cases, a successful separation of sc pDNA from oc pDNA and other impurities was achieved by applying a mobile phase with a high salt content. Plasmid DNA obtained with these methods has been reported as highly pure, being observed an increase in supercoiled pDNA recovered directly related with the concentration of salt used (98±1.2, 91±1.4 and 77±1.7% for, respectively, 3, 2.5 and 2 M ammonium sulfate) [64]. Use of sodium citrate has also been attempted and, even though the use of this salt results in samples with smaller HPLC purity (80% when a buffer with 1.0 M sodium citrate is used) when

compared to samples eluted with ammonium sulfate (96 and 90% for elution buffers with, respectively 2.5 and 2.0 M ammonium sulfate), its environmental impact is significantly lower than the one presented by ammonium sulfate what could be advantageous from a large-scale processing point of view [67]. Another important aspect is the impurity content reduction to values recommended by regulatory agencies for pDNA quality by ammonium sulfate in hydrophobic chromatography and the need of a second chromatographic method for reaching a highly pure product when sodium citrate is used [67].

1.3.2.3 Size exclusion chromatography

Size exclusion chromatography (SEC) explores differences on molecule size, being larger molecules eluted early during the chromatographic run whilst smaller molecules will have a residence time inversely correlated with their size. Size exclusion chromatographic methods are characterized by elution patterns depending on the access of molecules to the matrix pores being that typical matrices present pores smaller than pDNA molecules and thus these biomolecules usually elute early during the chromatographic run [48].

Although this type of chromatography is typically used for pDNA sample desalting, buffer exchange and removal of RNA and small solutes, it has also been studied for separation of pDNA isoforms [48, 54, 68, 69]. With purpose of different DNA molecules fractionation, the appropriate SEC support has to be selected. An elution pattern consisting in exclusion of pDNA and gDNA whereas RNA, proteins and endotoxins are retained in the column has been reported for two chromatographic supports, both from Pharmacia Biotech: Sephacryl S1000 and Superose 6. Sephacryl S1000 presents an exclusion limit of 20 kbp for linear DNA and loading of pDNA containing samples into columns with this stationary phase results in two resolved peaks being DNA forms (gDNA, oc and sc pDNA and also linear and denatured DNA) eluted in the first peak. However, analyzing samples collected is observed that the latest fractions of this peak are enriched in sc pDNA being thus obtained separation of this DNA form. This support is characterized by yields close to 70% for recovery of sc pDNA being the best results obtained with injection of low pDNA concentration samples [54, 69]. Regarding the Superose 6, the exclusion limit for DNA is 450 bp, significantly lower than the one presented by Sephacryl S1000, and its application for pDNA separation during SEC methods results in shorter runs (20 minutes compared to the 200 to 300 minutes needed for Sephacryl S1000), being observed separation of pDNA from low molecular weight impurities that are retained in the support. This support allows also injection of samples with higher pDNA concentration but is however characterized by a more difficult separation of gDNA and oc pDNA from sc pDNA [54, 69].

For the use of size exclusion chromatography with pDNA purification purposes, small loading volumes and low flow rates should be used in order to improve the separation. Due to the fact that separation is favored by small impurity content of loaded samples, this method is usually recommended as a final polishing step during downstream processing of plasmid DNA [69].

2. Background and objectives

Gene therapy and DNA vaccination are fields in which researchers have demonstrated a growing interest. Although viral vectors are amply used in cell transduction and can efficiently deliver DNA to the cell nucleus, the possible integration of viral sequences to the receptor cells genome is a concern during therapeutic scenarios. The possible application of these technologies to human patients for the treatment, cure or prevention of a specific disorder has rendered the pursuit and development of safe and efficient non-viral vectors a requirement.

In this context, minicircles present several advantages over conventional plasmids for the aforementioned applications, namely their lack of bacterial sequences, which in turn results in low immunogenicity, and smaller size that renders their access to cell nucleus an easier task and thus confers higher transfection efficiency. These vectors are produced by the replication and recombination *in vivo* of a parental plasmid in *E. coli*. The recombination event results in the formation of two products mainly in a supercoiled state – a miniplasmid and the therapeutically valuable minicircle – meaning that the production should be optimized to guarantee maximum minicircle production while ensuring a minimum content in un-recombined parental plasmid. After *in vivo* production, the *E. coli* cells have to be disrupted in order to release the recombination products and thus cell debris, host proteins, endotoxins, RNA and host genomic DNA will also be present in the solution obtained. To be used in therapeutic applications, the minicircle will have to be purified in processes that allow the removal of these contaminants while ensuring a high minicircle recovery and minimal conversion of supercoiled form to open circular, due to higher biological activity of the first isoform mentioned. However, the minicircle purification step where miniplasmid and un-recombined parental plasmid are removed is hindered by the physicochemical similarities between these three DNA forms and almost identical size of minicircle and miniplasmid molecules.

The main focus of the present work was to develop a simple, robust, cost-effective and scalable process for minicircle (MC) purification that guarantees the obtainment of a product complying with the recommendations of regulatory agencies for therapeutic grade quality. The simplicity of the proposed method relies in the use of a single chromatographic step for the separation of impurities (RNA, miniplasmids, un-recombined parental plasmids) from minicircles after an enzymatic digestion with a nicking-enzyme. The only DNA forms harboring a restriction site for this enzyme, which introduces a single cut in one of the chains of double stranded DNA molecules, are the parental plasmid (PP) and miniplasmid (MP). This nicking-enzyme thus converts supercoiled MP and PP into the corresponding open circular forms. The difference between the hydrophobicity of open circular isoforms (MP and PP) and supercoiled DNA molecules (MC) is then explored next by performing a hydrophobic interaction chromatography step.

3. Materials and methods

The methods described in the sections below were performed using the reagents presented in Table 3.1, along with manufacturer, grade and purity. Current laboratorial material was used for preparation, handling and storage.

Table 3.1 – Description of manufacturer, grade and purity of reagents used in the methods applied for minicircle production and purification in the present work.

Reagent	Manufacturer	Grade	Minimum Purity
Luria Bertani Broth (LB)	Sigma	molecular biology	-
NaOH	Fisher Chemical	p.a	98.7%
Glucose	Fisher Chemical	p.a	-
Kanamycin	AMRESCO	USP	ultra pure
L-(+)arabinose	Merck	for microbiology	-
Seakem LE agarose	Lonza	for electrophoresis	-
Tris	Eurobio	molecular biology	-
Acetic acid	Fisher Chemical	p.a	99.8%
EDTA	Merck	p.a	98.0%
Sucrose	Fisher Chemical	p.a	-
Bromophenol blue	Sigma	for electrophoresis	-
Ethidium bromide solution	Sigma	-	-
EDTA base	Panreac	PRS	98.0%
HCl	Liedel-de Häen	p.a, ACS, ISO	37.0%
SDS	Merck	for synthesis	90.0%
Potassium acetate	Acros Organics	ACS	99.5%
Isopropanol	Sigma	ACS	99.5%
Ammonium Sulfate	Panreac	ACS	99.0%
NaCl	Panreac	p.a, ACS, ISO	99.5%
KCl	Merck	p.a	99.5%
KH_2PO_4	Panreac	p.a	99.0%
Azide	Panreac	-	99.0%

3.1 Bacterial strain and plasmid

In this work, the parental plasmid was produced in *E. coli* BW2P, a strain developed in our laboratory during the PhD of Michaela Šimčíková [70]. This strain was constructed by disruption of the *endA* gene via the insertion of a single copy of the P_{BAD}/*araC-RBS-parA* cassette into the commercial BW27783 *E. coli* strain chromosome. The ribosome binding site (RBS) was optimized on the basis of an equilibrium statistical thermodynamic model and introduced upstream of *parA* by replacing the original RBS via SOEing PCR. After successful cassette insertion, the construction of BW2P was finalized by knocking out the *recA* gene [70].

The parental plasmid used (pMINILi-CMV-VEGF-GFP) was constructed by Liliano Brito ([71], in preparation) by modifying plasmid pMINI8 constructed by Michaela Šimčíková [70]. Briefly, the

P_{BAD}/araC-parA cassette on pMINI8 was eliminated by directed mutagenesis by resorting to the introduction of AgeI restriction sites. The resulting plasmid, pMINILi, was then modified to construct pMINILi-CMV-VEGF-GFP. With this purpose, plasmid pVAX-VEGF-GFP available in our laboratory was subjected to a double digestion with EcoRI and KpnI. This plasmid only presents one restriction site for each of these restriction enzymes, thus originating two DNA fragments after digestion – the fragment of the VEGF gene and a fragment corresponding to the remaining pVAX-VEGF-GFP sequence. The same enzymes were also used for digestion of pMINLi, cutting the plasmid in two fragments of 35 and 3952 bp. The latter fragment was subjected to ligation with the VEGF fragment (611 bp), resulting in the construction of pMINILi-CMV-VEGF-GFP (4563 bp). Construction of pMINILi and pMINILi-CMV-VEGF-GFP was successfully evaluated by digestion with restriction enzymes and sequencing.

3.2 Cell banks

For the preparation of cells banks, 50 µL of chemically competent *E. coli* BW2P cells were gently thawed on ice and mixed with ≈ 50 ng of pMINILi-CMV-VEGF-GFP. Mixture was kept in ice for 20 minutes after which a heat shock at 42°C for 1 minute was performed. The mixture was then immediately placed on ice and left to rest for 2 minutes. After this incubation, mixture was resuspended with 800 µL of sterile LB medium and incubated at 37°C for 1 hour for cell recuperation. Transformed cells were then plated on LB agar plates prepared with medium supplemented with 0.5% (w/v) glucose and 30 µg/mL kanamycin, which were incubated overnight at 30°C. Isolated colonies were then picked and used to inoculate 5 mL of sterile LB medium supplemented with 0.5% (w/v) glucose and 30 µg/mL kanamycin in a 15 mL Falcon tube (VWR). This inoculum was incubated at 30°C and 250 rpm until reaching mid-exponential phase, correspondent to an OD_{600nm} (optical density at 600 nm) of 1 to 1.2. At this stage, cell banks were prepared by mixing 500 µL of bacterial growth with the same volume of 30% (v/v) glycerol and stored at -80°C.

Banks created were confirmed by digestion of purified plasmids (High Pure Plasmid Isolation Kit, Roche) with restriction enzyme NcoI for which the parental plasmid presents three restriction sites (at 3350, 952 and 533 bp on the plasmid sequence). After restriction the presence of two expected bands (2398 and 1746 bp) confirmed successful bank construction.

3.3 Cellular growth and minicircle production

To perform cellular growth, LB medium was prepared following manufacturer indications and pH was corrected to 7.5, before thermal sterilization at 121°C for 21 min, by addition of 1 M NaOH. A pre-inoculum was prepared in 15 mL Falcon tubes (VWR) by inoculating 5 mL of LB supplemented with 30 µg/mL kanamycin and 0.5% (w/v) glucose with a loop of frozen BW2P cells harboring the pMINILi-CMV-VEGF-GFP plasmid, and incubating overnight at 37°C and 250 rpm. The next day, the inoculum was prepared in 30 mL of LB in 100 mL shake flasks, supplemented with 30 µg/mL kanamycin and 0.5% (w/v) glucose. The appropriate volume of pre-inoculum to obtain an optical density of 0.1,

measured at 600 nm (OD_{600nm}), was determined and centrifuged in a tabletop centrifuge (Eppendorf centrifuge 5417R) at 6000 g for 3 minutes prior to inoculum seeding. Cultures were incubated at 37°C and 250 rpm until an $OD_{600nm} \approx 2.5$ (late exponential phase) was reached. Cellular growth was performed at 37°C and 250 rpm with 250 mL of LB in 2 L shake flasks, supplemented with 30 µg/mL of kanamycin, after inoculation with the cells obtained by centrifugation (6000 g for 3 minutes in a Eppendorf centrifuge 5810R) of the volume of inoculum needed to achieve an OD_{600nm} of 0.1. Cells were incubated for about 3 hours to reach late exponential phase ($OD_{600nm} \approx 2.5$). At this moment, recombination was induced by addition of 0.01% (w/v) L-(+)arabinose directly into the medium. Samples of 2 mL of culture were collected at 0, 1 and 2 hours of recombination, centrifuged at 6000 g for 3 minutes (Eppendorf centrifuge 5417R) and stored at -20°C until further analysis of recombination efficiency. After 2 hours of recombination, cell growth was suspended and the medium was centrifuged at 6000 g for 15 minutes in a Sorval RC 6 centrifuge with a SLA 3000 rotor. The cell pellets obtained were stored at -20°C until further processing.

3.4 Gel electrophoresis

Horizontal gel electrophoresis was performed using an EPS 301 power supply (Amersham Pharmacia Biotech) and two electrophorese units (HE99X Max Horizontal Unit from Hoefer for large gels and a kuroGEL Miniplus 10 from VWR for small gels). The gels were prepared with 1% (w/v) agarose (SeaKem LE agarose) in TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA, pH 8) and loaded with samples mixed with a 6X loading buffer (40%(w/v) sucrose, 0.25%(w/v) bromophenol blue), using NZYDNA ladder III (NZYTech) as molecular weight marker. Gel electrophoresis was performed with 1% TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA, pH 8) at 100 V for 60 minutes and 120 V for 90 minutes for small or larger gels, respectively. Gels were stained in an ethidium bromide solution (0.4 µg/mL) and images were obtained with an Eagle Eye II gel documentation system (Stratagene).

3.5 Analysis of recombination efficiency by densitometry

Samples of 2 mL of culture, collected after induction of recombination, were processed with the High Pure Plasmid Isolation Kit (Roche) following the manufacture's protocol, in order to purify the plasmid produced during cellular growth. The concentration of total plasmid purified was measured using a Nanovue Plus (GE) spectrophotometer. The volume equivalent to 500 ng of plasmid isolated from samples collected at the end of bacterial growth (after 2 hours of recombination) was digested for 2 hours at 37°C with 0.2 µL of SacII (Promega) – a restriction enzyme for which only the parental plasmid and miniplasmid present a recognition site (localized in the 3.9 kbp region of the parental plasmid) – in the presence of the appropriate reaction buffer (buffer C, Promega). The digested samples were subjected to gel electrophoresis, as described in section 3.4. Recombination efficiency (Re) was determined by densitometry analysis of the band intensities of parental plasmid and linearized miniplasmid, using the ImageJ software and equation 3.1, where A_{PP} and A_{MP} are, respectively, the areas of the band intensities

of parental plasmid and linearized miniplasmid and r_m is the molar ratio of parental plasmid to miniplasmid [39].

$$\mathrm{Re}\ (\%) = \left(1 - \frac{\mathrm{A_{PP}}}{\mathrm{A_{PP}} + r_m \mathrm{A_{MP}}}\right) \times 100 \quad \textbf{(3.1)}$$

3.6 Primary purification

3.6.1 Alkaline lysis

The method for alkaline lysis performed in this work was based in the protocol described by Birnboim *et al.* [72]. The cells harvested at the end of cellular growth were resuspended in P1 buffer (50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA base, pH 8) resorting to vortex. The pH of buffer P1 was previously corrected by addition of 1 M HCl. The volume of resuspended cells was divided into 50 mL centrifuge tubes and the volume of P2 buffer (0.2 N NaOH, 1%(w/v) SDS) was added to each tube in a 1:1 ratio (v/v), in order to perform alkaline lysis of the cells. The mixture was gently homogenized and left to rest at room temperature for 10 minutes. To stop cell lysis and neutralize the mixture, the appropriate volume of P3 buffer (5 M potassium acetate, acetic acid) was added. After gentle homogenization, the tubes were placed on ice and left to rest for 10 minutes. The volumes of P1, P2 and P3 buffers (respectively V_{P1}, V_{P2} and V_{P3}) used in the steps above were calculated to concentrate the solution to an OD_{600nm}=60, taking into account the final OD_{600nm} and volume of the respective cellular growth (V_{cg}), following the equation 3.2.

$$\mathrm{V_{P1}} = \mathrm{V_{P2}} = \mathrm{V_{P3}} = \frac{\mathrm{OD_{600nm}} \times \mathrm{V_{cg}}}{60} \quad \textbf{(3.2)}$$

Immediately after neutralization, the mixture was centrifuged at 4°C for 30 min at 13000 rpm in a Sorvall RC 6 centrifuge to remove cell debris and some genomic DNA and proteins. The supernatant was transferred into new tubes and subjected to an additional centrifugation in order to guarantee a better clarification of the alkaline lysate.

3.6.2 DNA pre-purification

The volume of clarified alkaline lysate obtained in section 3.6.1 was measured and 0.7 volumes of pure isopropanol were added. The mixture was gently homogenized and distributed by 50 mL Falcon tubes (VWR) that were left to rest at -20°C for 2 hours to promote precipitation of nucleic acids. After this period, the mixture was centrifuged at 12000 rpm and 4°C for 30 min in an Eppendorf centrifuge 5810R. After centrifugation, the supernatant was discharged and the pellet was left to dry overnight at 4°C.

The pellets in each tube were resuspended in 500 µL of 10 mM Tris-HCl, pH 8 and pooled in a single 50 mL Falcon tube (VWR). A second passage with the same buffer was performed in each tube, in order to minimize DNA loss.

The DNA rich solution obtained was conditioned to 2.5 M of ammonium sulfate by dissolution of appropriate amount of salt. Following homogenization, solution was left to rest for 15 min on ice and then centrifuged in an Eppendorf centrifuge 5417R for 30 min at 13000 rpm and 4°C. The supernatant was

collected into a 15 mL Falcon tube and diluted with an equal volume of MilliQ water, in order to dilute salt concentration.

3.7 Diafiltration and concentration

To perform desalting, the solution obtained in section 3.6.2 was subjected to centrifugation at 4000g and 4°C (Eppendorf centrifuge 5810R) in a 2 mL Amicon with a membrane cut-off of 30kDa (Amicon Ultra-2 30K from Millipore). The total volume obtained from 1 litter of cellular growth was processed in several passages with a single Amicon unit, discharging the permeate. The DNA solution, retained by the membrane, was washed with MilliQ water, using the same volume as the initial diluted solution. To collect the concentrated DNA, the Amicon unit was inverted and centrifuged for 5 min at 4000g and 4°C, being recovered a volume between 40 and 500 µL, depending on duration of the previous centrifugation step. To guarantee maximum recovery of DNA, the unit was again washed with 2 mL of MilliQ water and the concentrate was recovered using the strategy described above.

3.8 Miniplasmid enzymatic digestion

The concentration of the solutions, obtained in section 3.7, was measured in a NanoValue Plus (GE) spectrophotometer, using the appropriate dilution. To convert the supercoiled miniplasmid in its relaxed form, 1 µL of Nb.BbvCI from New England Biolabs (10 U/µL) per 700 µg of nucleic acids was used. The volume of reaction buffer (NEB2, New England Biolabs) added was equivalent to one tenth of the final digestion volume. The mixtures were incubated at 37°C overnight and, in the next morning, a sample of 0.5 µL was diluted 1:10 in MilliQ water and analyzed by gel electrophoresis, as described in section 3.4, to evaluate if all miniplasmid had been converted to the open circular form.

3.9 Final purification

3.9.1 Column preparation

To prepare the HIC column, 10 mL of Phenyl Sepharose 6 Fast Flow resin (GE Healthcare) were washed with MilliQ water to remove any traces of ethanol, used as storage buffer by the manufacturer. The slurry was slowly added to a column with inner diameter of 10 mm (Tricorn 10/100, GE Healthcare), using a 2 mL Pasteur pipette and a flow rate of 1 mL/min imposed by a peristaltic pump connected to the column bottom. The addition of the slurry to the column was performed having special care to not form air bubbles in the packed volume. After all slurry was added, the column was closed with the top adjustable cap and connected to an ÄKTApurifier100 system (GE Amersham). The final packing was performed using a continuous flow of MilliQ water that was gradually increased, from 1 mL/min to 18 mL/min, until no alteration on resin level was observed at the imposed flow rate. For column storage, 40 mL (4 CV) of PBS+0.02% azide (0.9% (w/v) sodium chloride in 10 mM phosphate buffer, pH 7.4, 0.02% (w/v) azide)

were passed through the column at 5 mL/min. The column was then disconnected from the system and stored at room temperature until further utilization.

3.9.2 Hydrophobic interaction chromatography

Hydrophobic interaction chromatography was performed using the column prepared in section 3.9.1 connected to an ÄKTApurifier100 system (GE Amersham) under the control of UNICORN 5.11 software (GE). The column was washed with MilliQ water until a conductivity inferior to 0.08 mS/cm was reached and no variation on controlled parameters was observed. The mobile phase consisted on mixtures of buffer A (2.2 M Ammonium Sulfate in 10 mM Tris-HCl pH8) and buffer B (10 mM Tris-HCl, pH 8) prepared by the system on a 2 mL Mixer M-925 Mixing Chamber (GE Amersham). The absorbance of the eluate was continuously measured at 254 nm by a UV detector positioned after the column outlet.

For separation by linear gradient, a flow rate of 2 mL/min was imposed. The column was equilibrated with 3 column volumes (CV) of buffer A and 1 mL of sample conditioned with 2.5 M ammonium sulfate (by dissolution of the appropriate amount of salt) was manually injected into a 1 mL loop on the system. After column equilibration, the sample was injected into the column by washing the loop with 3 mL of buffer A. All unbound material was washed out of the column in 3 CV at 0% B. The linear gradient was performed from 0% B to 100% B in 15 CV. The eluate was collected during the course of the chromatographic run in 1.5 mL fractions in 2 mL eppendorf tubes positioned on a Frac-920 collector (GE Amersham).

