UNIVERSIDADE DA BEIRA INTERIOR
Ciências da Saúde

Purification of HPV16 E6/E7 DNA plasmid-based vaccine using a modified monolithic support

Alexandra Sabrina Antunes Soares

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

Orientadora: Profª. Doutora Ângela Sousa
Coorientadora: Profª. Doutora Fani Sousa

Covilhã, junho de 2013

"You've got to find what you love. And that is as true for your work as it is for your lovers. Your work is going to fill a large part of your life, and the only way to be truly satisfied is to do what you believe is great work. And the only way to do great work is to love what you do. If you haven't found it yet, keep looking. Don't settle. As with all matters of the heart, you'll know when you find it. And, like any great relationship, it just gets better and better as the years roll on. So keep looking until you find it. Don't settle."

Steve Jobs

To the most important people of my life,

whom I love very much,

My parents…

Acknowledgments

First, I specially thank my supervisors Professor Ângela Sousa e Professor Fani Sousa for believing in me and my ability to develop this work, and for all the patience, help and knowledge they transmitted to me along the past year.

To Professor Doctor João Queiroz, rector from University of Beira Interior, I would like to express my gratitude for making possible the development of this investigation project.

I would also like to express my gratitude to the people at the Biotechnology and Biomolecular Sciences group of the Health Sciences Research Centre at University of Beira Interior for all their help, with a special acknowledge to Patricia Pereira and Marta Silva for their friendship, complete availability and support.

I would like to thank my friends for all the coffee, long talks, patience and full support, especially to Cátia, Elisabete, Pedro, Rui, Stéphanie and Vanessa.

To my family, especially my godparents, Maria de Lurdes e José Orlando, my grandmother Carminda, my "sister" Andrea and my brother Óskar for all their love and support in the most difficult times.

Finally, I will be eternally grateful to my parents for all the love, sacrifices, good advice and support they gave me all my life.

Abstract

Human papillomavirus (HPV) is one of the most common sexually transmitted diseases in the world and has been associated with several human cancers, like cervical cancer. Thus, effective vaccination against Human papillomavirus represents an opportunity for the control of this cancer. The development of therapeutic Human papillomavirus vaccines is required to facilitate the control and eliminate on of a preexisting Human papillomavirus infection. In the last years, the expansion of efficient plasmid DNA purification processes has fostered therapeutics applications like gene therapy and DNA vaccination. Recently, the application of chromatographic operations based on affinity interactions between plasmid DNA or impurities with specific amino acids immobilized in stationary phases has demonstrated good results in the supercoiled plasmid DNA purification. Despite of selectivity achieved with these ligands, conventional matrices present limitations such as the low binding capacity and diffusivity for plasmid DNA samples. Owing to bottlenecks associated to conventional matrices, monolithic supports have emerged as interesting alternatives due to the versatility of their structural characteristics. The research work present in this thesis describes a new strategy that combines the selectivity of arginine as affinity ligand with the versatility of the epoxy-based monoliths to efficiently purify the supercoiled HPV-16 E6/E7 plasmid from other plasmid isoforms and *Escherichia coli* impurities present in clarified lysate. Additionally, breakthrough experiments were designed to compare the dynamic binding capacity of plasmid DNA to the conventional arginine-agarose matrix with the modified monolithic support. The dynamic binding capacity obtained for the arginine-epoxy monolith was significantly higher than the capacity achieved in the arginine conventional support. Quality control tests indicated that the plasmid sample resultant from the purification step presented a purity degree approximately 100% and an homogeneity higher than 97% of supercoiled isoform, with an extremely reduced level of impurities (RNA, proteins, genomic DNA and endotoxins). Overall, given that the plasmid DNA final product meets regulatory specifications, this combined support can be the key to obtain an adequate non-viral vaccine against a Human papillomavirus infection.

Keywords

Affinity chromatography, arginine ligand, Human papillomavirus, DNA vaccines, modified monolithic support, supercoiled plasmid DNA

Resumo

O Papiloma Vírus Humano é um vírus sexualmente transmissível que está relacionado com o desenvolvimento de vários cancros, como o cancro do colo do útero. Este tipo de cancro é a segunda maior causa de morte em mulheres, afetando mundialmente cerca de meio milhão de mulheres, das quais aproximadamente 274 mil morrem. A evidente associação que existe entre o Papiloma Vírus Humano e o cancro do colo do útero torna o Papiloma Vírus Humano um alvo interessante para o desenvolvimento de vacinas, no sentido de prevenir ou tratar o desenvolvimento do cancro. Atualmente, já são comercializadas duas vacinas preventivas, a Gardasil da Merck e a Cervarix da GlaxoSmithKline. Apesar de ambas mostrarem ser eficazes e seguras, possuem algumas limitações como elevado custo, não protegem contra todos os tipos de Papiloma Vírus Humano e trata-se de vacinas exclusivamente preventivas. Assim o desenvolvimento de vacinas terapêuticas pode ser uma estratégia promissora para colmatar estas falhas.