The purification of minicircles was performed using a step gradient at 2 mL/min. The column was equilibrated with 3 column volumes (CV) of buffer 17% B (≈ 204 mS/cm) and 1 mL of sample, obtained in section 3.8 and conditioned with 2.5 M ammonium sulfate by dissolution of the appropriate amount of salt, was manually injected into a 1 mL loop on the system. After column equilibration, the sample was injected into the column by washing the loop with 3 mL of buffer 17% B (≈ 204 mS/cm). All unbound sample was washed out of the column in 4 CV at 17% B (≈ 204 mS/cm). The first and second steps were performed at, respectively, 35% B (≈ 173 mS/cm) and 100% B (≈ 2 mS/cm) during 2 CV each. The eluate was collected during the course of the chromatographic run in 1.5 mL fractions in 2 mL eppendorf tubes positioned on a Frac-920 collector (GE Amersham).

In between runs, the column was washed with MilliQ water until a conductivity inferior to 0.08 mS/cm was reached and no variation on controlled parameters was observed. For column storage this procedure was also applied being, in this case, followed by the passage through the column of 40 mL (4 CV) of PBS+0.02% azide (0.9% (w/v) sodium chloride in 10 mM phosphate buffer, pH 7.4, 0.02%(w/v) azide) at 5 mL/min. The column was then disconnected from the system and stored at room temperature until further utilization.

A cleaning in place was performed after each 5 runs with 2 CV of 1 M NaOH followed by 3 CV of MilliQ water, 2 CV of ethanol 20% (v/v) and 3 CV of MilliQ water.

3.10 Micro-dialysis

For identification of the purified samples, 200 µL of peak fractions were collected into 0.5 mL eppendorf tubes. The tube caps were removed and the top of each tube was covered with a dialysis membrane with a thickness of 23 µm and 12000-14000 molecular weight cut-off (OrDial D14, Orange Scientific) that was kept in place with a rubber band. The tubes were inverted, assuring that all the solution was near the membrane, and placed floating in a beaker with 2 L of MilliQ water for desalting. The dialysis was performed for 48 hours at 4°C with agitation, after which 15-24 µL of desalted sample were analyzed by gel electrophoresis, as described in section 3.4.

3.11 Restriction enzyme mapping for DNA identification

To confirm the identity of purified forms, 15 µL of desalted samples, corresponding to DNA peaks of the chromatogram, were digested with several enzymes (XhoI from Promega, Bsp1407I (isoschizomer of BsrGI) and PvuII, both from ThermoScientific). The digestions were prepared for 20 µL of total digestion volume using 0.3 µL of enzyme, 2 µL of the corresponding reaction buffer (respectively, buffer D from Promega, buffer Tango and buffer G, both from ThermoScientific) and 2.7 µL of autoclaved MilliQ water. Digestions were performed by incubation at 37°C for 2 hours and analyzed by gel electrophoresis as described in section 3.4.

The restriction enzyme XhoI was also used for identification of minicircle bands on samples collected at the end of cellular growth. Samples of 2 mL of culture, collected after 2 hours of induction of recombination, were processed with the High Pure Plasmid Isolation Kit (Roche) following the manufacture's protocol, in order to purify the plasmid produced during cellular growth. The concentration of total plasmid purified was measured using a Nanovue Plus (GE) spectrophotometer. The volume equivalent to 500 ng of plasmid was digested for 2 hours at 37°C with 0.2 µL of XhoI (Promega) – a restriction enzyme for which only the parental plasmid and minicircle present a recognition site (localized in the 864 bp region of the parental plasmid) – in the presence of the appropriate reaction buffer (buffer D, Promega).

4. Results and discussion

4.1 Process overview

The process developed in the present work for the production and purification of minicircles involved a sequence of shake flask fermentation for bacterial growth and minicircle production, cell harvest, processing and purification that can be summarily described by the flowchart presented in Figure 4.1.

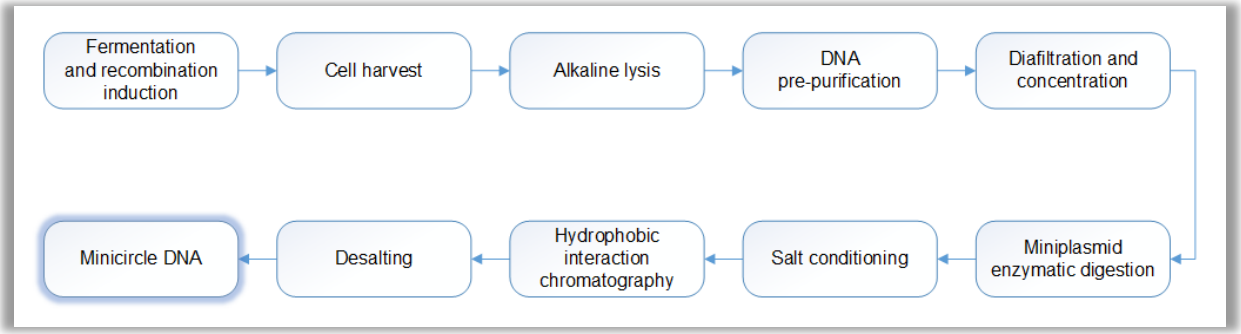

Figure 4.1: Summary flowchart of the procedural steps used in the present work for the production and purification of minicircle DNA.

4.2 Bacterial strain and parental plasmid

The BW2P *E. coli* strain used for plasmid production and recombination in the present work was constructed by Šimčíková, M. [70] by disruption of the *endA* gene (responsible for non-specific digestion of plasmid DNA) via the insertion of a single copy of the P_{BAD}*/araC-RBS-parA* cassette into the commercial BW27783 *E. coli* strain chromosome to obtain a ParA resolvase-expressing strain. After successful insertion, the construction of BW2P was finalized by knocking out the *recA* gene (responsible for homologous recombination) [70]. The BW27783 used for BW2P construction is a strain improved for arabinose uptake due to expression of the gene encoding the low-affinity, high capacity arabinose permease (*AraE)* from a constitutive promoter (P_{CP8}) instead of the native promoter, thus eliminating the natural all-or-none induction of P_{BAD} promoter [73]. The cassette inserted into the *E. coli* chromosome presented an optimized ribosome binding site (RBS) with a strength selected to increase expression of ParA resolvase by tuning the RNA translation initiation [70].

The parental plasmid backbone pMINILi-CMV-VEGF-GFP used in this work was constructed by Liliana Brito during her Master Degree Thesis ([71], in preparation) and was derived from pVAX1GFP by a series of modifications [70]. The parental plasmid (4563 bp) obtained harbors a high-copy replication origin (pUC), a kanamycin resistance selection marker (KanR), an eukaryotic cassette (with a CMV promoter, VEGF and GFP genes and the bovine growth hormone polyadenylation signal) and two multimer resolution sites (MRS). These sites are recognized by the ParA resolvase to catalyze intramolecular site-specific recombination (illustrated on Figure 4.2), originating two DNA molecules: i) a

miniplasmid (2106 bp), comprised of the bacterial backbone and ii) a minicircle (2457 bp), formed by the eukaryotic cassette. The parental plasmid contains also a specific recognition site in the bacterial backbone for the nicking endonuclease Nb.BbvCI [74]. This site is used for the *in vitro* conversion of the supercoiled miniplasmid (sc MP) into its open circular form (oc MP) and is crucial for the purification of supercoiled minicircle (sc MC) proposed in this work.

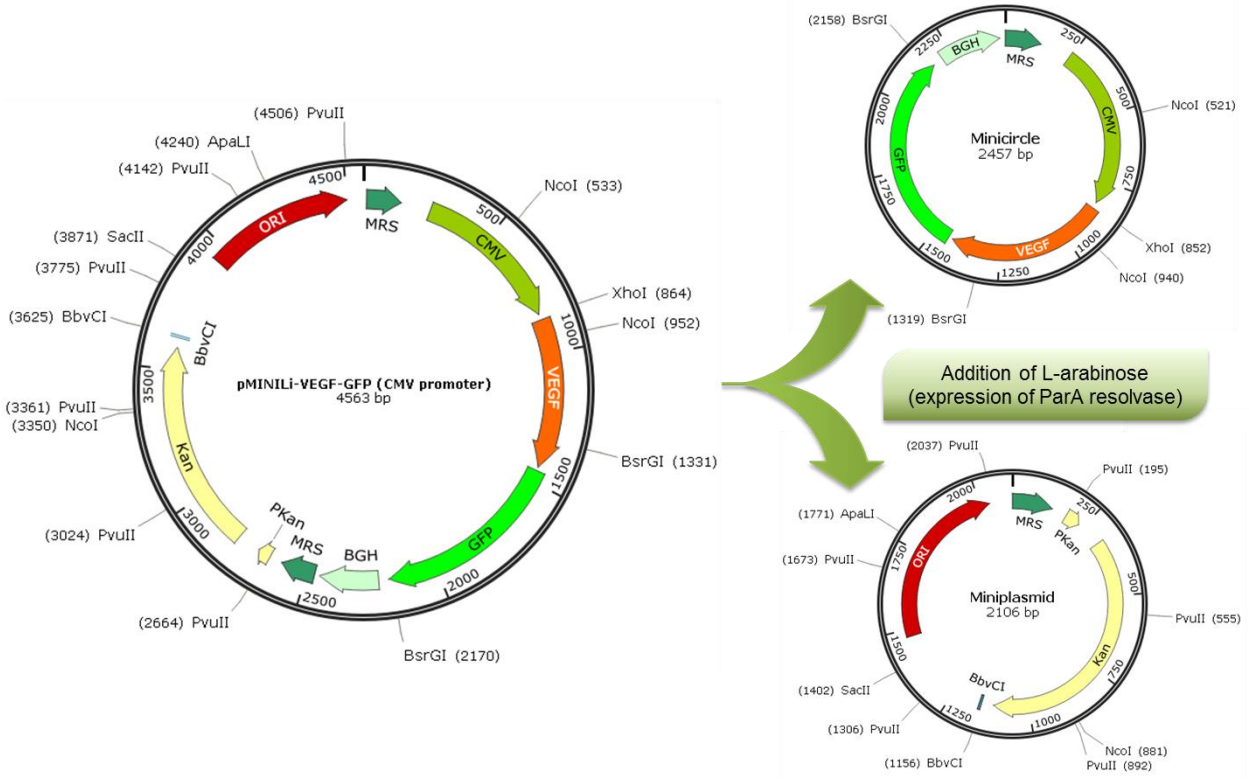

Figure 4.2: Schematic representation of the *in vivo* recombination event mediated by expression of ParA resolvase. Upon addition of 0.01% of L-arabinose to the cellular growth medium, the p*arA* gene, which is under tight transcription control of the P_{BAD}/*araC* arabinose promoter hosted on the *E. coli* BW2P chromosome, is induced to express ParA resolvase. The ParA catalyzes the intramolecular recombination between the two MRS on the parental plasmid pMINILi-CMV-VEGF-GFP, resulting in the excision of a DNA segment and giving origin to two circular DNA molecules, one miniplasmid containing the bacterial backbone and one minicircle containing the eukaryotic expression cassette. MRS - multimer resolution sites of the *parA* resolvase system; CMV – cytomegalovirus immediate early promoter; VEGF - vascular endothelial growth factor; GFP – green fluorescent protein; BGH – bovine growth hormone polyadenylation signal; pKan – kanamycin resistance gene promoter; kan – kanamycin resistance gene; ORI – origin of replication; ApaLI, BbvCI, BsrGI, NcoI, PvuII, SacII, XhoI –recognition sites for the respective restriction enzymes.

Digestion of pMINILi-CMV-VEGF-GFP purified from BW2P cell banks with NcoI yielded two DNA bands on an agarose gel, corresponding to 2398 and 1746 bp fragments (Figure 4.3). Although the lower molecular weight band (419 bp) expected from the plasmid map (Figure 4.2) was not visualized in the gel, the presence of the two observable bands was sufficient to confirm plasmid identity.

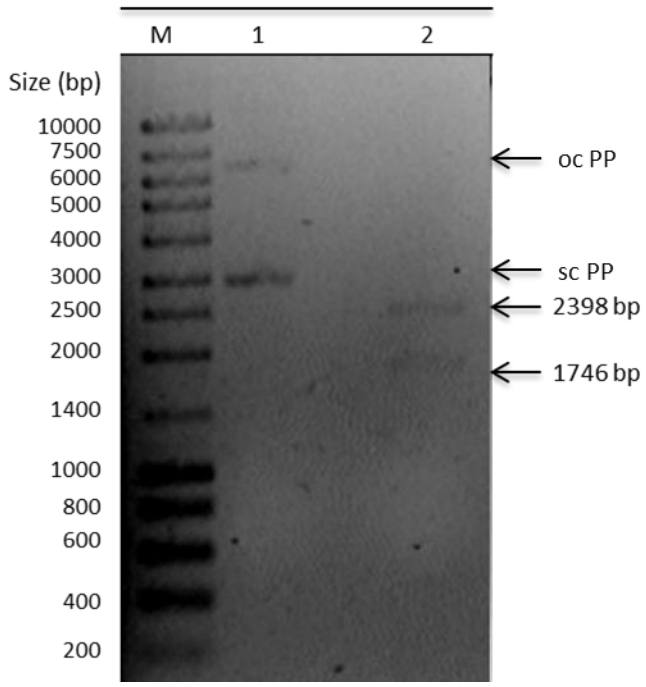

Figure 4.3: Agarose gel corresponding to the digestion of pMINILi-CMV-VEGF-GFP purified from BW2P cell banks with the restriction enzyme NcoI. Lane M – molecular weight marker; lane 1 – non-digested parental plasmid; lane 2 – pattern of digestion of pMINILi-CMV-VEGF-GFP with NcoI showing two (2398 and 1746 bp) of the three fragments expected after digestion. oc PP – open circular parental plasmid; sc PP – supercoiled parental plasmid.

4.3 Cellular growth and parental plasmid recombination

The strategy for cellular growth and plasmid recombination was previously developed by Šimčíková, M. [70] and Liliana Brito ([71], in preparation). For parental plasmid replication and minicircle production, seeding cultures of BW2P cells harboring the pMINILi-CMV-VEGF-GFP plasmid were conducted in the presence of 0.5% (w/v) glucose for recombination inhibition [39] and allowed to reach an OD_{600nm} of approximately 2.5 (late exponential phase). The inclusion of glucose at this phase of the production process is important due to its effect on repression of the P_{BAD}/*araC* arabinose promoter by reduction of 3',5'-cyclic AMP. This results in the prevention of leaky expression of ParA resolvase, thus avoiding premature recombination of the parental plasmid before arabinose induction [39, 42].

These cultures were then used to inoculate 250 mL of LB in 2 L shake flasks with a starting OD_{600nm} of 0.1 and cellular growth was followed by hourly measure of OD_{600nm} until late exponential phase was reached (OD_{600nm}≈2.5 reached at approximately between 2.5 to 3 hours of growth). At this point, ParA resolvase expression was induced by addition of 0.01% (w/v) of L-arabinose directly to the growth medium. Cells were then allowed to grow for two additional hours after induction (Figure 4.4 (A)) to promote complete recombination of parental plasmid. The analysis of the bacterial exponential log phase (Figure 4.4 (B)) results in the determination of a maximum specific growth rate of 1.04 h^{-1} for the bacterial culture (data obtained from eight independent cultures). Plasmid DNA production during cellular growth was estimated by purifying pDNA from 2 mL samples collected at the end of cellular growth with a commercial purification kit (High Pure Plasmid Isolation Kit, Roche). This kit results in DNA containing

samples that are free from RNA due to the use of RNase. Considering 20 independent cultures, the mass of plasmid DNA determined to be produced was 6.6 ± 1.7 mg per liter of bacterial growth. Although these values are merely indicative and do not consist on a precise analysis, the deviation registered can be explained by batch to batch variations.

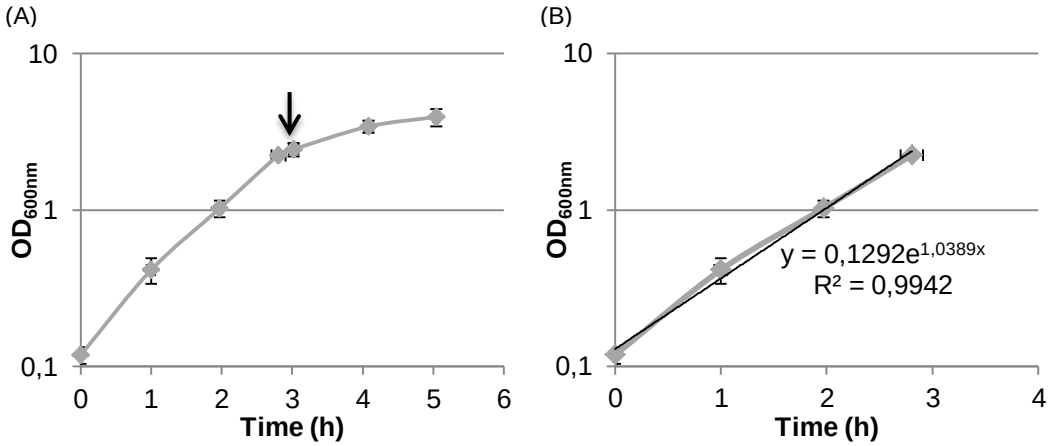

Figure 4.4: (A) Average growth curve of *E. coli* BW2P harboring pMINILi-CMV-VEGF-GFP in LB at 37°C and 250 rpm. Black arrow indicates the time of induction of recombination with 0.01%(w/v) L-arabinose. (B) Average representation of the exponential log phase of *E. coli* BW2P harboring pMINILi-CMV-VEGF-GFP in LB at 37°C and 250 rpm. Average growth curves were obtained from eight independent cell cultures, by calculation of the mean time and mean OD_{600nm} at which each point was collected and measured. Error bars are indicative of the standard deviation obtained for the mean time (horizontal bars) and mean OD_{600nm} (vertical bars).

To analyze the recombination efficiency, 2 mL samples of cell culture were collected before induction and hourly during the subsequent two hours. These samples were processed with a commercial pDNA purification kit (High Pure Plasmid Isolation Kit, Roche) following the manufacturer's protocol. An analysis of the agarose gel obtained with the purified samples (Figure 4.5) shows that samples corresponding to 0 h of induction (lanes 1, 4, 7 and 10) do not present bands with sizes lower than the size of the parental plasmid (4563 bp). This indicates that recombination did not happen before addition of L-arabinose, confirming a successful tight control of the expression of ParA resolvase. The fact that these bands are not exactly located in the 4.5 kbp region is related with the molecule topology. Due to the supercoiled state, plasmids will migrate slightly more on the gel to a region corresponding to ≈3 kbp bands of the molecular weight marker. The fact that the molecular weight marker used is composed of non-supercoiled DNA fragments explains the differences observed between the migration of supercoiled molecules on samples studied and bands of the molecular weight marker [65].

The presence of a tenuous band with higher molecular weight (≈5 kbp) can be attributed to the open circular form of parental plasmid that, having a more distended conformation, will migrate less on the gel. The tenuous bands of higher molecular weight can be attributed to the formation of multimers during cell growth. As for the intense band around 7 kbp, the fact that site-specific recombinases like ParA act upon supercoiled molecules [33] together with the absence of this band after recombination (lanes 2, 3, 5, 6 ,8, 9, 11 and 12), suggests that this is not an open circular form of the parental plasmid. Taking into

account the band apparent molecular weight (≈ 7 kbp), the most probable explanation is that this band corresponds to dimers of the parental plasmid.

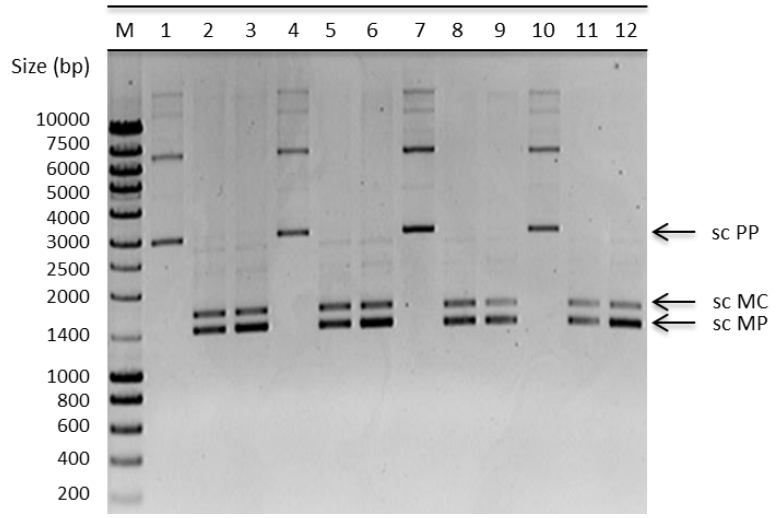

Figure 4.5: Gel electrophoresis analysis for qualitative quality control of samples obtained at 0, 1 and 2 hours after induction of recombination and purified with the High Pure Plasmid Isolation Kit (Roche). Lane M – molecular weight marker; lanes 1, 4, 7 and 10 – 500 ng of DNA from samples collected at 0 hours of recombination induction; lanes 2, 5, 8 and 11 – 500 ng of DNA samples collected at 1 hour of recombination induction; lanes 3, 6, 9 and 12 – 500 ng of DNA from samples collected at 2 hours of recombination induction. sc PP – supercoiled parental plasmid (4563 bp); sc MC – supercoiled minicircle (2457 bp); sc MP – sc miniplasmid (2106 bp).

The bands visible in the other lanes represent DNA molecules obtained after 1 hour (lanes 2, 5, 8 and 11) and 2 hours (lanes 3, 6, 9 and 12) of recombination induction. In all cases, a recombination efficiency of practically 100% and the presence of supercoiled (sc) forms of miniplasmid (lowest molecular weight band with ≈ 1.5 kbp) and minicircle (second lowest molecular weight band with ≈ 1.7 kbp) are observed. Two tenuous bands slightly above the bands before mentioned can also be detected, which are attributed to the open circular isoforms of MP (≈ 2.4 kbp) and MC (≈ 2.9 kbp). An interesting observation results from the comparison of the apparent MC/MP ratio in samples obtained after 1 and 2 hours of recombination. A slightly higher content of MP relative to MC is observed when analyzing the bands from samples with 2 hours of recombination (lanes 3, 6, 9 and 10). This can be explained by the fact that MP molecules contain the bacterial origin of replication and thus keep replicating after recombination, whilst only one molecule of MC is obtained from each parental plasmid. The fact that this difference on MC/MP ratio is more noticeable on the samples with 2 hours of recombination than in the ones obtained only after 1 hour of L-arabinose addition and that apparently with 1 hour of induction recombination is complete, indicates that the allowed time for growth after induction could be reduced to only 1 hour. From the results obtained, the second hour appears to mainly result in MP replication without significant increase in the content of the form of interest (MC).

For band identity confirmation and determination of recombination efficiency, the samples obtained after 2 h of recombination were subjected to enzymatic restriction with SacII and XhoI. Digestion of samples containing the MC and MP supercoiled isoforms will result in different patterns depending on the enzyme used. As shown in Figure 4.2, the MC presents a single recognition site for XhoI (at 852 bp on

the MC sequence) and does not harbor a recognition sequence for SacII, whilst in the case of the MP the inverse situation is observed (recognition site for SacII at 1402 bp on the MP sequence). As both these sites are present on the parental plasmid, un-recombined parental plasmids will be digested and linearized by both enzymes. An analysis of digestion of plasmid DNA samples obtained from four independent cell cultures with XhoI (lanes 2-5) and SacII (lanes 7-10) is presented in Figure 4.6.

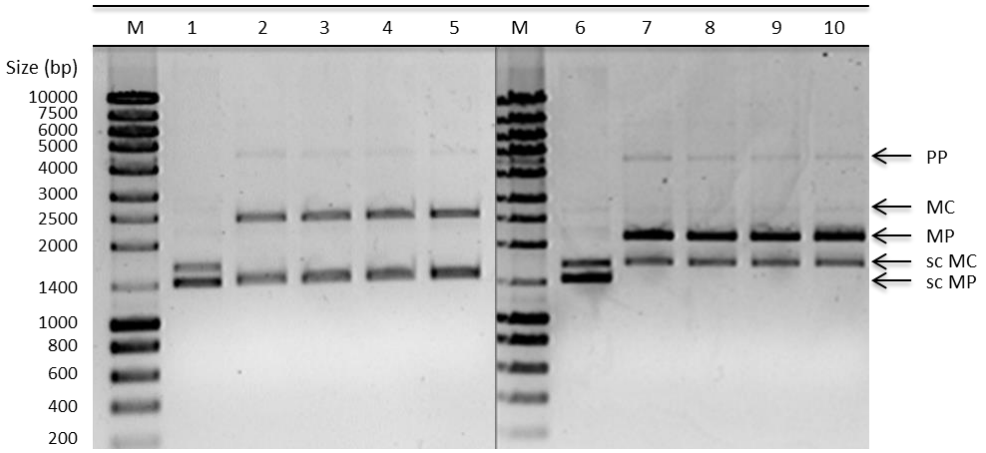

Figure 4.6: Agarose gel electrophoresis of pDNA (500 ng) from samples obtained after 2 hours of induction of recombination and digested with SacII and XhoI for 2 hours at 37°C. Lane M – molecular weight marker; lanes 1 and 6 – non-digested sample; lanes 2 to 5 – samples digested with XhoI; lanes 7 to 10 – samples digested with SacII. PP – linearized parental plasmid; MC – linearized minicircle; MP – linearized miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

The gel presented on Figure 4.6 shows that band identity is consistent with the expected behavior. When subjected to digestion with XhoI, the band of higher molecular weight visible on the non-digested sample (i.e. the sc MC band) is linearized, originating a band with a size equivalent to linearized MC (2457 bp), whilst no enzymatic activity is observed on the band attributed to the sc MP (lanes 2-5). On the contrary, upon digestion with SacII the band of lower molecular weight (i.e. the sc MP band) is digested, originating a band with a size equivalent to linearized MP (2106). As for the tenuous band in the digested samples at approximately 4.5 kbp, there is no difference in the migration pattern obtained with SacII or XhoI. This indicates that both enzymes result in equal linearization of this pDNA species, resulting in the conclusion that it is indeed representative of parental plasmid. The presence of this band indicates however that the recombination efficiency did not reach 100% and thus that un-recombined parental plasmid is still present after the 2 hours of recombination. To quantify the recombination efficiency, a densitometry analysis of the agarose gel was performed on the basis of the intensity of linearized PP and linearized MP bands by resorting to Equation 3.1, presented in section 3.5. The rationale behind the use of these bands for this calculation is due to the fact that both the parental plasmid and miniplasmid possess a replication origin and will thus continue to replicate during cellular growth, whilst the minicircle has a non-replicative nature. Thus, adding un-recombined parental plasmid and miniplasmid formed during recombination will result in the determination of initial parental plasmid [32, 39]. With this purpose, the bands obtained after enzymatic digestion with SacII of samples with 2 hours of recombination (to account for both miniplasmid isoforms originated after recombination induction and present in the purified

samples) from 30 independent cellular growths were analyzed, being determined that recombination event presented an efficiency of 96.6% ± 2.6%. This value is significantly higher than the values obtained when the ParA is expressed under the control of P_{BAD}/*araC* promoter from a low-copy number plasmid (50%) [38]. It is also higher compared with the values obtained when using a P_{BAD}/*araC-RBS-parA* cassette with the original ribosome binding site (70%) [70]. These results indicate that the insertion of the *parA* gene under the control of the P_{BAD}/*araC* promoter on the bacterial chromosome and the improvement of the RBS led to an efficient ParA resolvase expression upon recombination induction, allowing an almost complete recombination of parental plasmid in 2 hours after addition of L-arabinose.

Analysis of the minicircle/miniplasmid (MC/MP) ratio after 1 and 2 hours of recombination confirms that the second hour of recombination is mainly resulting in MP replication. In fact, the MC/MP ratio decreases from 0.7 ± 0.2 at 1 hour of recombination to approximately half of this value (0.4 ± 0.1) at the end of cellular growth (based on densitometry analysis of the respective band intensities from samples collected from 20 independent cellular growths). This result needs however to be validated by the determination of recombination efficiency at 1 hour after recombination induction. This parameter was not determined during the development of this work, being only calculated for four cellular growths that were purposely induced at a higher optical density (≈3.5) than the OD_{600nm}=2.5 previously determined as optimal ([71], in preparation). For these four cases, the recombination efficiency observed was 85.1% ± 2.0%. Although this value is lower than the one obtained in cultures in which the recombination event is analyzed 2 hours after induction (96.6% ± 2.6%), the fact that the induction was not performed at the appropriate cell growth phase is in itself a factor that would reduce recombination efficiency. However, this result shows that even under non-optimal induction conditions, the recombination efficiency after 1 hour of recombination is high. Taking into account the MC/MP plasmid ratio and provided that the purification method efficiently separates un-recombined parental plasmid from the sc MC, studies should be performed to determine if the total recombination time can indeed be reduced.

4.4 Primary purification and intermediate recovery

After fermentation and minicircle production, the resulting cells have to be disrupted in order to recover the therapeutically valuable minicircle from the intracellular space. With this purpose, a modified alkaline lysis method based on the protocol described by Birnboim *et al.* [72] was performed to disrupt cells and denature gDNA and proteins that were then removed by centrifugation [48, 63]. The following steps involved recovery of nucleic acids from the resulting clarified alkaline lysate (containing host cell proteins and gDNA not removed in the previous step, RNA, endotoxins, and pDNA [54]) by isopropanol precipitation, their resuspension in a Tris buffer and clarification of the solution by ammonium sulfate precipitation to reduce the impurity load. The reduction of the impurity load as well as pDNA concentration are crucial to prepare the solution for the subsequent purification steps and should typically result in a solution were pDNA accounts for at least 50% of all solutes [48].

Although an accurate value for impurities content was not determined, on the basis on studies using similar strategies for cell disruption and pDNA intermediate recovery [55, 57, 63], it is expected that the impurity content has been reduced when compared to the initial content in the alkaline lysate.

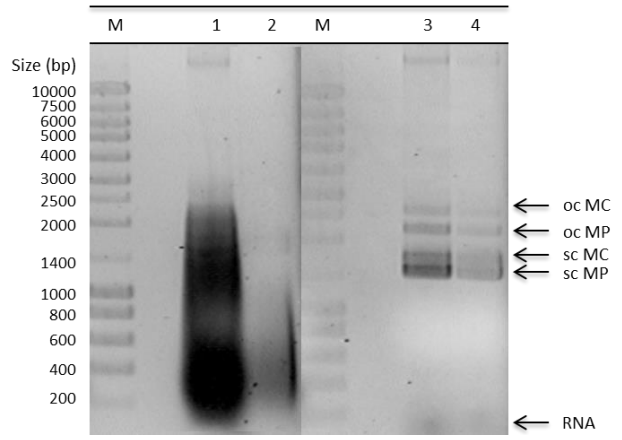

Figure 4.7: Agarose gel of samples collected from the clarified alkaline lysate (lanes 2 and 4, respectively 3 and 15 µL) and from the solution obtained by resuspension of the pDNA containing pellets precipitated with isopropanol (lanes 1 and 3, respectively 3 and 15 µL). Lanes 3 and 4 correspond to solutions obtained by incubation of samples on lane 1 and 2 with a RNase containing buffer for 2 h at 37°C. Lanes M – molecular weight marker; oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; scMP – supercoiled miniplasmid.

This was confirmed by performing an agarose gel electrophoresis analysis of samples of the clarified alkaline lysate (lanes 2 and 4 of Figure 4.7) and of the solution obtained after resuspension of the pDNA precipitated with isopropanol (lanes 1 and 3 of Figure 4.7). The solution obtained after addition of ammonium sulfate was not analyzed due to its high salt content. Given the high RNA content, 20 µL of the samples analyzed (lanes 1 and 2 on Figure 4.7) were incubated with 40 µL a RNase containing buffer (buffer 1 of the High Pure Plasmid Isolation Kit, Roche) for 2 hours at 37°C to facilitate the visualization of MC and MP bands on the gel (lanes 3 and 4 of Figure 4.7). Observation of lanes 3 and 4 shows the presence of open circular and supercoiled MC and MP species. Furthermore, as the sample volume used in the gel was the same (15 µL), it is also possible to observe an increase in concentration after resuspension of the pDNA-containing pellet obtained by isopropanol precipitation.

4.5 Miniplasmid enzymatic digestion

During the recombination event, two molecules of almost identical size are produced – a minicircle and a miniplasmid. In addition to this, both molecules are obtained in a supercoiled state. This renders the separation of these molecules a very difficult process. The purification method developed in this thesis relies on the enzymatic modification of the topology of the miniplasmid species to facilitate the desired separation. In particular, the nicking enzyme Nb.BbvCI is used to convert supercoiled miniplasmid molecules into the respective covalently closed open circular isoform by nicking one of the strands.

Nb.BbvCI is a DNA-nicking enzyme obtained by mutation of the restriction enzyme R.BbvCI [74]. This restriction enzyme acts on DNA substrates by duplex cleavage within a seven-base pair asymmetric

recognition sequence (CCTCAGC/GCTGAGG→ CC^TCAGC/GC^TGAGG). Enzymatic hydrolysis of both DNA strands requires the presence of two catalytic sites within the enzyme, one for hydrolyzing each strand. Most naturally occurring enzymes that act on duplex DNA present identical catalytic sites (homodimers or tetramers), comprising two identical subunits or assemblages of subunits, bound in symmetrical opposition, being that each subunit possesses one catalytic site. However, the existence of rare heterodimeric enzymes possessing two catalytic sites has made it possible to engineer new enzymes that act specifically only on one DNA strand of a double-stranded substrate, cleaving it and creating a nick on that DNA strand. The creation of these new enzymes is achieved by mutation of one of the catalytic sites in order to obtain an enzyme in which only the non-mutated site remains functional. The restriction enzyme R.BbvCI from *Bacillus brevis* is one of these rare enzymes in which the mentioned mutation was successfully performed on both catalytic sites, thus giving origin to two new enzymes – Nb.BbvCI that cleaves on the bottom strand (3'-5) and Nt.BbvCI that cleaves the top strand (5'-3'). The generation of these enzymes involved the identification of the subunits on the original enzyme (R_1 and R_2) with catalytic activity. Mutation of the respective catalytic site resulted in Nb.BbvCI (mutated on the R2 subunit – $R_1^+R_2^-$) and Nt.BbvCI (mutated on the R1 subunit – $R_1^-R_2^+$), both presenting strand and sequence-specificity and allowing DNA nicking rather than cleavage [74]. However, despite mutated and thus not active, the presence of both subunits is necessary for the catalytic activity of the non-mutated subunit [75].

Although the applications most addressed for this type of enzymes are in methods such as amplification of specific DNA segments, *in vivo* gene targeting and studies on DNA repair, among others [75], the use of such a specific nicking enzyme is extremely interesting for minicircle purification purposes.

The fact that the recognition site for the enzyme is located on the bacterial backbone of the parental plasmid assures that only these and miniplasmid molecules will be nicked upon digestion, whilst the minicircle will remain in a supercoiled state. This makes it then possible to explore differences on interactions between supercoiled and open circular molecules and chromatographic matrices.

The simplest method to perform the enzymatic digestion of the miniplasmid and parental plasmid is to directly add Nb.BbvCI to the solution obtained after re-suspension of the pDNA-containing pellets generated by isopropanol precipitation. With this purpose, two strategies were tested: re-suspension of pellets with 10 mM Tris-HCl (pH 8) and re-suspension with the diluted reaction buffer of the enzyme (1X NEB2). The first test was performed by resuspending the pellets in 10 mM Tris-HCl (pH 8) and adding 1 unit of Nb.BbvCI per µg of total DNA. Given that direct loading of digestion mixtures in agarose gel did not allow a clear visualization of the MC and MP bands due to the large amount of RNA (lanes 1 and 2 in Figure 4.8), 20 µL samples were incubated for 2 h at 37°C with 40 µL of a RNase-containing buffer (buffer 1 of the High Pure Plasmid Isolation Kit, Roche) and run in the gel. Observation of the corresponding lanes (lanes 3 and 4 of Figure 4.8) shows that Nb.BbvCI did not digest the sc miniplasmid, as the corresponding band (at ≈1.5 kbp) is present on the gel.

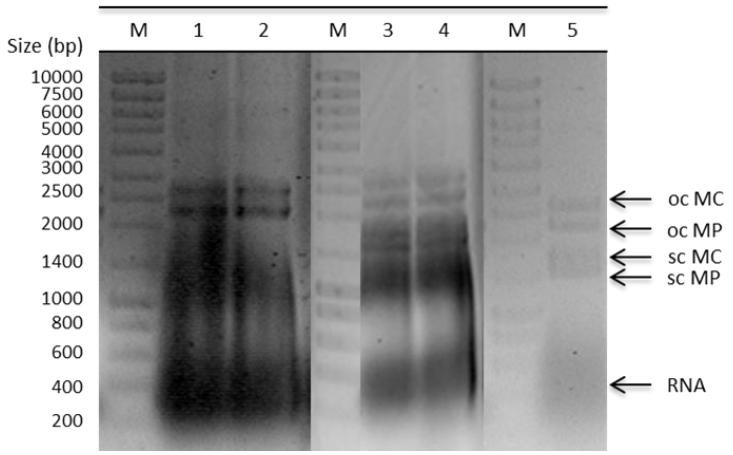

Figure 4.8: Agarose gel of Nb.BbvCI digestion of DNA in solutions obtained after isopropanol precipitation. Lanes M – molecular weight marker; lanes 1 and 2 – samples (2.1 µL) of pellets resuspended in 10 mM Tris-HCl (pH 8) digested overnight at 37°C with Nb.BbvCI; lanes 3 and 4 – samples (5 µL) obtained after overnight incubation at 37°C of 20 µL of solutions corresponding to lanes 1 and 2 with 40 µL of a RNase containing buffer; lane 5 – sample (1 µL) obtained after overnight digestion with Nb.BbvCI of pellets resuspended in the reaction buffer (1X NEB2) followed by 2 hours of incubation with a 20 µL of a RNase containing buffer at 37°C. oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

As an alternative strategy, the digestion was performed in solutions obtained after isopropanol precipitation and pellet resuspension in the enzyme reaction buffer. In this case, a small fraction (1.5 mL) of a clarified lysate was treated with 0.7 volumes of isopropanol and the resulting pellet was resuspended in 20 µL of 1X NEB2. To this mixture, 10 units of Nb.BbvCI were added and digestion was performed overnight at 37°C. In the following morning 20 µL of the RNase-containing buffer were added and the sample was incubated for 2 hours at 37°C. Observing lane 5 in Figure 4.8 it is possible to conclude that this strategy was also unsuccessful since the sc MP band is still present.

In view of the previous results, the decision was made to pre-purify the solutions obtained after isopropanol precipitation with 2.5 M ammonium sulfate. Following removal of precipitated impurities by centrifugation, supernatants were diafiltered with MilliQ water and concentrated in Amicon filter units (see section 3.7) in preparation for digestion. Although a concentration of the nucleic acids on the supernatant was also an objective in order to obtain a smaller volume to treat by enzymatic digestion and thus have a more economic and efficient use of the enzyme, the main reason for performing this step was the necessity of a salt-free solution on the subsequent step (miniplasmid enzymatic digestion).

The diafiltration and concentration of the solution obtained after ammonium sulfate precipitation resulted in concentrated, salt-free solutions containing the DNA (lanes 5 and 6 in Figure 4.9). Furthermore, no losses to the permeate were observed (lanes 1 to 4 in Figure 4.9).

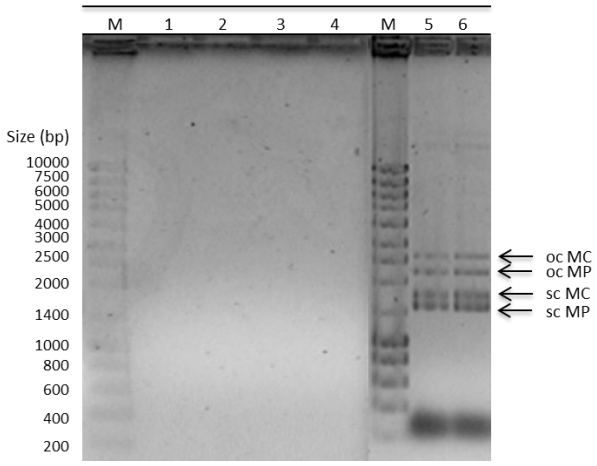

Figure 4.9: Agarose gel of permeates obtained during diafiltration of salt conditioned solutions with nucleic acids and samples obtained after concentration of nucleic acids in MilliQ water. Lanes M – molecular weight marker (5µL); lanes 1 and 2 – 20 µL of permeates obtained after the first centrifugation of Amicon units; lanes 3 and 4 - 20 µL of permeates obtained after the second centrifugation of Amicon units; lanes 5 and 6 – 500 ng of samples obtained by concentration after diafiltration. oc MC – opencircular minicircle; oc MP – opencircular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

Samples prepared by this method were then subjected to enzymatic digestion with Nb.BbvCI (1 unit of enzyme per µg of DNA). The first digestions performed involved the study of the enzyme activity over different time periods to determine the time of incubation needed to achieve complete digestion of MP and evaluate if long digestion periods could led to an un-specific enzymatic conversion of sc MC to oc MC. With this purpose, samples with 2 µg of nucleic acids were incubated at 37°C with 2 units of Nb.BbvCI and 5 µL of reaction buffer (1X NEB2) in a total digestion volume of 50 µL for 1, 2, 3, 4 and 5 hours and overnight (Figure 4.10). Observing the gel on Figure 4.10 is possible to conclude that in all samples subjected to digestion the sc MP was completely converted to the corresponding oc isoform without loss of sc MC. This indicates that long digestions can be used to assure the complete digestion of sc MP without degrading sc MC.

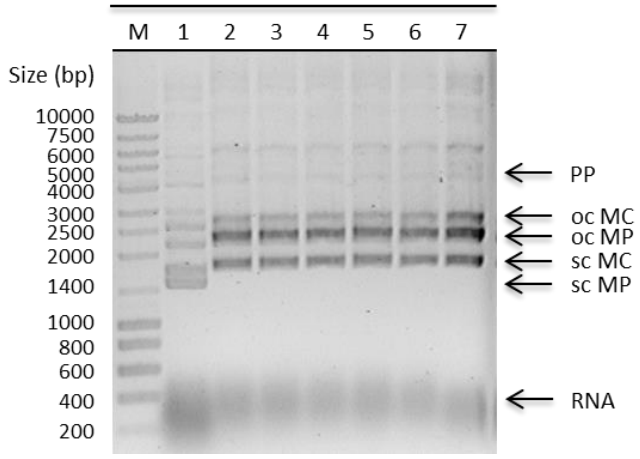

Figure 4.10: Agarose gel analysis of pre-purified DNA samples that were digested with Nb.BbvCI. Lane M – molecular weight marker (5 µL); lane 1 – non-digested sample (5 µL). Samples (50 µL) were digested for 1 h (lane 2), 2 h (lane 3), 3h (lane 4), 4 h (lane 5), 5 h (lane 6) and overnight (lane 7). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

The digestion mixtures analyzed in the gel of Figure 4.10 contained 1 unit of Nb.BbvCI per µg of nucleic acids. However, given the fact that the solution has other DNA molecules besides the MP over which the enzyme will not present enzymatic activity due to absence of recognition site, the amount of enzyme used can be significantly reduced. The determination of a combination of enzyme amount and digestion volume that result in the desired complete digestion is particularly relevant when solutions present a high concentration of nucleic acids. If digestions are performed with a ratio of 1 enzyme unit per µg of DNA in these cases, a large and unnecessary amount of Nb.BbvCI would be used resulting in an non-wise use of a costly material. Due to this, several attempts were made during the development of this work to achieve a complete sc MP digestion at the expense of the minimum enzyme. Given the results shown in Figure 4.10, digestions were incubated overnight to assure complete digestion of sc MP even with small amounts of enzyme. Four samples with ≈513 µg of nucleic acid content each were digested overnight with different amounts of enzyme (250, 200, 150 and 100 units of Nb.BbvCI) in 320 µL of total digestion volume at 37°C (Figure 4.11 A). The results obtained show complete sc MP digestion, thus allowing further reduction of enzyme. A second attempt (Figure 4.11 B) involved the digestion of three samples with approximately 1650 µg of nucleic acids in the presence of 100, 50 and 10 units of Nb.BbvCI and in a total digestion volume of 320 µL. In this last case, although the bands on gel are tenuous, is possible to observe the presence of sc MP in the samples digested with 10 units of enzyme (Figure 4.11 B, lane 9). The digestion was finally tested by subjecting a sample with 3.5 mg of nucleic acids to 50 units of Nb.BbvCI overnight at 37°C in a total digestion volume of 190 µL (Figure 4.11 C, lane 11), resulting also on complete MP digestion. This volume of enzyme seems thus be sufficient to digest samples with up to 3.5 mg of nucleic acids, provided that the reaction volume is maintained in the range of values tested.

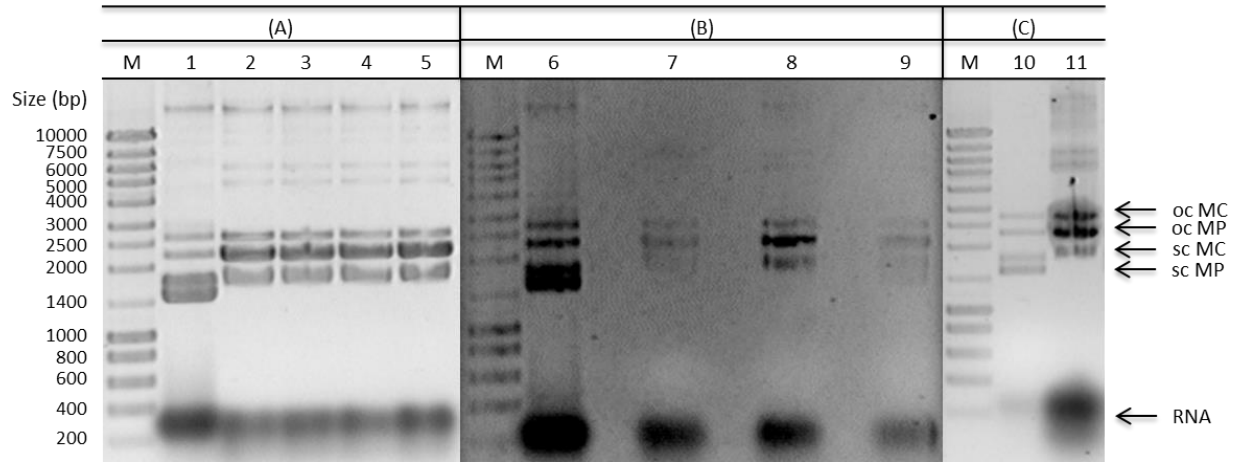

Figure 4.11: Effect of Nb.BbvCI amount on the overnight digestion of pre-purified DNA samples at 37°C. (A) Samples with 513 µg of DNA were digested with 250 units (lane 2), 200 units (lane 3), 150 units (lane 4) and 100 units (lane 5) of Nb.BbvCI. A non-digested, 2.2 µg DNA sample is shown as control (lane 1). (B) Samples with 1650 µg of DNA were digested with 100 units (lane 7), 50 units (lane 8) and 10 units (lane 9) of Nb.BbvCI. A non-digested, 1.8 µg DNA sample is shown as control (lane 6). (C) A sample with 3500 µg of DNA was digested with 50 units of Nb.BbvCI (lane 11). A non-digested, 490 ng DNA sample is shown as control (lane 10). Lanes M – molecular weight marker. oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

4.6 Purification by hydrophobic interaction chromatography

The nucleic acid species present on the samples obtained after Nb.BbvCI digestion are essentially sc MC, oc MC, oc MP and RNA (see Figure 4.11). The isolation of the target MC species from the nucleic acid impurities relied in a single step of hydrophobic interaction chromatography (HIC). The rationale for the choice of HIC as a purification method was based on the fact that several reports have attested the feasibility of using HIC with a mobile phase with high content of a kosmotropic salt (>2 M) to separate sc pDNA from oc forms [54, 67, 64]. A preliminary HIC run with a phenyl-Sepharose stationary phase and a mobile phase at 1.5 M ammonium sulfate in 10 mM Tris-HCl (pH 8) resulted, to some extent, in the separation between sc and oc isoforms in DNA containing samples (Figure 4.12 and Figure 4.13).

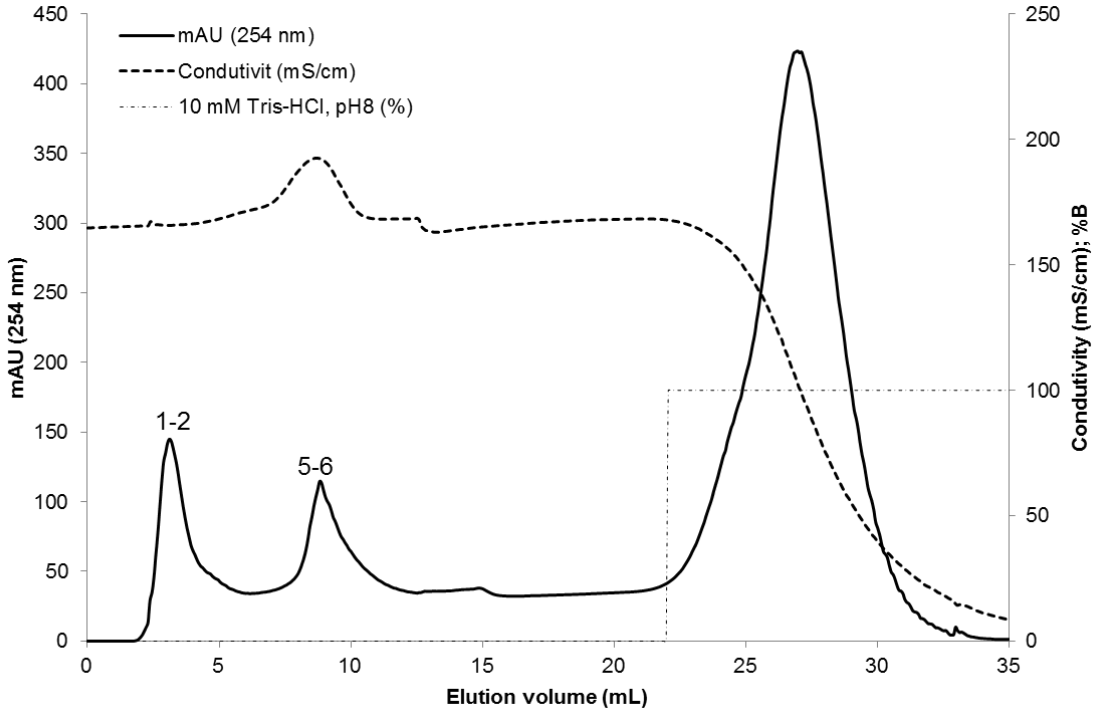

Figure 4.12: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms. 1 mL of a DNA sample obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation was loaded onto a phenyl-Sepharose column (100 mm x 10 mm). A step-wise elution was performed at 2.5 mL/min by changing the mobile phase from 1.5 M ammonium sulfate in 10 mM Tris-HCl, pH8 to 10 mM Tris-HCl, pH8. Peak elution fractions are indicated by numbers over the respective peak.

The agarose gel analysis of collected fractions (respectively numbered in Figure 4.13) allowed the identification of chromatographic peaks in Figure 4.12. Although the first eluted peak (fractions 1 and 2) presents both open circular and supercoiled isoforms, the fact that in fractions 3 to 9 apparently only supercoiled molecules are eluted gave an indication that this separation could be further explored to examine if a method relying on differences in hydrophobic interactions could be applied for minicircle purification.

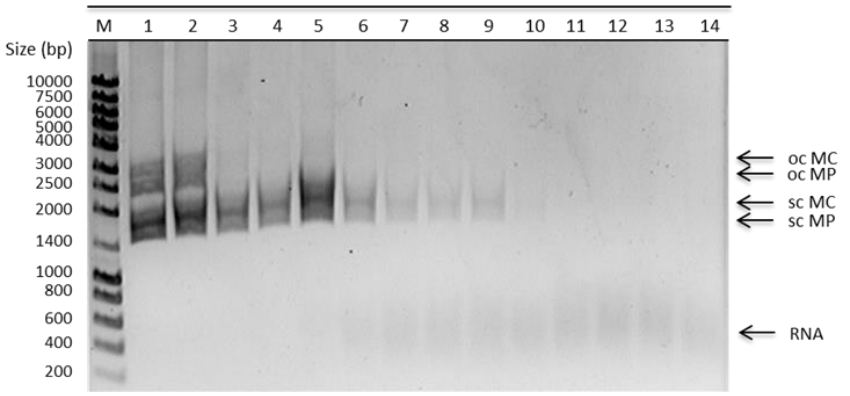

Figure 4.13: Agarose gel analysis of fractions collected during the chromatographic run presented on Figure 4.12. Lane M – molecular weight marker (5 µL); lanes 1 to 14 – samples (20 µL) of the respective numbered fractions collected during the chromatographic run. oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

In order to enable the desired separation, the correct chromatographic conditions had to be screened. Although analysis of Figure 4.13 shows that some fractions only present supercoiled isoforms (lanes 3 and 4), applying a chromatographic method relying on 1.5 M ammonium sulfate as elution buffer, after digestion of miniplasmid with Nb.BbvCI, would clearly result in a high loss of therapeutically valuable minicircle either by being co-eluted with open circular forms in the flowthrough or with RNA after change of mobile phase to 10 mM Tri-HCl pH 8. The obvious attempt to avoid this occurrence was the increase on hydrophobicity difference of open circular and supercoiled forms.

Hydrophobic interaction chromatography (HIC) is a chromatographic technique that has been widely used for purification of therapeutic proteins and has also found applications for pDNA purification from RNA, denatured gDNA and endotoxins due to higher hydrophobicity of these impurities when compared with the plasmid molecule [48, 54, 63]. These differences arise from the effect of the pDNA molecule topology on shielding the hydrophobic bases inside the double helix whilst the sugar-phosphate chains typically present on the sugar-phosphate backbone are available to establish hydrogen bonds with surrounding water molecules, thus resulting in a hydrophilic molecule. On the other hand, the single-stranded nature of denatured gDNA and RNA result on exposure of the hydrophobic bases and the portion of lipid A on endotoxins confer a hydrophobic character to these molecules [48, 67]. However, the hydrophilic nature of pDNA molecules can be circumvented by using stationary and mobile phases that promote the binding of pDNA molecules. With this purpose, a high concentration of a kosmotropic salt, such as ammonium sulfate or sodium citrate, can be used to promote water-ion and ion-macromolecule interactions due to their inherent size and surface charge. The addition of such salts will have a competitive effect with the ionic groups on the pDNA molecule resulting in water sequestration and dehydration of pDNA molecules. These effects have a direct influence on increasing compaction and promoting conformational changes on pDNA, resulting in the establishment of hydrophobic interactions between the pDNA and the ligands on the chromatographic support [48, 67].

The chromatographic method applied in this work involved the use of a chromatographic support with phenyl ligands (Phenyl Sepharose 6 Fast Flow, High Sub) being intended the determination of the

best chromatographic conditions, relying on a ammonium sulfate buffer as the mobile phase, to achieve an efficient and reproducible separation of supercoiled minicircle molecules from the remaining nucleic acids (including open circular DNA forms) and impurities in solution with minimal loss of the desired product.

4.6.1 Assays with non-digested samples

4.6.1.1 Chromatographic conditions screening

The influence of flow rate and conductivity on the elution behavior of open circular and supercoiled molecules, as well as of RNA, on non-digested samples obtained after recombination induction during bacterial growth, alkaline lysis, isopropanol precipitation and ammonium sulfate precipitation (as described in section 3) was studied by performing a chromatographic run with a linear gradient elution. The method, developed for the column described in section 3.9.1 (a column packed with 10 mL of resin) and implemented on a ÄKTApurifier100 system, consisted in column equilibration with 3 column volumes (CV) of 2 M ammonium sulfate in 10 mM Tris-HCl pH 8 (buffer A) at 2.5 mL/min, injection of a 1 mL sample conditioned to 2.5 ammonium sulfate and washing of unbound material with 2 CV of buffer A followed by a 5 CV linear gradient until 100% of 10 mM Tris-HCl, pH 8 (buffer B). During chromatographic run fractions of 1.5 mL were collected at the column outlet. The resulting chromatogram and agarose gel analysis of eluted fractions are presented in Figure 4.14 and Figure 4.15, respectively.

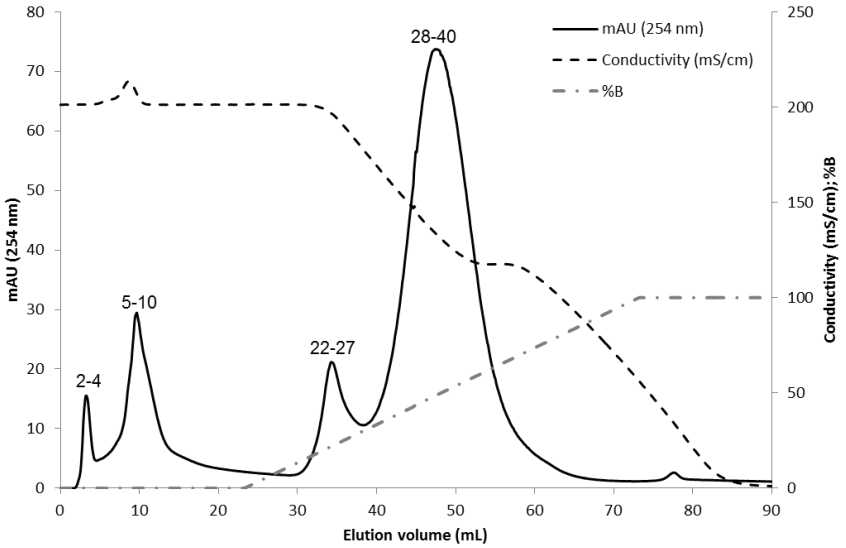

Figure 4.14: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a linear gradient elution scheme. 1 mL of a DNA sample obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation was loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound material was performed at 2.5 mL/min with 2 CV of 0% B and followed by elution with a linear gradient from 0 to 100% B in 5 CV. Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

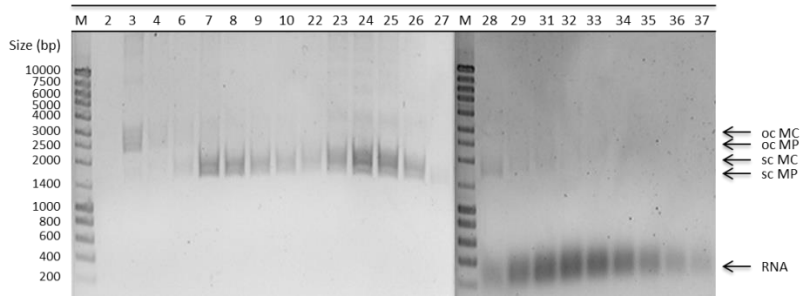

Figure 4.15: Agarose gel analysis of 10 µL from fractions collected during the chromatographic run presented in Figure 4.14. Lane M – molecular weight marker; numbered lanes correspond to fractions collected during the chromatographic run (lanes 2 to 4 – first peak; lanes 6 to 10 – second peak; lanes 22 to 27 – third peak; lanes 28 to 37 – fourth peak); oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

Analyzing Figure 4.15 it is possible to observe that a separation between open circular (lanes 3 to 6) and supercoiled DNA molecules (lanes 7 to 26), as well from RNA (28 to 37) is obtained to some extent. However, is possible to visualize tenuous bands of sc molecules on lanes 3 to 6 and on lanes 28 to 31. The presence of tenuous high molecular weight bands above sc bands is also present on lanes 23 to 25. In addition to the results of gel electrophoresis, the peaks of the chromatograms on Figure 4.14 are not well resolved (principally the third and fourth peaks). The main reasons for the low peak resolution can possibly be attributed to: (1) use of a high flow-rate that results in peak broadening due to slow diffusion of DNA molecules from the matrix support into the buffer during elution [65]; (2) use of an elution buffer with insufficient salt concentration [64]; (3) use of a short linear gradient that results in a continuous variation of conductivity that could be too abrupt, resulting in the co-elution of molecules that could probably be separated by using the appropriate conditions.

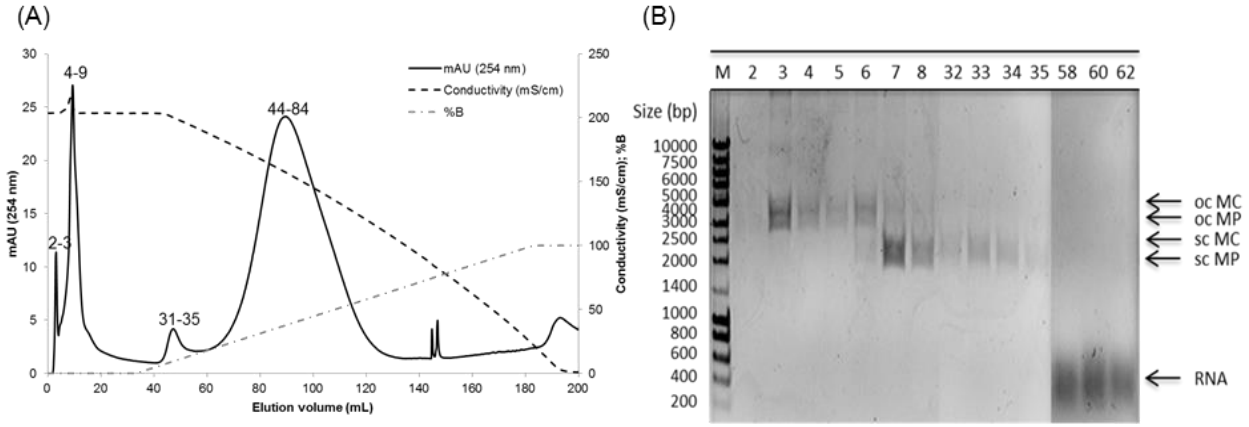

Figure 4.16: (A) Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a linear gradient elution scheme. 1 mL of a DNA sample obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation was loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2 M ammonium sulfate in 10 mM Tris HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound material was performed at 1.5 mL/min with 3 CV of 0% B and followed by elution with a linear gradient from 0 to 100% B in 15 CV. Numbers over peaks are indicative of the respective peak range fractions. (B) Agarose gel of 15 µL from fractions collected during the chromatographic run presented in (A). Lane M – molecular weight marker; numbered lanes correspond to fractions collected during chromatographic run (lanes 2 and 3 – first peak; lanes 4 to 8 – second peak; lanes 32 to 35 – third peak; lanes 58 to 62 – fourth peak); oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

Building up on these results, the flow-rate was decreased to 1.5 mL/min, the length of the washing step was increased to 3 CV to assure complete removal of open circular forms and the length of the linear gradient was increased to 15 CV. These modifications resulted in a broadening of peaks 3 and 4 without significant improvement on peak resolution (Figure 4.16, (A)). The pattern obtained by gel electrophoresis (Figure 4.16, (B)) of collected fractions is also similar to the one obtained with the first elution strategy (Figure 4.15). It is also important to refer that in this case, as well as for all chromatograms that will be presented along this discussion, the column feed was conditioned with 2.5 M ammonium sulfate before injection.

In a second attempt to resolve the chromatographic peaks, the use of a flow rate equal to 2 mL/min was tested and concentration of salt on buffer A was increased to 2.2 M ammonium sulfate (to increase differences between hydrophobicity of oc and sc DNA forms). The remaining chromatographic parameters (length of washing and linear gradient) were maintained as in the previous attempt. The resulting chromatogram and agarose gel electrophoresis of fractions collected during this run are presented in Figure 4.17 and Figure 4.18, respectively.

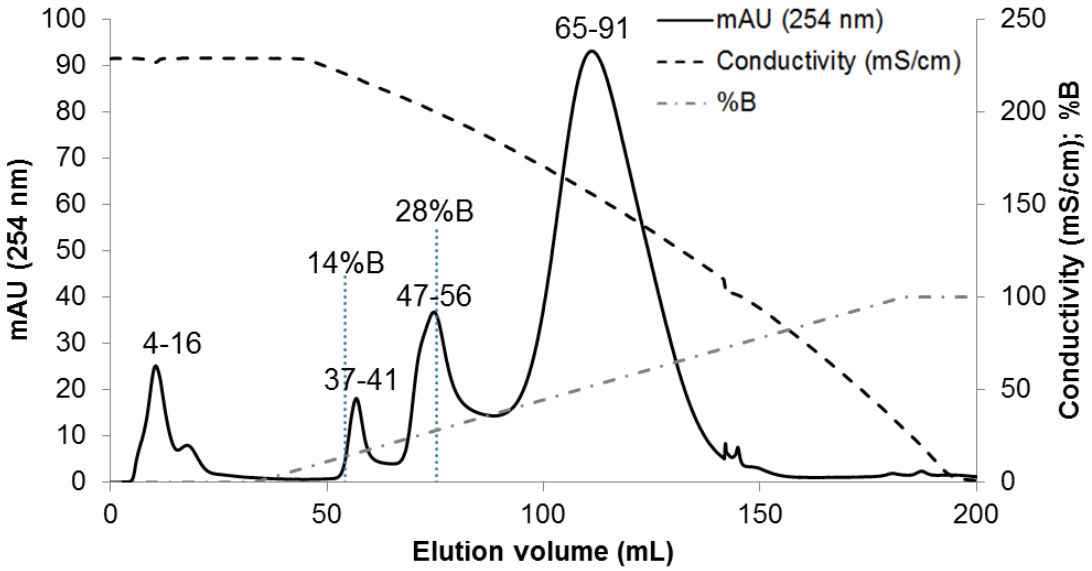

Figure 4.17: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a linear gradient elution scheme. 1 mL of a DNA sample obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation was loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound material was performed at 2 mL/min with 0% B for 3 CV, followed by linear gradient, at the same flow rate, from 0 to 100% B in 15 CV. Numbers over peaks are indicative of the respective peak range fractions. Numbers over dotted blue line indicates the % of B in the beginning of second peak (14% B) and at the maximum of the third peak (28% B). Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

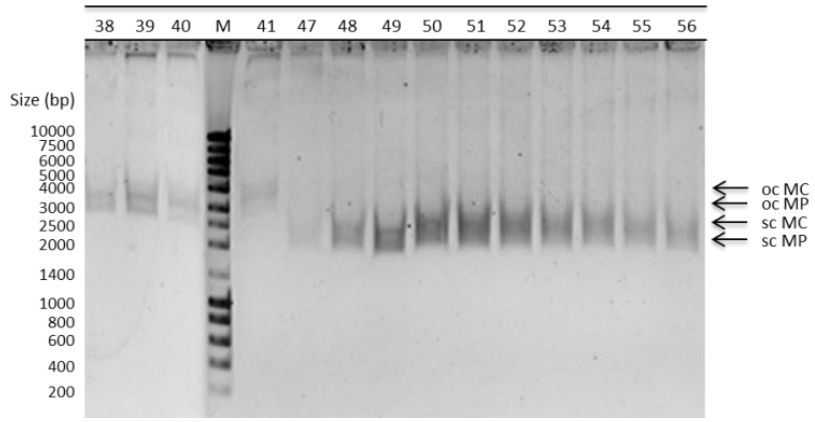

Figure 4.18: Agarose gel analysis of 15 µL from fractions collected during the chromatographic run presented in Figure 4.17. Lane M – molecular weight marker; numbered lanes correspond to fractions collected during chromatographic run (lanes 38 to 41 – second peak; lanes 47 to 56 – third peak); oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

Analyzing the results obtained up to this point, the last attempt (Figure 4.17 and Figure 4.18) yielded the best results. A reasonable resolution of peaks on the chromatogram was obtained without a too broad nature. Also, comparing the gel in Figure 4.18 with the two gels obtained before (Figure 4.15 and Figure 4.16 (B)) it is possible to notice that the co-elution of sc and oc forms that was visible before (lanes 3, 4 and 6 on Figure 4.15 and lanes 6 and 7 on Figure 4.16 (B)), is no longer observed. An apparent complete separation between oc and sc forms is noticeable, with oc molecules eluting only in fractions of the second peak and sc only in the third peak.

4.6.1.2 Establishment and optimization of a step-wise chromatographic method

The results obtained in the separation shown in Figure 4.17 were used as the basis for the development of a stepwise elution method for the desired purification. In order to establish the steps in the elution profile, the relation between the % of B and the peaks of oc (second peak, fractions 37 to 41) and sc isoforms (third peak, fractions 47 to 56) obtained with the linear gradient of the chromatographic run on Figure 4.17 was analyzed. The percentage of buffer B on the beginning of the second peak (14% B) and at the third peak (28%) were selected to define steps that would lead to the elution of oc and sc DNA forms, respectively. To elute any unbound species after sample injection (1 mL), the method was set to start with a washing segment of 2 CV of 0% of buffer B. Then, the percentage of buffer B was set to 14% for 2 CV after which a second step at 28% B for 2 CV was performed. Finally, for the elution of RNA and other impurities bound to the column, a third step was programed at 100% B for 3 CV. In between runs and column storage the column was further washed with MilliQ water until all parameters measured stabilized and conductivity reached a value bellow 0.08 mS/cm.

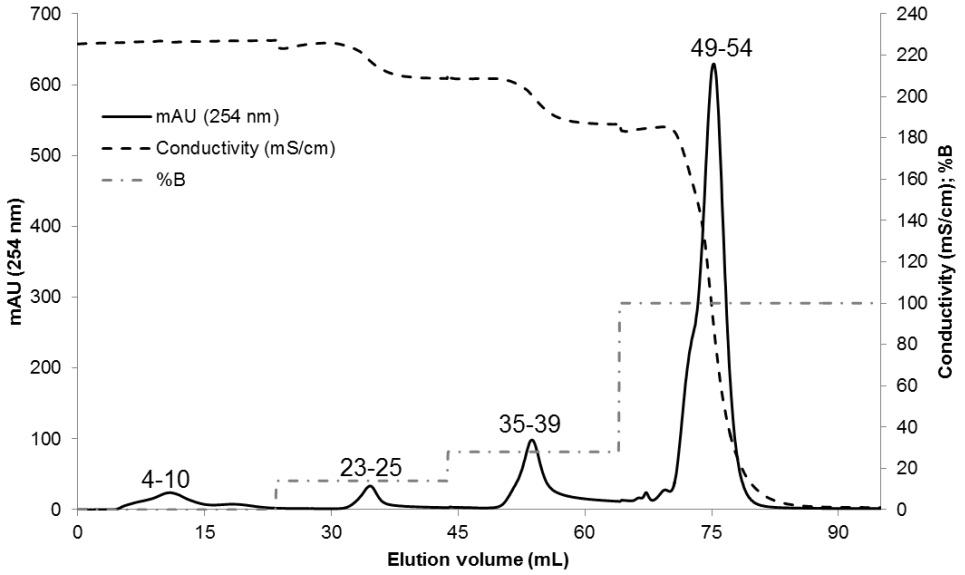

Figure 4.19: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a stepwise gradient elution scheme. 1 mL of a DNA sample obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation were loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound material was performed at 2 mL/min with 0% B for 2 CV, followed by a step-wise elution with 3 successive steps at 14% B (2 CV), 28% B (2 CV) and a 100% B (3 CV). Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

The mobile phase used for this chromatographic run was obtained by mixing 2.2 M ammonium sulfate in 10 mM Tris-HCl, pH 8 (buffer A) and 10 mM Tris-HCl, pH 8 (buffer B). The chromatographic run was performed at 2 mL/min (see chromatogram presented in Figure 4.19). Analysis of Figure 4.19 shows that sharp and resolved peaks were obtained using the elution strategy applied, with one peak obtained for each step, indicating that, according to the results obtained with the linear gradient, sc forms were being separated from oc DNA and RNA. To evaluate if this was the case, an agarose gel analysis of the fractions collected was performed (Figure 4.20).

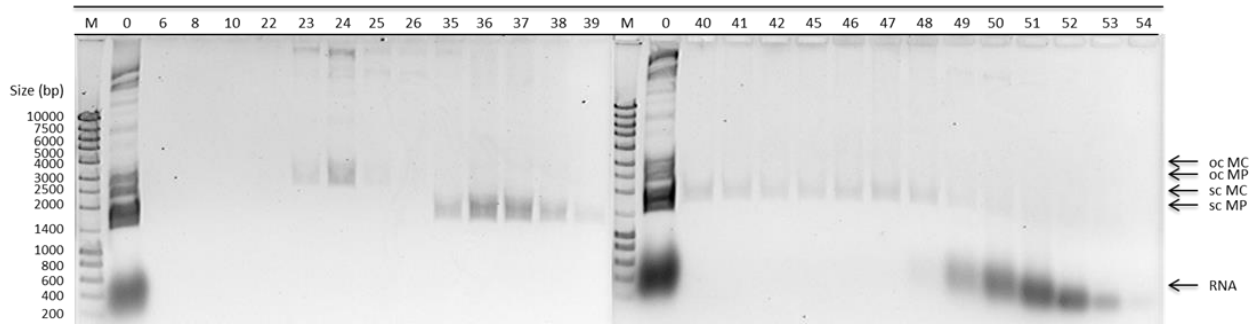

Figure 4.20: Agarose gel analysis of 10 µL from fractions collected during the chromatographic run presented in Figure 4.19. Lane M – molecular weight marker; Lanes 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lanes 6 to 10 – first peak; lanes 23 to 25 – second peak; lanes 35 to 39 – third peak; lanes 49 to 54 – fourth peak); oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

Observation of the gel presented above shows that no bands are visible on the lanes corresponding to the first chromatographic peak (lanes 6 to 10), meaning that the absorbance measured can be probably attributed to the elution of proteins still present in the injected sample instead of DNA molecules. Regarding the fractions from the second and third peak, the migration of DNA forms on the gel confirms that the desired separation is occurring, with oc DNA forms eluting on the second peak and sc forms eluting in the third peak. However, and although RNA is only eluted after fraction 39, sc molecules also elute in this last peak (lanes 49 to 53) and in fractions between the two last peaks.

At this stage, the samples injected in the column had not been subjected to enzymatic digestion, so only separation of oc and sc forms of both MP and MC DNA was being observed. Nevertheless, and considering the results obtained, it was clear that the desired separation of DNA isoforms was occurring and that the method was valid. Optimization of the method was thus the next objective.

The following chromatographic runs were performed with elimination of the first step by changing the percentage of buffer B used to equilibrate and wash the column, with the goal of determining the conditions that would avoid the binding of oc DNA forms to the column, which would thus be eliminated in the flow through. The percentage of B during washing was thus changed to 14% B on the basis of the results shown in Figure 4.19. Additionally, given the fact that sc molecules were not only eluted during the third peak of chromatogram in Figure 4.19, being co-eluted with RNA during the fourth peak, a set of experiments were performed with the objective of determining the appropriate percentage of buffer B to elute sc forms. Thus, tests were performed using with 32, 35 and 40% B during the first elution step (corresponding to the second step in Figure 4.19). The comparison of the three chromatograms obtained is presented in Figure 4.21.

Observing Figure 4.21 it is visible that the change of the percentage of buffer B to 40% in the first elution step leads to a clear loss of peak resolution. Regarding the other two values used, at least by observing the chromatogram, the separation obtained appears to be comparable. To better analyze the DNA forms eluted in the three chromatographic runs, samples from relevant peak fractions were run on agarose gel (Figure 4.22). Analyzing the gels presented in Figure 4.22 it is possible to observe that an increase in the percentage of B in the first step from 32% (gel A) and 35% (gel B) to 40% (gel C) resulted in the co-elution of RNA in the second peak (see fractions 24 and 25). Taking into account the chromatogram in Figure 4.21, this was already expected due to the increase in absorbance observed between the third and fourth peaks. Oc forms are observed in the fractions corresponding to the third peak when using 32% of B (fractions 23-25 in gel A), which are not visible when using 35 % B (gel B). the presence of these bands in gel (A) indicates that an optimization of the percentage of buffer B used during washing was still needed.

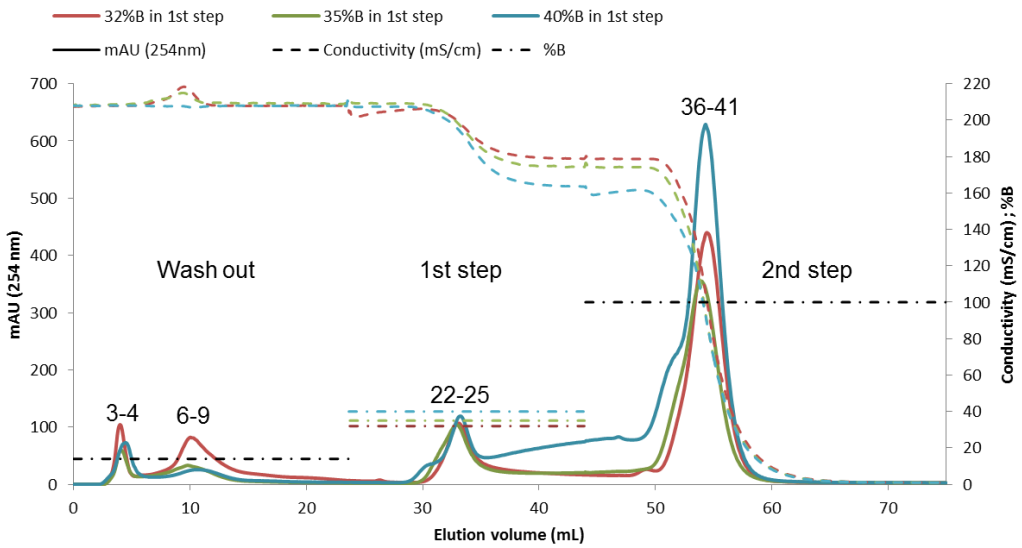

Figure 4.21: Optimization of hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a stepwise gradient elution scheme. 1 mL of a DNA sample obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation were loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound material was performed at 2 mL/min with 14% B (2 CV), followed by a step-wise elution at 32-40% B (2 CV) and 100% B (3 CV). In the first step three percentages of buffer B were studied: 32% B (red), 35% B (green), 40% B (blue). Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

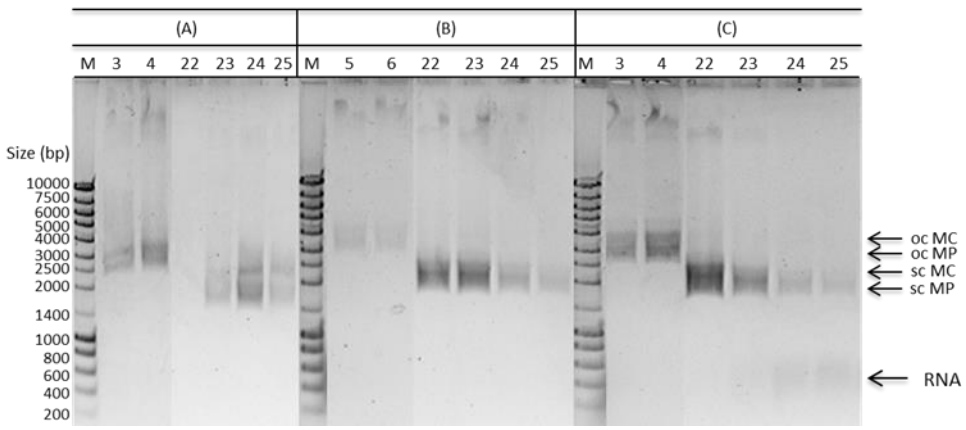

Figure 4.22: (A): Agarose gel analysis of 20 µL from fractions collected during the chromatographic run with first step at 32% B (Figure 4.21, red). Lane M – molecular weight marker; numbered lanes correspond to fractions collected during chromatographic run (lanes 3 and 4 – first peak; lanes 22 to 25 – third peak). (B): Agarose gel of 20 µL from fractions collected during the chromatographic run with first step at 35% B (Figure 4.21, green). Lane M – molecular weight marker; numbered lanes correspond to fractions collected during chromatographic run (lanes 5 and 6 – second peak; lanes 22 to 25 – third peak). (C): Agarose gel of 20 µL from fractions collected during the chromatographic run with first step at 40% B (Figure 4.21, blue). Lane M – molecular weight marker; numbered lanes correspond to fractions collected during chromatographic run (lanes 3 and 4 – first peak; lanes 22 to 25 – third peak). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

4.6.2 Assays with digested samples

4.6.2.1 Final optimization of the step-wise chromatographic method

In order to find the best percentage of buffer B during washing, three new chromatographic runs were performed at 16, 17 and 18% of B. To assure complete elution of oc forms before the elution of sc DNA, the length of washing was also increased from 2 to 4 CV. At this stage, samples used had been previously digested with Nb.BbvCI to convert sc MP molecules into the oc form. In this particular case, in all three chromatographic runs the sample injected (1 mL) contained 510 µg of nucleic acids. Furthermore, 0.5 mL fractions were collected from the third peak instead of the usual 1.5 mL fractions to improve the accuracy of the analysis. Also with this objective, from this stage on collected samples were subjected to micro-dialysis, as described in section 3.10, before agarose gel analysis.

From the analysis of the three chromatograms obtained (Figure 4.23) is possible to observe that all three percentages of buffer B during the washing result in sharper peaks than the ones obtained at 14% B. This is clear an indication that, as long as the eluted isoforms on these peaks are only the desired ones (open circular forms), all three percentages could be chosen for the final method. Another interesting aspect is the high similarity between the chromatograms - all peaks are superimposed in the three cases studied with minimum differences in the maximum absorbance. These results suggest a strong robustness of the method developed.

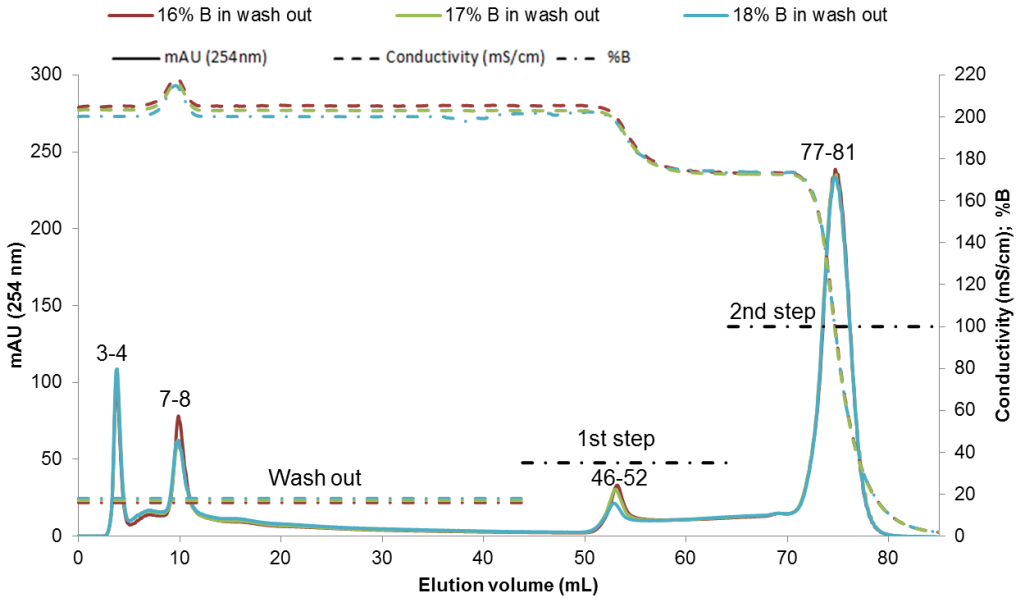

Figure 4.23: Optimization of hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a stepwise gradient elution scheme. 1 mL of DNA samples obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation were digested with Nb.BbvCI and loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound material was performed at 2 mL/min with three different percentages of buffer B for 4 CV, followed by a step-wise elution with a first step at 35% B (≈ 173 mS/cm) for 2 CV and a second step at 100% B (≈ 2 mS/cm) for 2 CV. Three percentages of buffer B during washing were studied: 16% B (red), 17% (green), 18% B (blue). Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

The agarose gel electrophoresis analysis in Figure 4.24 shows that performing the washing at either 16 (gel A) or 17% B (gel B) results in a good separation of oc and sc forms, with no sc MC bands observed in fractions 3, 7 or 8. Comparing the lanes corresponding to the first and second peaks of the chromatograms (lanes 3, 7 and 8) in gels (A) and (B) in Figure 4.24 it is visible that only oc forms of MP and MC are eluted. However, observing gel (C) in Figure 4.24 it is possible to detect a tenuous band of sc MC in lane 7. This indicates that a percentage higher than 17% of B should not be used during washing in order to avoid sc MC loss. Regarding the lanes corresponding to the third peak of the chromatograms (lanes 46 to 53 directly representing the fractions collected), although no change was made on the percentage of buffer B for the first step, the bands of eluted fractions do not have a consistent proportion in band intensity of oc and sc isoforms. In addition, gel (B) presents lanes with oc and sc MC bands (lanes 46 and 47) that are not visible in gel (A). Some oc MC forms appear thus to either interact with the ligands on the chromatographic matrix or result from sc MC relaxation during or after the chromatographic run.

The insertion of a new step between washing at 17% B and elution at 35% B did not result in the improvement of the separation, as neither oc MP or oc MC were eluted during this new step and co-elution of oc MC with sc MC during the step at 35% B was still observed (results not shown). However, the fact that oc MP is completely separated from the MC species is remarkable and prompted further investigation into the binding capacity of the matrix and applicability of the method for other minicircles and plasmids.

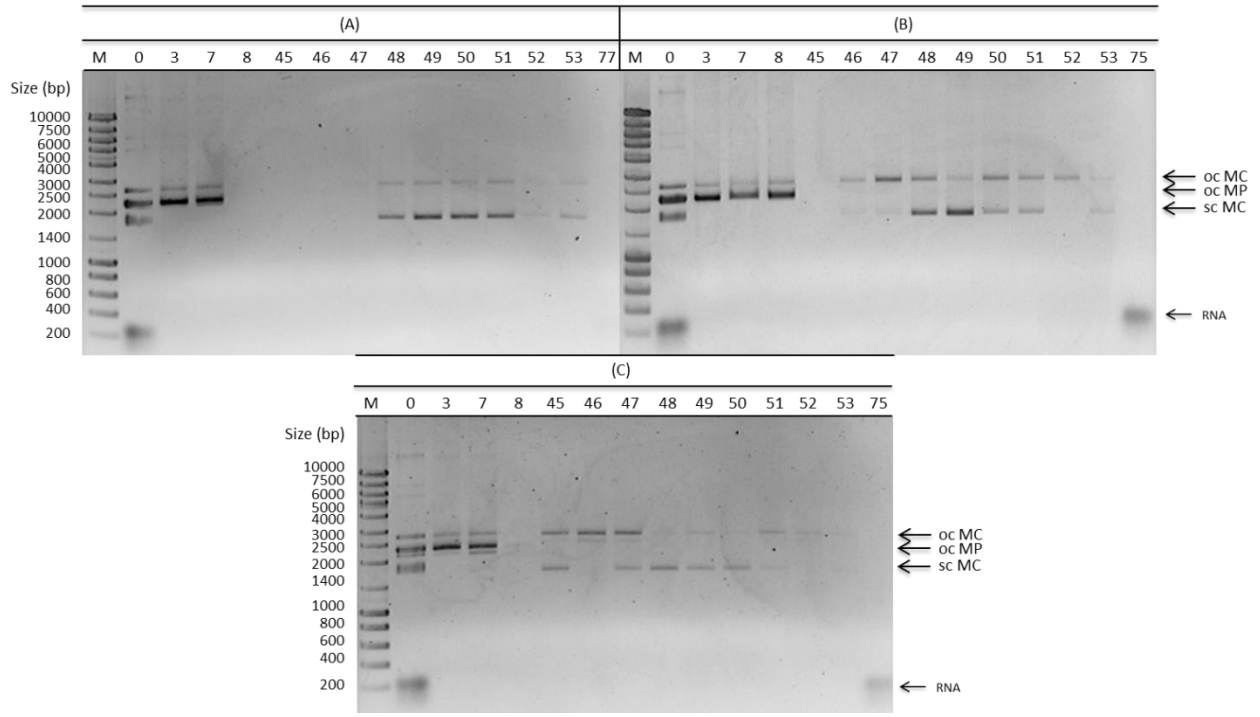

Figure 4.24: (A): Agarose gel analysis of 24 µL fractions collected during the chromatographic run with washing of unbound material at 16% B (Figure 4.23, red). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lane 3 – first peak; lanes 7 to 8 – second peak; lanes 46 to 52 – third peak; lane 77 – fourth peak). (B): Agarose gel of 24 µL fractions collected during the chromatographic run with washing of unbound material at 17% B (Figure 4.23, green). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lane 3 – first peak; lanes 7 to 8 – second peak; lanes 46 to 52 – third peak; lane 75 – fourth peak). (C): Agarose gel of 24 µL fractions collected during the chromatographic run with washing of unbound material at 18% B (Figure 4.23, blue). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lane 3 – first peak; lanes 7 to 8 – second peak; lanes 46 to 52 – third peak; lane 75 – fourth peak). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle; sc MP – supercoiled miniplasmid.

In view of the results obtained, the method used from this stage on involved an initial washing of unbound material at 17% of buffer B (≈ 204 mS/cm) for 4 CV, followed by a first step at 35% B (≈ 173 mS/cm) for 2 CV for MC elution and a final step at 100% B (≈ 2 mS/cm) for 2 CV to remove RNA and other strongly bound molecules from the matrix.

4.6.3 Confirmation of eluted DNA forms

To further confirm the identity of the bands visible in the agarose gel analysis and assure that the species eluted in each peak were the desired DNA forms, samples collected during HIC were dialyzed (as described in section 3.10) and digested with restriction enzymes chosen according to the MC and MP maps (Figure 4.2). With this purpose, three enzymes were chosen: (1) PvuII, which has no restriction sites on the MC and cuts the MP into six fragments (264, 337, 360, 364, 367 and 414 bp), (2) XhoI, which only has one restriction site on the MC and results in its linearization (2457 bp) and (3) BsrGI, which has restriction sites only on the MC that result in two fragments with 839 bp and 1618 bp. The enzymes XhoI and BsrGI were used in order to confirm the identity of the oc form that is co-eluted with the sc MC. As the

oc and linearized forms will migrate a similar distance on the agarose gel, using an enzyme which only has one restriction site in the MC would only allow confirming, without doubt the identity of the sc MC, whereas identification of the oc form could remain dubious. A similar result could be obtained with BsrGI as one of the fragments obtained from digestion of MC with this enzyme will result in a fragment of 1618 bp and would thus result in a band in the same region of the sc MC. With this purpose, these enzymes were used to digest the column feed sample, a sample from the first chromatographic peak and a third from the peak in which sc MC is eluted. The results obtained are presented in Figure 4.25.

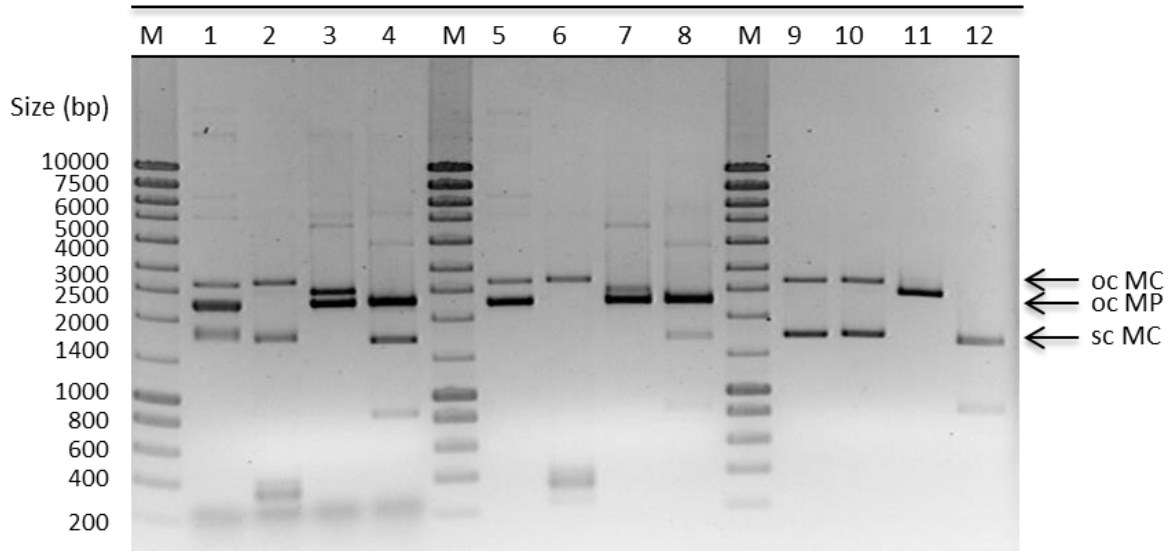

Figure 4.25: Restriction analysis of DNA species present in the column feed and in eluted fractions from the first (oc MC and oc MP) and third peaks (MC) obtained with the HIC method developed. Column feed (lane 1) was digested with PvuII (lane 2), XhoI (lane 3) and BsrGI (lane 4). A fraction from the first peak (lane 5) was digested with PvuII (lane 6), XhoI (lane 7) and BsrGI (lane 8). A fraction from the third peak (lane 9) was digested with PvuII (lane 10), XhoI (lane 11) and BsrGI (lane 12). Lanes M – molecular weight marker.

Analyzing the figure confirms that the conclusions appointed for the HIC separation given before were in fact correct. Observing the lanes corresponding to the column feed (1 to 4) shows that PvuII only digests the band appointed as oc MP (lane 2 at ≈2.4 kbp) resulting in fragments of low molecular weight, whilst the other two bands (1.7 and 2.6 kbp, approximately), previously identified as oc and sc MC, remain unchanged. Regarding digestions with XhoI and BsrGI (lanes 3 and 4, respectively), it is possible to confirm that there is no enzymatic activity on the MP (≈2.4 kbp) and that the bands at 1.7 and 2.6 kbp are digested. In the case of digestion with XhoI (lane 3), a new band at approximately 2.5 kbp appears which corresponds to linearized MC. In the case of digestion with BsrGI, two new bands appear at 843 and 1618 bp which correspond to the expected MC fragments. The bands at 200 bp in all these lanes (lanes 1 to 4) correspond to the presence of RNA. Similar observations are made when digesting DNA species from fractions of the first (lanes 5 to 8) and third (lanes 9 to 12) peaks: digestions with PvuII only affect the MP, whereas digestions with XhoI and BsrGI only affect oc and sc MC.

4.6.4 Method robustness evaluation

As a next step, the ability of the HIC step to separate oc and sc forms of two other minicircles ([71], in preparation) and of the plasmid pVAX-*LacZ* (6.05 kbp) was studied. Minicircles (pMINILi-hEF1α-VEGF-GFP and pMINILi-hEF1α(CpGfree)-VEGF-GFP) were, similarly to the one used in this work (pMINILi-CMV-VEGF-GFP), derived from the pMINILi and thus both harbor a restriction site for Nb.BbvCI. After recombination, pMINILi-hEF1α-VEGF-GFP (4476 bp) will originate a MP with 2106 bp and a MC with 2370 bp, while pMINILi-hEF1α(CpGfree)-VEGF-GFP (4125 bp) will originate a MP with 2106 bp and a MC with 2019 bp. The method by which these minicircles were recovered and pre-purified before digestion with Nb.BbvCI and HIC is, however, different from the method used with pMINILi-CMV-VEGF-GFP. Now, a commercial kit (NucleoBond Xtra Midi EF, MACHEREY-NAGEL), which relies on alkaline lysis for cell disruption and on an anion-exchange column for minicircle pre-purification were used. Furthermore this kit makes use of a RNase containing buffer, thus resulting in samples containing both oc and sc isoforms of MP and MC that are free from RNA. The use of RNase explains the absence of a peak during the step at 100% B in the respective chromatograms (red and blue lines in Figure 4.26). Regarding the plasmid pVAX-LacZ (green line in Figure 4.26, a sample containing oc and sc pDNA isoforms was obtained by application of methods described in section 3.6. Digestion with Nb.BbvCI was not performed on this sample.

Having as basis the elution patterns and agarose gel analysis performed previously with pMINILi-CMV-VEGF-GFP, the analysis of the chromatograms presented in Figure 4.26 suggests that the separation of oc and sc DNA forms is occurring for all the three samples studied. Regarding elution at 100% B, an RNA peak was observed only in the case of pVAX-*LacZ* for the reasons described above. An interesting observation is the fact that peaks are obtained around the same elution volumes as before, indicating that the method can be applied to plasmids with different sizes and is not restricted to the DNA forms for which it was initially developed. Although the other minicircles and miniplasmids studied present sizes similar to the recombination products of pMINILi-CMV-VEGF-GFP, the fact that separation of oc and sc forms can also be obtained for a significantly larger plasmid such as pVAX-*LacZ* is an indication that the method efficiency is not dependent on plasmid size.

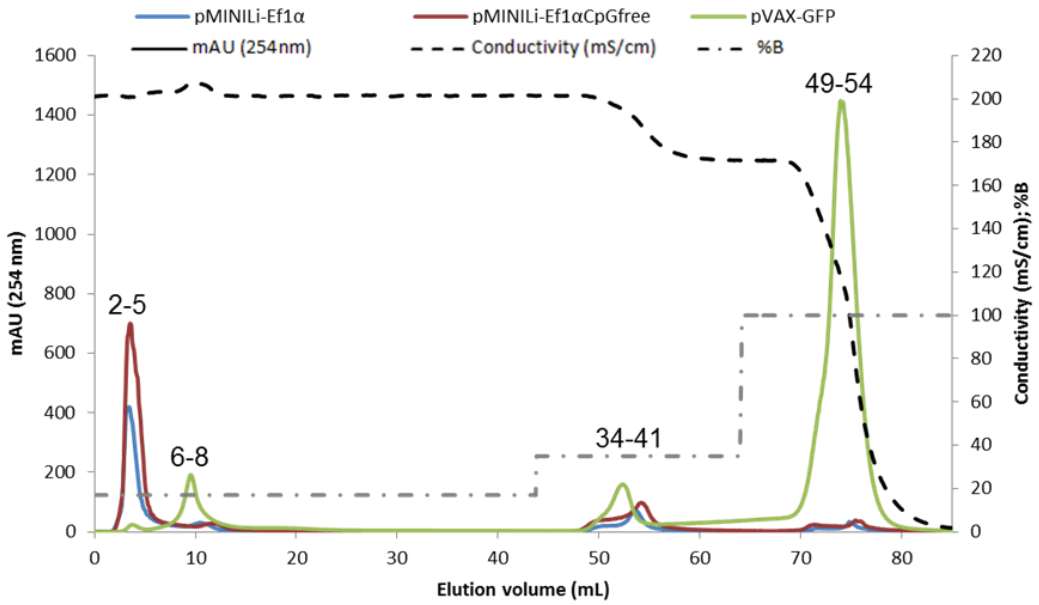

Figure 4.26: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a stepwise gradient elution scheme. 1 mL of DNA samples obtained after pre-purification were loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound sample was performed at 2 mL/min with 17% B (≈ 204 mS/cm) for 4 CV, followed by a step-wise elution with a first step at 35% B (≈ 173 mS/cm) for 2 CV and a second step at 100% B (≈ 2 mS/cm) for 2 CV. Method was applied to study three different samples: pMINILi-hEF1α-VEGF-GFP (blue), pMINILi-hEF1α(CpGfree)-VEGF-GFP (red), both previously digested with Nb.BbvCI, and pVAX-*LacZ* (green). Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

Fractions collected during the three runs were analyzed by agarose gel electrophoresis. The results obtained for the chromatographic run of the pVAX-*LacZ* sample are presented in Figure 4.27.

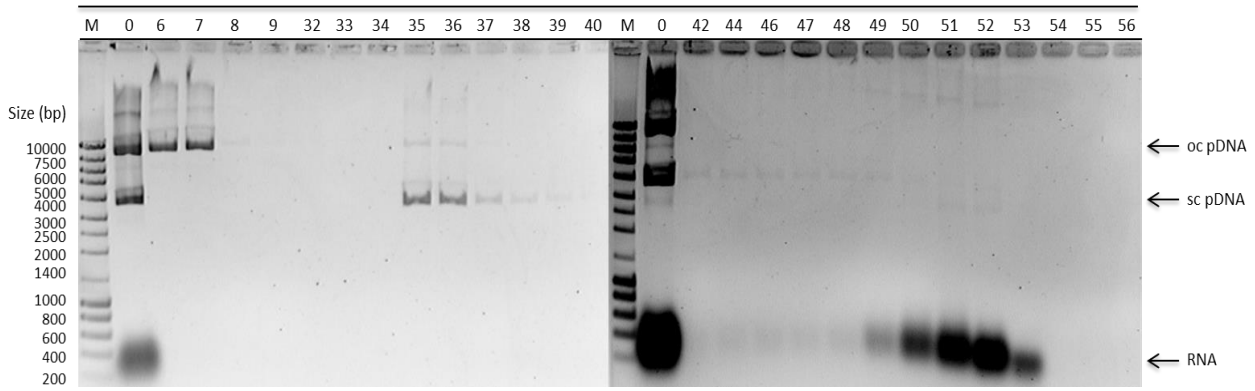

Figure 4.27: Agarose gel analysis of 20 µL from fractions collected during the chromatographic run of pVAX-LacZ, presented in Figure 4.26 (green line). Lane M – molecular weight marker; lanes 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run presented in green (lanes 6 to 8 – first peak; lanes 34 to 40 – second peak; lanes 49 to 54 – third peak). oc pDNA – open circular plasmid DNA; sc pDNA – supercoiled plasmid DNA.

Isolated oc pDNA can be observed in lanes 6 to 8, without co-elution of sc pDNA, similarly to what had been observed previously with pMINILi-CMV-VEGF-GFP. In addition, and also in line with previous observations, oc and sc pDNA are co-eluted in the peak at 35% B (lanes 34 to 40). Although sc pDNA is

also co-eluted with RNA in the fourth peak, the majority of sc pDNA is eluted at 35% B (lanes 35 and 36), indicating that the separation is reasonable. This fact is supported by densitometry analysis of the bands on the mentioned lanes - 90% sc pDNA (lane 35) and 93% sc pDNA (lane 36). This is considered a good homogeneity result in terms of the sc pDNA/total pDNA ratio (the minimum value recommended by FDA is 80% [27]).

The separation of pMINILi-hEF1α-VEGF-GFP and pMINILi-hEF1α(CpGfree)-VEGF-GFP minicircles shown in Figure 4.26 was analyzed by agarose gel electrophoresis (data not shown). Minicircle containing fractions (fractions 34 to 41 of chromatograms represented in red and blue on Figure 4.26) were selected on the basis of this analysis, pooled, desalted and concentrated with Amicon units and analyzed by electrophoresis (see lane 1 and lane 3 in Figure 4.28). To confirm band identity, a sample of concentrated fractions was digested with BsrGI. The differences between the two minicircles and pMINILi-CMV-VEGF-GFP reside only on the eukaryotic promoter, being the remaining sequence common to all three minicircles. Consequently, the two minicircles being analyzed in Figure 4.28 have a sequence map similar to the one presented in Figure 4.2, with the exception of the CMV sequence. This map reveals the existence of two restriction sites for BsrGI, one in the VEGF sequence and another at the end of the GFP gene. Thus, and similarly to what was observed for pMINILi-CMV-VEGF-GFP (Figure 4.25, lane 12), BsrGI digestion of either of the minicircles originates two fragments (lanes 2 and 4). According to the respective sequence maps (data not shown), a fragment of 839 bp should be obtained for both minicircles, as observable in lanes 2 and 4. The second fragment will have 1531 bp (pMINILi-hEF1α-VEGF-GFP) or 1180 bp (pMINILi-hEF1α(CpGfree)-VEGF-GFP), depending on the presence or absence of CpG motifs on the promoter. This is also confirmed in lanes 2 (band of ≈ 1.5 kbp) and 4 (band of ≈ 1.2 kbp). Densitometry analysis of lanes 1 and 3 shows that separation was better in the case of pMINILi-hEF1α-VEGF-GFP, being obtained a homogeneity of 69% for sc MC whilst in the second case 50% of oc and sc MC are present.

On the basis of the results shown in Figure 4.26 to Figure 4.28, is possible to conclude that the method developed is applicable to the separation of oc and sc DNA forms others than the ones for which it was originally developed, proving the method robustness and reproducibility.

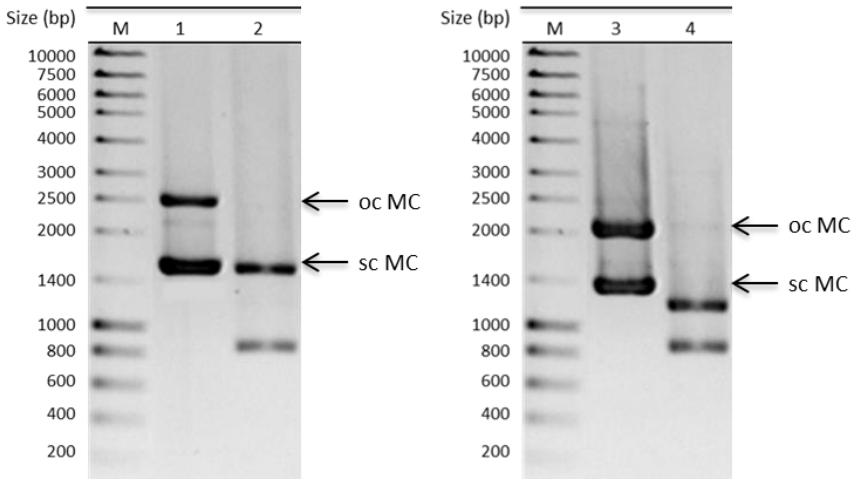

Figure 4.28: Agarose gel analysis of samples obtained by pooling and concentrating fractions 34 to 41 in the chromatograms shown in Figure 4.26. Lanes M – molecular weight marker; lane 1 – concentrated pool of pMINILi-hEF1α-VEGF-GFP; lane 2 - concentrated pool of pMINILi-hEF1α-VEGF-GFP digested with BsrGI; lane 3 – concentrated pool of pMINILi-hEF1α(CpGfree)-VEGF-GFP; lane 4 - concentrated pool of pMINILi-hEF1α(CpGfree)-VEGF-GFP digested with BsrGI. oc MC – open circular minicircle; sc MC – supercoiled minicircle.

4.6.5 Loading study

The final stage of the present work had as aim the study of the loading capacity of the Phenyl-Sepharose matrix. With this purpose, preliminary studies were performed with five samples with different nucleic acid content (ranging from 0.4 to 1.6 mg of total nucleic acids), which were digested with Nb.BbvCI, conditioned to 2.5 M ammonium sulfate, and loaded into the column. Reinforcing the conclusions drawn from Figure 4.23 and Figure 4.26, when samples with different amounts of nucleic acids (MC, MP and RNA) are injected into the column, and eluted with the method previously established and detailed in section 3.9.2, the resulting chromatograms present a highly similar elution pattern, with the main differences registered in peak intensities (Figure 4.29 and Figure 4.30).

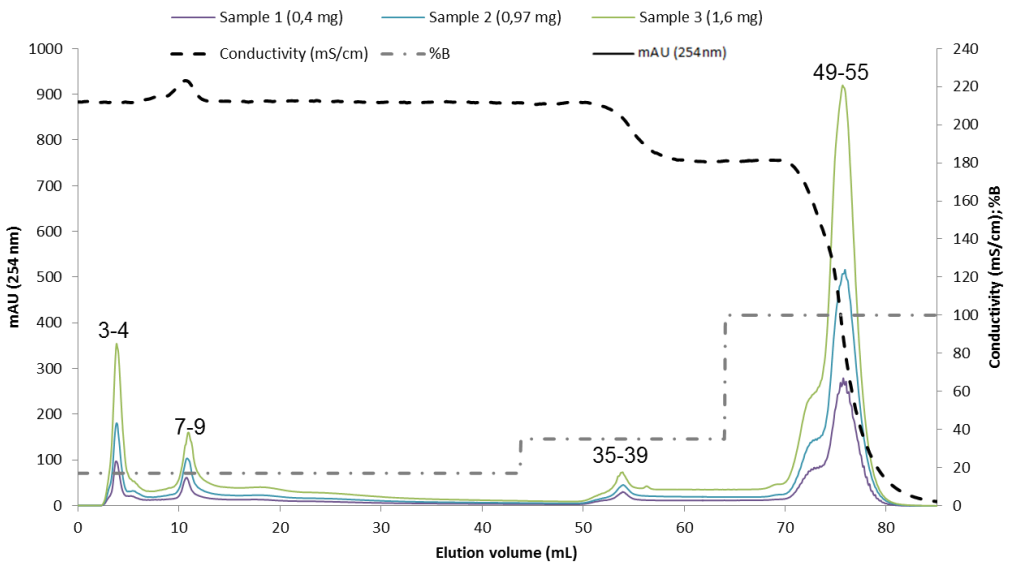

Figure 4.29: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a stepwise gradient elution scheme. 1 mL of DNA samples obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation were digested with Nb.BbvCI and loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound sample was performed at 2 mL/min with 17% B (≈ 204 mS/cm) for 4 CV, followed by a step-wise elution including 2 steps: first step at 35% B (≈ 173 mS/cm) for 2 CV and a second step at 100% B (≈ 2 mS/cm) for 2 CV. Three runs of samples with different nucleic acids amounts are represented: Sample 1 – 0.4 mg (purple line); sample 2 – 0.97 mg (blue line) and sample 3 – 1.6 mg (green line). Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

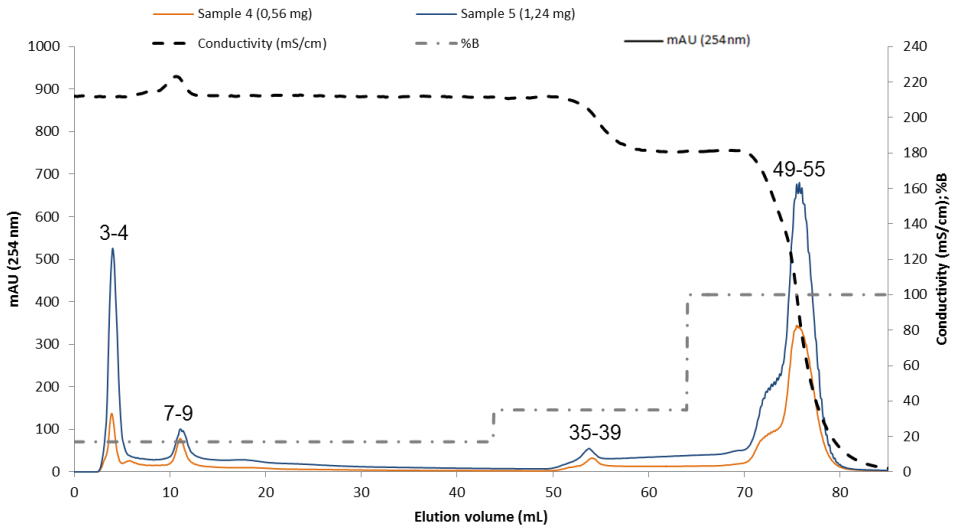

Figure 4.30: Hydrophobic interaction chromatography separation of covalently-closed DNA isoforms using a stepwise gradient elution scheme. 1 mL of DNA samples obtained after recombination, alkaline lysis and pre-purification with isopropanol and ammonium sulfate precipitation were digested with Nb.BbvCI and loaded onto a phenyl-Sepharose column (100 mm x 10 mm). Buffer A: 2.2 M ammonium sulfate in 10 mM Tris-HCl pH 8. Buffer B: 10 mM Tris-HCl pH 8. Washing of unbound sample was performed at 2 mL/min with 17% B (≈ 204 mS/cm) for 4 CV, followed by a step-wise elution including 2 steps: first step at 35% B (≈ 173 mS/cm) for 2 CV and a second step at 100% B (≈ 2 mS/cm) for 2 CV. Two runs of samples with different nucleic acids amounts are represented: sample 4 – 0.56 mg (orange line) and sample 5 – 1.24 mg (blue line). Numbers over peaks are indicative of the respective peak range fractions. Continuous line – mAU (254 nm); dashed line – conductivity (mS/cm); dashed dotted line – percentage of buffer B.

The results presented in Figure 4.29 and Figure 4.30 demonstrate the method robustness and high reproducibility as the exact same peaks are always obtained in the same elution regions, without deviation in peak fractions. To confirm this conclusion, an agarose gel analysis of all chromatographic peaks present on this figure for injections of higher mass (1.6 and 1.24 mg; respectively green line in Figure 4.29 and blue line in Figure 4.30) was performed. This analysis, shown in Figure 4.31, confirmed the separation of sc MC from RNA and oc MP forms.

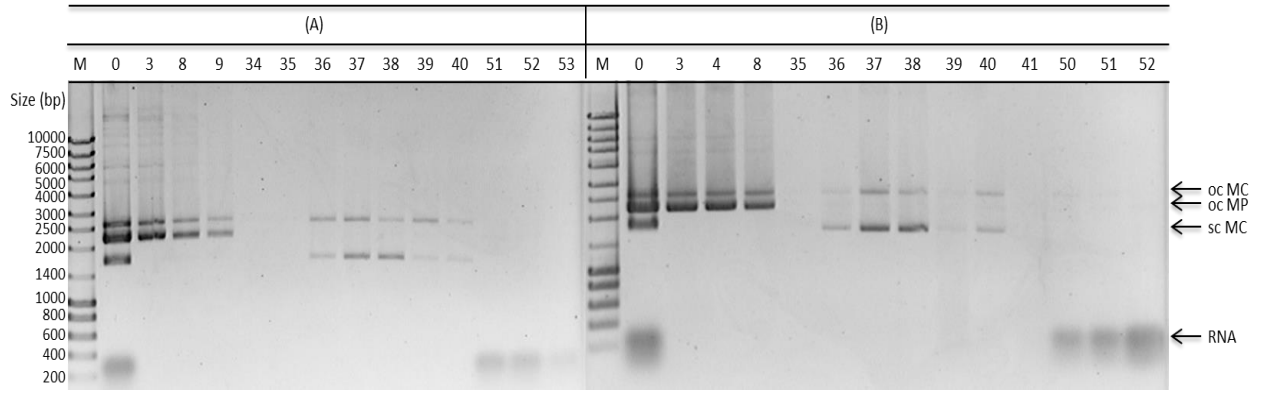

Figure 4.31: (A) Agarose gel analysis of 20 µL from fractions collected during the chromatographic run with injection of 1.24 mg of total nucleic acids, presented in Figure 4.30 (blue line). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lane 3 – first peak; lanes 8 and 9 –second peak; lanes 35 to 39 – third peak; lanes 51 to 53 – fourth peak). (B) Agarose gel of 20 µL from fractions collected during the chromatographic run with injection of 1.6 mg of total nucleic acids, presented in Figure 4.29 (green line). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lanes 3 and 4 – first peak; lanes 8 –second peak; lanes 35 to 39 – third peak; lanes 51 to 52 – fourth peak). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle.

Regretfully, it was not possible to determine the column dynamic binding capacity (DBC). An accurate and precise quantification of the performance of the chromatographic runs was not possible due to analytical problems with the HPLC column used for that purpose. In face of these difficulties, the only aspects possible to estimate are the percentage of sc MC in the purified fractions studied by gel electrophoresis and the percentage of minicircle recovered during the chromatographic runs by analysis of peak intensity. These values cannot however be viewed as accurate, and were only determined in order to provide an estimation of minicircle recovered and percentual content of sc isoforms in purified fractions.

Regarding the estimation of minicircle recovered, calculations were performed having as basis the chromatograms presented in Figure 4.29 and Figure 4.30. Resorting to the UNICORN software, the area of the third peak (corresponding to elution of MC isoforms) and total peak area of each chromatogram were considered. The resulting values indicate a MC recovery of 3.0% ± 0.9% from the total nucleic acids injected into the column. This small value can be explained by the low content of MC in the initial sample. With the available data, a correlation between total nucleic acids injected and MC recovered cannot be determined. Samples were originated not only from different cell cultures, as well as from pre-purification procedures and thus the sample injected in each run can present variations on nucleic acids content, rendering the establishment of a correspondence not valid.

In order to contextualize these values, an analysis of the DNA mass produced during fermentation is also needed. This parameter was intended to be determined by HPLC quantification and due to the aforementioned problems with the system, the only present alternative is an estimation based on the concentration of DNA obtained for samples collected when cellular growth was suspended (after 2 hours of recombination) and purified with a commercial kit (High Pure Plasmid Isolation Kit, Roche). These purified samples are RNA free and thus the mass determined is only relative to the MC and MP content (and eventually to un-recombined parental plasmid). The estimation involved densitometry analysis of lanes on agarose gels loaded with these samples for estimation of percentage of MC and MP. For this calculation, samples obtained from 20 independent 250 mL cellular growths were analyzed, being measured on agarose gel the intensity of bands of residual un-recombined parental plasmid (when present), miniplasmid and minicircle. The intensity obtained for each band was accounted and the total was considered correspondent to 100% of all nucleic acids. Percentage of MC was calculated with the respective band intensity in relation to the total band intensities obtained. These calculations resulted in an estimated value of 25.6% ± 6.3% for MC produced. Taking into account that these samples are RNA free and this type of nucleic acids account for the majority of the sample subjected to HIC (typically samples obtained after all pre-purification steps present a RNA content of ≈ 50% [48]), the estimations presented for HIC purification can be explained, to some extent, as a consequence of low MC production. In addition, a possible loss of MC during all the procedural steps to prepare the sample for purification has also to be considered. However, it is always important to highlight that these conclusions are only indicative as an accurate quantification of the DNA obtained after each procedural step was not possible to perform and only this could result in a correct and meaningful evaluation of the process developed.

For the relation of oc and sc MC, a densitometry analysis of agarose gels loaded with fractions from the third peak of each chromatogram presented in Figure 4.29 and Figure 4.30 was performed. In this case, the intensities of the oc and sc MC bands lanes of the gel corresponding to the third peak were evaluated. The percentage of sc MC was then calculated taking into account the intensity of the band of this isoform and total intensity of oc and sc MC bands in each lane. The estimated average value for sc MC was 70% with a standard deviation of 18%.

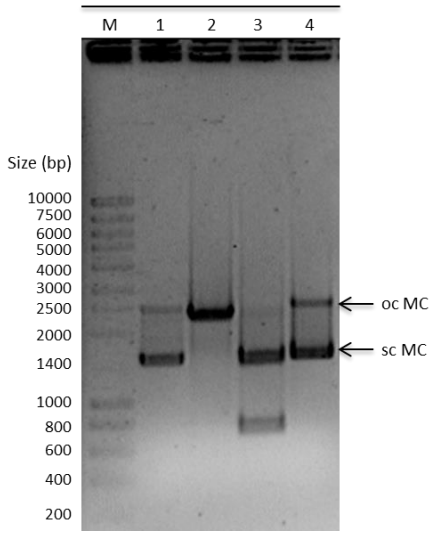

Figure 4.32: Agarose gel analysis of samples of the MC pool. Samples with 570 ng of DNA were digested with XhoI (lane 2), BsrGI (lane 3), PvuII (lane 4). A non-digested, 570 ng DNA sample is shown as control (lane 1). Lane M – molecular weight marker. oc MC – open circular minicircle; sc MC – supercoiled minicircle.

Samples prepared by pooling the MC fractions of three runs were also studied. In this case the samples used for injection into the HIC column presented a total mass of nucleic acids of 0.40, 0.44 and 1.24 mg and the pool of fractions from the three runs resulted in a sample with 9 mL of final volume. The pool obtained was subjected to diafiltration and concentration in a Amicon unit. The resulting concentrated sample presented a volume of 100 µL with a concentration of 286.25 ng/µL (measured in NanoVue Plus). These results are translated into an estimated MC recovery of 1.4% from the initial 2.08 mg of total nucleic acids injected into the column. This concentrated sample was also studied by agarose gel and digested with XhoI (which cuts the MC in one restriction site resulting in one fragment with 2457 bp), BsrGI (which cuts the MC into two fragments of 843 and 1618 bp) and PvuII (for which the MC harbors no restriction site), being the results presented in Figure 4.32. Analyzing this gel it is possible to observe the expected digestion pattern, being noticeable in lane 2 the digestion of sc MC with XhoI, from which results a single band at approximately 2.5 kbp (corresponding to the MC size). Digestion with BsrGI and PvuII also confirm bands identity by clear digestion of oc MC band and presence of the two expected bands in lane 3 (digestion with BsrGI), whilst in lane 4 it is clear that the enzyme (PvuII) had no enzymatic activity over DNA forms on the sample. A densitometry analysis of the non-digested sample (lane 1) allows the estimation of a percentage of 79.6% of sc MC in the sample, which is close to the minimum recommended by FDA [27] and can be considered a reasonable value. However the application of an additional step after HIC appears to be needed in order to increase product homogeneity.

5. Conclusions and future work

The main objective of the present work, i.e. the development of a robust and reproducible chromatographic method for minicircle purification, was successfully achieved. The process developed involves bacterial cell culture and recombination of parental plasmid into target minicircle, cell harvesting and alkaline lysis, precipitation with isopropanol and ammonium sulfate and diafiltration/concentration by microfiltration. Then, the key step in the process is performed which involves the selective modification of the structure of miniplasmid impurities by resorting to a nicking enzyme that recognizes a specific sequence in the molecule. As a consequence, supercoiled miniplasmids are converted into their open circular forms, whereas minicircles are unaffected. Finally, supercoiled and open circular minicircle species are isolated from miniplasmid and RNA by hydrophobic interaction chromatography using a downward step elution with ammonium sulfate. Agarose gel electrophoresis analysis confirmed that the method is able to produce minicircles that are virtually free from miniplasmid, parental plasmid and RNA.

Regarding the production stage, an optimization of the cell culture/recombination step is advised to increase MC production. A possible alternative could be the utilization of a richer growth medium such as Terrific Broth [76] or optimization of culture in bioreactor. This last hypothesis could, however lead to a decrease in the recombination efficiency and thus to a higher content of un-recombined parental plasmid at the end of fermentation. However, preliminary studies indicate that the HIC method developed is also able to separate parental plasmid species from minicircles. This observation could be explained by the fact that, having a restriction site for Nb.BbvCI, the parental plasmid will be nicked together with miniplasmid during enzymatic digestion. A clear and final confirmation of this separation would result in the addition of a positive characteristic to the method developed and be a promising factor for optimization of a bioreactor culture process. Admitting that the current production methodology is kept, another possible improvement in the cell culture/recombination step could be the reduction of time allowed for the recombination process as the analysis of samples obtained after one and two hours of recombination indicate that the second hour is only being used mainly for miniplasmid replication.

Given the preliminary results obtained in terms of the content of supercoiled minicircle in the purified samples, some measures are advised to increase of product homogeneity. In order not to add steps after HIC purification, the alkaline lysis conditions should be studied and optimized. During this step, a pH higher than 12.5 was measured after homogenization of the solution obtained by addition of lysis buffer. This occurrence could be one of the reasons for the high content of open circular DNA forms in samples to be subjected to purification and consequently could influence the homogeneity of the final product. If an improvement of this step is not successful, another possible alternative is the addition of a polishing step after HIC. For this, a size-exclusion chromatography (SEC) step could be considered and evaluated due to the fact that the conformational state of open circular molecules results in a larger

molecule. Sample homogeneity could be increased given that oc MC could probably be separated from the supercoiled minicircle during SEC, being eluted earlier than sc molecules during this chromatographic run. However, in order to maintain the simplicity of the method developed and avoid product loss in a new chromatographic step, the first hypothesis should be firstly examined.

For environmental concerns the study of alternative elution buffers, such as sodium citrate, for the hydrophobic chromatographic method here developed should be evaluated. This salt is more environmentally friendly and its use has already been reported as successfully in hydrophobic interaction chromatography (although at expense of product purity) [67]. Provided that method efficacy and reproducibility is maintained its use would be advisable, especially at large scale.

Another important aspect is related with the nicking enzyme used. This enzyme is a costly resource and a similar alternative should be studied and produced directly on the lab in order to reduce process costs. Ideally, an inducible production of an enzyme with this characteristic could be produced during cellular growth for miniplasmid relaxation in a similar strategy as used by Kay *et al.* [50] in a *E. coli* strain with expression of ϕC31 integrase and I-SceI homing nuclease. However, this strategy applied with a nicking enzyme would probably have an unsuccessful outcome as the cellular machinery could repair the nick on the miniplasmid chain and thus the final product would most probably present both supercoiled minicircle and miniplasmid.

An advantage of the use of hydrophobic interaction chromatography instead of an affinity-based chromatography is related with the cost-effectiveness of the process. Specifically addressing the matrix used in this work (Phenyl Sepharose 6 Fast Flow High Sub, GE), its cost (506€/200 mL of resin as of August 2014, from Fisher Scientific) is significantly lower than other two matrices considered for amino-acid affinity-based chromatography (479€/10 mL of arginine-separopore 4B (bioPLUS) and 253€/10 mL of L-histidine-separopore 4B-CL (BioPLUS) both as May 2014, from Biogen Cientifica), and thus its use in a laboratory-scale or, more importantly, in large-scale is comparatively more cost-effective. Comparing HIC with a chromatographic method using these affinity matrices, temperature constrains on the method efficacy are also a point of advantage as the affinity chromatography would have to be conducted at a specific and controlled temperature [58, 62] whilst the present HIC method was applied at room temperature without detection of significant variations on efficacy and reproducibility.

As final remark, is also relevant to mention that this method has been applied for transfection studies ([71], in preparation). Although in this case the pre-purification relied on a commercial kit that generates endotoxin-free samples, the application of HIC allowed the recovery of biological active minicircles that were successfully used for transfection. Regarding the complete method applied in this work, desalted purified minicircle samples were also used in agar antibiotic free plates, being verified that no colonies were formed after 5 days proving that samples were not contaminated with microorganisms.

6. References

[1] T. Friedmann and R. Roblin, "Gene Theraphy for Human Genetic Disease?," *Science.,* vol. 175, no. 4025, pp. 949-955, 1972.

[2] F. Meyer and M. Finer, "Gene Therapy: Progress and Challenges," *Cell Mol Biol.,* vol. 47, no. 8, pp. 1277-1294, 2001.

[3] M. Kay, "State-of-the-art gene-based therapies: the road ahead," *Nat Rev Genet.,* vol. 12, no. 5, pp. 316-328, 2011.

[4] John Wiley and Sons Ltd., "The Journal of Gene Medicine," [Online]. Available: http://www.abedia.com/wiley/. [Accessed October 2014].

[5] S. Rosenberg, P. Aebersold, K. Cornetta, A. Kasid, R. Morgan, R. Moen, E. Karson, M. Lotze, J. Yang, S. Topalian, M. Merino, K. Culver, D. Miller, R. Blaese and W. Anderson, "Gene Transfer into Humans - Immunotherapy of Patients with Advanced Melanoma, Using Tumor-Infiltrating Lymphocytes Modified by Retroviral Gene Transduction," *N Engl J Med.,* vol. 323, no. 9, pp. 570-578, 1990.

[6] M. Edelstein, M. Abedi, J. Wixon and R. Edelstein, "Gene therapy clinical trials worldwide 1989-2004 - an overview," *J Gene Med.,* vol. 6, no. 6, pp. 597-602, 2004.

[7] R. Blaese, W. Kenneth, A. Miller, C. Carter, T. Fleisher, M. Clerici, G. Shearer, L. Chang, Y. Chiang, P. Tolstoshev, J. Greenblatt, S. Rosenberg, H. Klein, M. Berger, C. Mullen, W. Ramsey, L. Muul, R. Morgan and W. Anderson, "T Lymphocyte-Directed Gene Therapy for ADA-SCID: Initial Trial Results After 4 Years," *Science,* vol. 270, no. 5235, pp. 475-480, 1995.

[8] J. Wolff, R. Malone, P. Williams, W. Chong, G. Acsadi, A. Jani and P. Felgner, "Direct gene transfer into mouse muscle in vivo," *Science.,* vol. 247, no. 4949, pp. 1465-1468, 1990.

[9] D. Tang, M. DeVit and S. Johnston, "Genetic immunization is a simple method for eliciting an immune response," *Nature.,* vol. 356, no. 6365, pp. 152-154, 1992.

[10] M. Kutzler and D. Weiner, "DNA vaccines: ready for prime time?," *Nat Rev Genet.,* vol. 9, no. 10, pp. 776-788, 2008.

[11] M. Edelstein, M. Abedi and J. Wixon, "Gene therapy clinical trials worldwide to 2007 - an update," *J Gene Med.,* vol. 9, no. 10, pp. 833-842, 2007.

[12] R. Gardlík, R. Pálffy, J. Hodosy, J. Lukács, J. Turňa and P. Celec, "Vectors and delivery systems in gene therapy," *Med Sci Monit.,* vol. 11, no. 4, pp. 110-121, 2005.

[13] K. Ponder, "Vectors of Gene Therapy," in *An Introduction to Molecular Medicine and Gene Therapy*, Bethesda, Wiley-Liss, Inc., 2001, pp. 77-112.

[14] U.S. Food and Drug Administration, "FDA," [Online]. Available: http://www.fda.gov/forindustry/userfees/prescriptiondruguserfee/ucm127153.htm. [Accessed September 2014].

[15] European Medicines Agency Science Medicines Health, "EMA," [Online]. Available: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002145/human_med_001480.jsp&mid=WC0b01ac058001d124. [Accessed November 2014].

[16] uniQure, "uniQure," [Online]. Available: http://www.uniqure.com/home/. [Accessed November 2014].

[17] T. Niidome and L. Huang, "Gene Therapy Progress and Prospects: Nonviral vectors," *Gene Ther.,* vol. 9, no. 24, pp. 1647-1652, 2002.

[18] K. Prather, S. Sagar, J. Murphy and M. Chartrain, "Industrial scale production of plasmid DNA for vaccine and gene therapy: plasmid design, production, and purification," *Enzyme Microb Technol.,* vol. 33, no. 7, pp. 865-883, 2003.

[19] H. Yin, R. Kanasty, A. Eltoukhy, A. Vegas, A. Dorkin and D. Anderson, "Non-viral vectors for gene-based therapy," *Nat Rev Genet.,* vol. 15, no. 8, pp. 541-555, 2014.

[20] J. Mairhofer and R. Grabherr, "Rational Vector Design for Efficient Non-viral Gene Delivery: Challenges Facing the Use of Plasmid DNA," *Mol Biotechnol.,* vol. 39, no. 2, pp. 97-104, 2008.

[21] A. Darquet, B. Cameron, P. Wils, D. Scherman and J. Crouzet, "A new DNA vehicle for nonviral gene delivery: supercoiled minicircle," *Gene Ther.,* vol. 4, no. 12, pp. 1341-1349, 1997.

[22] EMA, "Guideline on the non-clinical studies required before first clinical use of gene therapy medicinal products (EMEA/CHMP/GTWP/125459/2006)," European Medicines Agency, London, 2008.

[23] W. Walker, D. Porteous and A. Boyd, "The effects of plasmid copy number and sequence context upon transfection efficiency," *J Control Release.,* vol. 94, no. 1, pp. 245-252, 2004.

[24] W. Yin, P. Xiang and Q. Li, "Investigations of the effect of DNA size in transient transfection assay using dual luciferase system," *Anal Biochem.,* vol. 346, no. 2, pp. 289-294, 2005.

[25] F. Faurez, D. Dory, V. Le Moigne, R. Gravier and A. Jestin, "Biosafety of DNA vaccines: New generation of DNA vectors and current knowledge on the fate of plasmids after injection," *Vaccine.,* vol. 28, no. 23, pp. 3888-3895, 2010.

[26] K. Remaut, N. Sanders, F. Fayazpour, J. Demeester and S. De Smedt, "Influence of plasmid DNA topology on the trasnfection properties of DOTAP/DOPE lipoplexes," *J Control Release.,* vol. 115, no. 3, pp. 335-343, 2006.

[27] FDA, "Guidance for Industry: Considerations for Plasmid DNA Vaccines for Infectious Disease Indications," U.S Food and Drug Administration, 2007.

[28] H.-S. Wang, Z.-J. Chen, G. Zhang, X.-L. Ou, X.-L. Yang, C. Wong, J. Giesy, J. Du and S.-Y. Chen, "A Novel Micro-Linear Vector for in vitro and in vivo Gene Delivery and Its Application for EBV Positive Tumors," *PLoS ONE,* vol. 7, no. 10, p. e47159, 2012.

[29] N. Nafissi and R. Slavcev, "Construction and Characterization of an in-vivo Linear Covalently Closed DNA Vector Production System," *Microb Cell Fact.,* vol. 11, p. 154, 2012.

[30] N. Nafissi, S. Alqawiaq, E. Lee, M. Foldvari, P. Spagnuolo and R. Slavcev, "DNA Ministrings: Highly Safe and Effective Gene Delivery Vectors," *Mol Ther Nucleic Acids.,* vol. 3, no. 6, p. e165, 2014.

[31] System Biosciences, "The Development of Minicircles, Interview with Inventor Mark Kay, M.D, Ph.D," *SBInsights,* pp. 6-7, April 2012.

[32] M. Simcikova, K. Prather, D. Prazeres and G. Monteiro, "On the dual effect of glucose during production of pBAD/AraC-based minicircles," *Vaccine,* vol. 32, no. 24, pp. 2843-2846, 2014.

[33] B. Hallet and D. Sherratt, "Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangments," *FEMS Microbiol Rev.,* vol. 21, no. 2, pp. 157-178, 1997.

[34] M. Schleef, M. Blaesen, M. Schmeer, R. Baier, C. Marie, G. Dickson and D. Scherman, "Production on Non Viral DNA Vectors," *Curr Gene Ther.,* vol. 10, no. 6, pp. 487-507, 2010.

[35] B. Bigger, O. Tolmachov, J. Collombet, M. Fragkos, I. Palaszewski and C. Coutelle, "An araC-controlled Bacterial cre Expression System to Produce DNA Minicircle Vectors for Nuclear and Mitochondrial Gene Therapy," *J Biol Chem.,* vol. 276, no. 25, pp. 23018-23027, 2001.

[36] K. Nehlsen, S. Broll and J. Bode, "Replicating minicircles: Generation of nonviral episomes for the efficient modification of dividing cells," *Gene Ther Mol Biol.,* vol. 10, pp. 233-244, 2006.

[37] Z. Chen, C. He, A. Ehrhardt and M. Kay, "Minicircle DNA Vectors Devoid of Bacterial DNA Result in Persistent and High-Level Transgene Expression in Vivo," *Mol Ther.,* vol. 8, no. 3, pp. 495-500, 2003.

[38] W. Jechlinger, C. Tabrizi, W. Lubitz and P. Mayrhofer, "Minicircle DNA Immobilized in Bacterial Ghosts: In vivo Production of safe Non-Viral DNA Delivery Vehicles," *J Mol Microbiol Biotechnol,* vol. 8, no. 4, pp. 222-231, 2004.

[39] P. Mayrhofer, M. Blaesen, M. Schleef and W. Jechlinger, "Minicircle-DNA production by site-specific recombination and protein-DNA interaction chromatography," *J Gene Med.,* vol. 10, no. 11, pp. 1253-1269, 2008.

[40] P. Mayrhofer, M. Schleef and W. Jechlinger, "Use of Minicircle Plasmids for Gene Therapy," in *Gene Therapy of Cancer*, vol. 542, Humana Press, 2009, pp. 87-104.

[41] P. Kreiss, B. Cameron, A. Darquet, D. Scherman and J. Crouzet, "Production of a new DNA vehicle for gene transfer using site-specific recombination," *Appl Microbiol Biotechnol.,* vol. 49, no. 5, pp. 560-567, 1998.

[42] L. Guzman, D. Belin, M. Carson and J. Beckwith, "Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose PBAD Promoter," *J Bacteriol.,* vol. 177, no. 14, pp. 4121-4130, 1995.

[43] J. Carra and R. Schleif, "Variation of half-site organization and DNA looping by AraC protein," *EMBO J.,* vol. 12, no. 1, pp. 35-44, 1993.

[44] C. Chang, L. Christensen, M. Lee and S. Kim, "Efficient Expression of Vascular Endothelial Growth Factor Using Minicircle DNA for angiogenic Gene Therapy," *J Control Release.,* vol. 125, no. 2, pp. 155-163, 2008.

[45] K. Narsinh, F. Jia, R. Robbins, M. Kay, M. Longaker and J. Wu, "Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors," *Nat Protoc.,* vol. 6, no. 1, pp. 78-88, 2011.

[46] P. Mayrhofer and M. Iro, "Minicircle-DNA," in *Gene Vaccines*, SpringerWienNewYork, 2012, pp. 297-310.

[47] J. Urthaler, H. Schuchnigg, P. Garidel and H. Huber, "Industrial Manufacturing of Plasmid-DNA Products for Gene Vaccination and Therapy," in *Gene Vaccines*, SpringerWienNewYork, 2012, pp. 311-330.

[48] D. Prazeres, "Part III Manufacturing," in *Plasmid Biopharmaceuticals: Basics, Applications, and Manufacturing*, John Wiley & Sons, Inc., 2011, pp. 327-562.

[49] Z. Chen, C. He and M. Kay, "Improved Production and Purification of Minicircle DNA Vector Free of Plasmid Bacterial Sequences and Capable of Persistent Transgene Expression In Vivo," *Hum Gene Ther.,* vol. 16, no. 1, pp. 126-131, 2005.

[50] M. Kay, C. He and Z. Chen, "A robust system for production of minicircle DNA vectors," *Nat Biotechnol.,* vol. 28, no. 12, pp. 1287-1289, 2010.

[51] System Biosciences, "SBI," [Online]. Available: http://www.systembio.com/minicircle-dna-vectors/mcdna-production. [Accessed August 2014].

[52] M. Simcikova, N. Kubasova, G. Monteiro and D. Prazeres, "Processo Para a Produção e Purificação de Minicírculos". PT Patent 107663, 2014.

[53] J. Stadler, R. Lemmens and T. Nyhammar, "Plasmid DNA purification," *J Gene Med.,* vol. 6, no. S1, pp. S54-S66, 2004.

[54] M. Diogo, J. Queiroz and D. Prazeres, "Chromatography of plasmid DNA," *J Chromatogr A.,* vol. 1069, no. 1, pp. 3-22, 2005.

[55] F. Sousa, D. Prazeres and J. Queiroz, "Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography," *J Gene Med.,* vol. 11, no. 1, pp. 79-88, 2009.

[56] F. Sousa, C. Cruz and J. Queiroz, "Amino acids-nucleotides biomolecular recognition: from biological occurence to affinity chromatography," *J Mol Recognit.,* vol. 23, no. 6, pp. 505-518, 2010.

[57] F. Sousa, S. Freitas, A. Azzone, D. Prazeres and J. Queiroz, "Selective purification of supercoiled plasmid DNA from clarified cell lysates with a single histidine-agarose chromatography step," *Biotechnol Appl Biochem.,* vol. 45, no. Pt 3, pp. 131-140, 2006.

[58] F. Sousa, T. Matos, D. Prazeres and J. Queiroz, "Specific recognition of supercoiled DNA in arginine affinity chromatography," *Anal Biochem.,* vol. 374, no. 2, pp. 432-434, 2008.

[59] A. Sousa, F. Sousa and J. Queiroz, "Biorecognition of supercoiled plasmid DNA isoform in lysine-affinity chromatography," *J Chromatogr B.,* vol. 877, no. 27, pp. 3257-3260, 2009.

[60] F. Sousa, C. Tomaz, D. Prazeres and J. Queiroz, "Separation of supercoiled and open circular plasmid DNA isoforms by chromatography with histidine-agarose support," *Anal Biochem.,* vol. 343, no. 1, pp. 183-185, 2005.

[61] A. Sousa, F. Sousa and J. Queiroz, "Impact of lysine-affinity chromatography on supercoiled plasmid DNA purification," *J Chromatogr B.,* vol. 879, no. 30, pp. 3507-3515, 2011.

[62] F. Sousa, D. Prazeres and J. Queiroz, "Dynamic binding capacity of plasmid DNA in histidine-agarose chromatography," *Biomed Chromatogr.,* vol. 21, no. 9, pp. 993-998, 2007.

[63] M. Diogo, J. Queiroz, G. Monteiro, S. Martins, G. Ferreira and D. Prazeres, "Purification of a cystic fibrosis plasmid vector for gene therapy using hydrophobic interaction chromatography," *Biotechnol Bioeng.,* vol. 68, no. 5, pp. 576-583, 2000.

[64] H. Bo, J. Wang, Q. Chen, H. Shen, F. Wu, H. Shao and S. Huang, "Using a single hydrophobic-interaction chromatography to purify pharmaceutical-grade supercoiled plasmid DNA from other isoforms," *Pharm Biol.,* vol. 51, no. 1, pp. 42-48, 2013.

[65] S. Iuliano, J. Fisher, M. Chen and W. Kelly, "Rapid analysis of a plasmid by hydrophobic-interaction chromatography with a non-porous resin," *J Chromatogr A.,* vol. 972, no. 1, pp. 77-86, 2002.

[66] N. Schwierz, D. Horinek and R. Netz, "Anionic and Cationic Hofmeister Effects on Hydrophobic and Hydrophilic Surfaces," *Langmuir,* vol. 29, no. 8, pp. 2602-2614, 2013.

[67] S. Freitas, J. Santos and D. Prazeres, "Plasmid purification by hydrophobic interaction chromatography using sodium citrate in the mobile phase," *Sep Purif Technol.,* vol. 65, no. 1, pp. 95-104, 2009.

[68] D. Latulippe and A. Zydney, "Size exclusion chromatography of plasmid DNA isoforms," *J Chromatogr A.,* vol. 1216, no. 35, pp. 6295-6302, 2009.

[69] G. Ferreira, G. Monteiro, D. Prazeres and J. Cabral, "Downstream processing of plasmid DNA for gene theraphy and DNA vaccine applications," *Trends Biotechnol.,* vol. 18, no. 9, pp. 380-388, 2000.

[70] M. Šimčíková, "Development of a process for the production and purification of minicircles for biopharmaceutical applications," PhD Thesis in Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, 2013.

[71] L. Brito, "Minicircle production and delivery to human mesenchymal stem/stromal cells for angiogenesis stimulation," Master Thesis in Biotechnology, Instituto Superior Técnico, Universidade de Lisboa, 2014.

[72] H. Birnboim and J. Doly, "A rapid alkaline extraction procedure for screening recombinant plasmid DNA," *Nucleic Acids Res.,* vol. 7, no. 6, pp. 1513-1523, 1979.

[73] A. Khlebnikov, K. Datsenko, T. Skaug, B. Wanner and K. J.D., "Homogeneous expression of the P(BAD) promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter," *Microbiology,* vol. 147, no. Pt 12, pp. 3241-3247, 2001.

[74] D. Heiter, K. Lunnen and G. Wilson, "Site-Specific DNA-nicking Mutants of the Heterodimeric Restriction Endonuclease R.BbvCI," *J Mol Biol.,* vol. 348, no. 3, pp. 631-640, 2005.

[75] S. Chan, B. Stoddard and S. Xu, "Natural and engineered nicking endonucleases--from cleavage mechanism to engineering of strand-specificity," *Nucleic Acids Res.,* vol. 39, no. 1, pp. 1-18, 2011.

[76] V. Gaspar, C. Maia, J. Queiroz, C. Pichon, I. Correia and F. Sousa, "Improved Minicircle DNA Biosynthesis for gene therapy applications," *Hum Gene Ther Methods.,* vol. 25, no. 2, pp. 93-105, 2014.

Attachment

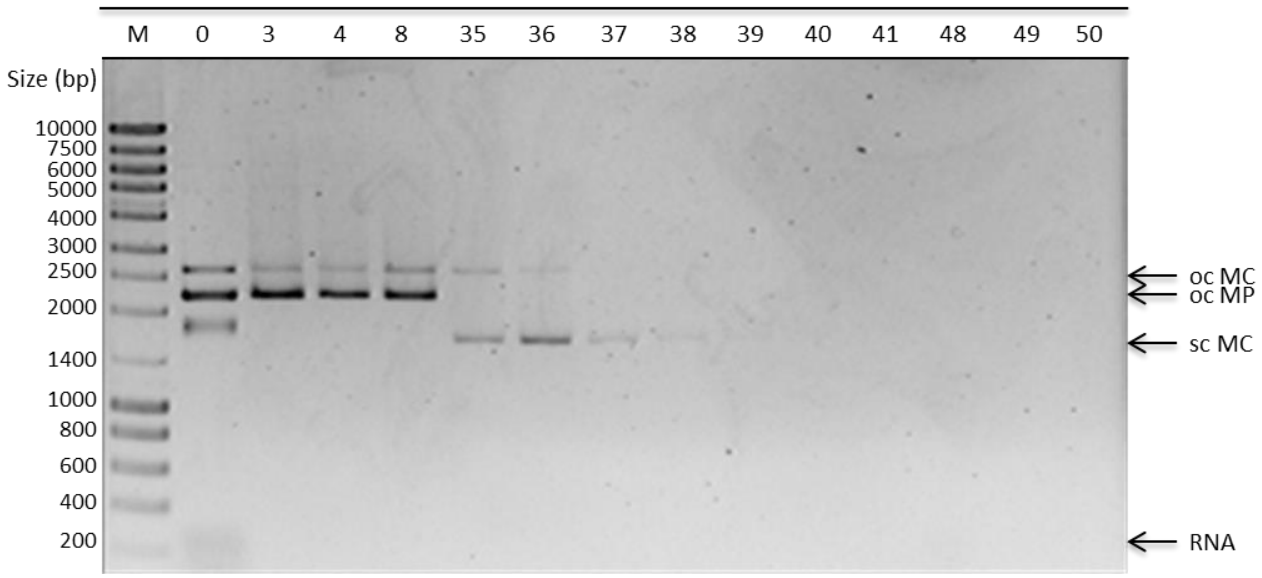

Figure i: Agarose gel analysis of 20 µL from fractions collected during the chromatographic run with injection of 0.4 mg of total nucleic acids, presented in Figure 4.29 (purple line). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lanes 3 and 4 – first peak; lane 8 –second peak; lanes 35 to 39 – third peak; lanes 49 and 50 – fourth peak). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle.

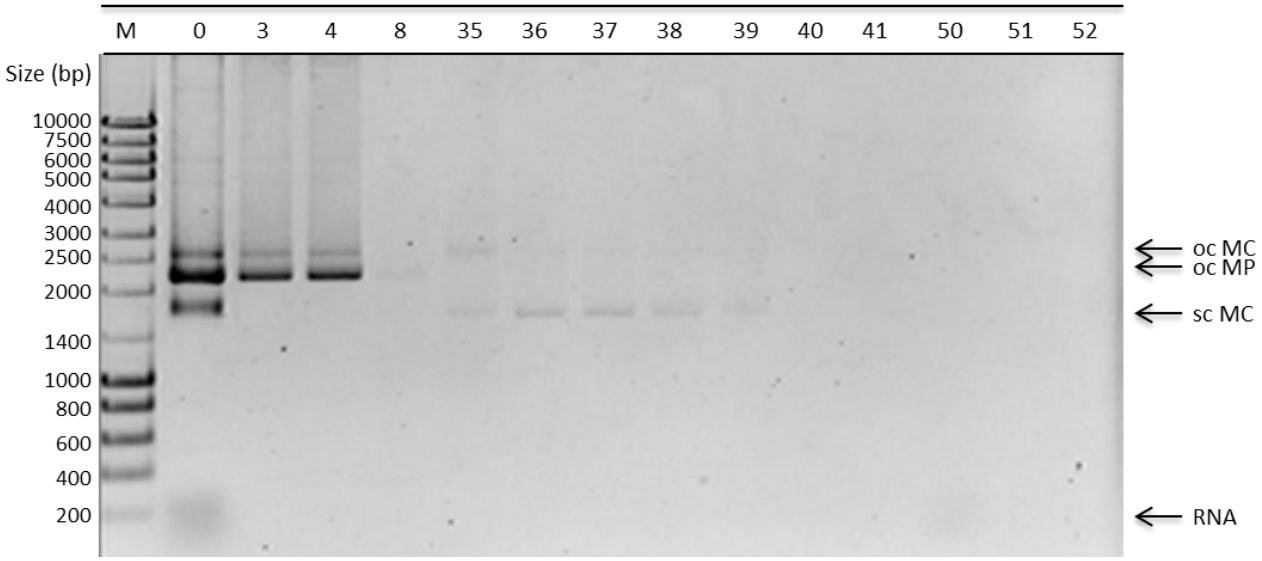

Figure ii: Agarose gel analysis of 20 µL from fractions collected during the chromatographic run with injection of 0.97 mg of total nucleic acids, presented in Figure 4.29 (blue line). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lanes 3 and 4 – first peak; lane 8 –second peak; lanes 35 to 39 – third peak; lanes 50 to 52 – fourth peak). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle.

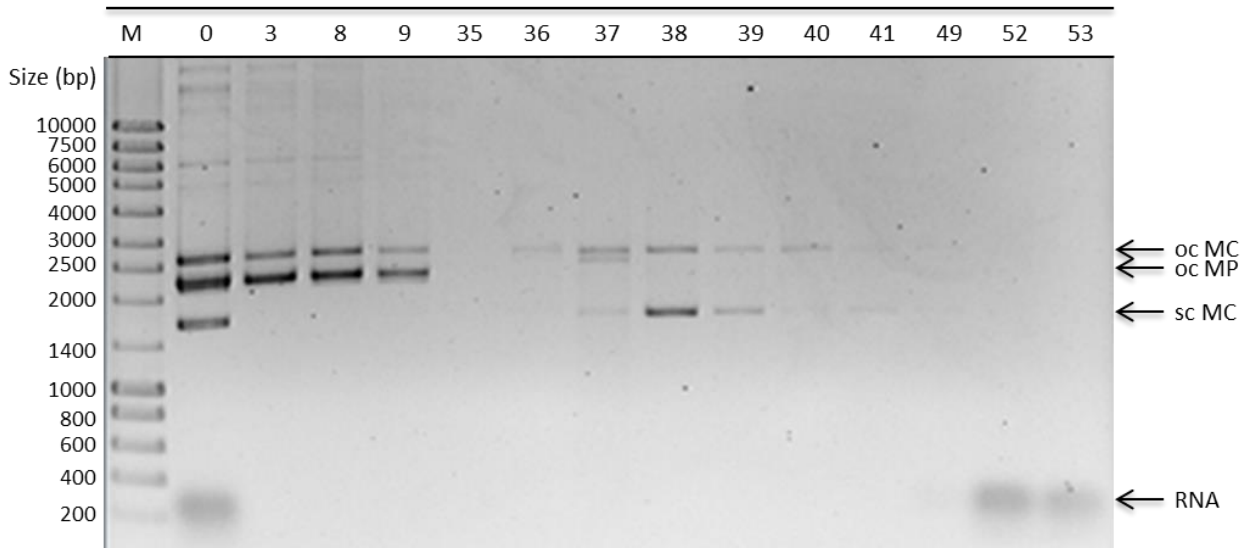

Figure iii: Agarose gel of analysis 20 µL from fractions collected during the chromatographic run with injection of 0.56 mg of total nucleic acids, presented in Figure 4.30 (orange line). Lane M – molecular weight marker; lane 0 – column feed; numbered lanes correspond to fractions collected during chromatographic run (lane 3 – first peak; lanes 8 and 9 –second peak; lanes 35 to 39 – third peak; lanes 49 to 53 – fourth peak). oc MC – open circular minicircle; oc MP – open circular miniplasmid; sc MC – supercoiled minicircle.