A utilização do DNA plasmídico (pDNA) como uma vacina não viral tem-se tornado numa potencial estratégia terapêutica para prevenir ou tratar determinadas doenças de forma menos invasiva e segura comparando com os vetores virais. O mecanismo de actuação desta vacina baseia-se na expressão de proteínas antigénicas que desencadeiam uma resposta imunológica, evitando a progressão da doença . A preparação destas vacinas de DNA plasmídico requer o desenvolvimento de processos de produção e purificação que permitam obter grandes quantidades de plasmídeo na sua forma biologicamente ativa (isoforma superenrolada (sc)), cumprindo os requisitos das agências reguladoras no que diz respeito ao grau de pureza. Da ocorrência natural de complexos proteínas-DNA em sistemas biológicos sugeriu o desenvolvimento de uma estratégia de cromatografia de afinidade, utilizando matrizes convencionais de agarose com determinados aminoácidos imobilizados que reconhecem especificamente a isoforma superenrolada do DNA plasmídico. No entanto, as matrizes convencionais apresentam algumas limitações quando comparadas com os suportes monolíticos que possuem excelentes propriedades de transferência de massa e elevada capacidade de ligação para moléculas de grandes dimensões como o DNA plasmídico.

Desta forma, o trabalho apresentado nesta tese consistiu na modificação de um monolito de epoxy por imobilização de aminoácidos de arginina, conjugando assim a selectividade deste ligando com a versatilidade do monolito de epoxy, tendo como finalidade purificar a isoforma sc do pDNA HPV-16 E6/E7. Numa fase inicial, foram realizados vários ensaios com amostras de DNA plasmídico pre-purificadas com o kit comercial, no sentido de avaliar e confirmar a presença dos ligandos de arginina, comparando o comportamento cromatográfico do monolito modificado com o mesmo suporte não modificado. Os resultados comprovaram que o monolito modificado reconhece especialmente o DNA plasmídico, permitindo a separação das isoformas superenrolada e circular aberta através de um gradiente por passos de NaCl. O mesmo não aconteceu com o monolito de epoxy não modificado, pois não ocorreu qualquer tipo de

interacção do plasmídeo com o suporte. Posteriormente foi realizado um estudo de capacidade de ligação dinâmica para completar a caracterização do monolito com ligandos de arginina e comparar com a coluna convencional de arginina-argarose. Os resultados comprovaram que, para as mesmas condições de caudal e concentração de pDNA HPV-16 E6/E7, o monolito possui uma capacidade de ligação significativamente superior em comparação com a coluna convencional. Uma vez conseguida a separação das isoformas e a caracterização do monolito modificado com a arginina, a segunda fase do trabalho consistiu na purificação da isoforma sc do pDNA HPV-16 E6/E7 a partir de um lisado complexo de *Escherichia coli*. Os testes de controlo de qualidade revelaram que a amostra de sc de pDNA HPV-16 E6/E7, resultante da purificação com o monolito de epoxy modificado, apresentava um grau de pureza de aproximadamente 100% e uma homogeneidade superior a 97%. Para além disso, constituintes do hospedeiro como RNA e proteínas não foram detectados na amostra purificada e a quantidade de DNA genómico e endotoxinas estavam a baixo dos valores referenciados pelas agências reguladoras como a Food and Drug Delivery.

Em suma, a combinação de ligandos de aminoácidos com suportes monolíticos pode ser uma solução promissora para obtenção de uma vacina não-viral baseada na isoforma sc do pDNA HPV-16 E6/E7, com o grau de pureza requerido para futuras aplicações terapêuticas contra a infecção por Papiloma Vírus Humano.

Palavras-chave

Cromatografia de afinidade, ligando de arginina, monolito de epoxy modificado, Papiloma Vírus Humano, Vacinas de DNA

Table of Contents

List of Figures

List of tables

List of Acronyms

AC	Affinity chromatography
AEC	Anion-exchange chromatography
AP 1	Activator protein 1
APC	Antigen-presenting cells
ATPase	Adenosine triphosphate
BCA	Bicinchoninic acid
BSA	Bovine Serum Albumin
°C	Celsius
CDI	CarbonylDiImidazole
cDNA	Complementary DNA
Cys	Cysteines
CR	Conserved Regions
CTL	Cytolytic T lymphocyte
DC	Dendritic cell
DNA	Desoxirribonucleic acid
E6AP	E6-associated protein
EAEMP	European Agency for the Evaluation of Medical Products
E. coli	*Escherichia coli*
EDTA	Ethylene-diamine tetraacetic acid
EGF	Epidermal growth factor
FDA	Food and Drug Administration
g	gram
gDNA	Genomic DNA
h	Hours
H_2SO_4	Sulfuric acid
H-bond	Hydrogen bond
HIC	Hydrophobic interaction chromatography
HCl	Chloride acid
HPV	Human Pappilomavirus
HSIL	High-grade squamous intraepithelial lesion
kDa	kilo dalton
kbp	kilo base pairs
KH_2PO_4	Monopotassium phosphate
K_2HPO_4	Dipotassium phosphate
KRF 1	Keratinocyte-specific transcriptional factor 1
L	liter
LAL	Limulus amebocyte lysate
LCR	Long control region

ln	Linear
MHC	Major histocompatibility complex
min	minutes
mL	milliliter
mm	millimeter
mRNA	Messenger RNA
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NF-I/CTF	Nuclear factor
$(NH_4)_2SO_4$	Ammonium sulfate
nm	Nanometer
oc	Open circular
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
pb	Base pairs
PCR	Polymerase chain reaction
PDGF	Platelet-derived growth factor
pDNA	Plasmid DNA
Rb	Retinoblastoma
RNA	Ribonucleic acid
rRNA	ribosomal RNA
rpm	Revolutions per minute
S. cerevisiae	*Saccharomyces cerevisiae*
sc	Supercoiled
SDS	Sodium dodecylsulfate
SEC	Size-exclusion chromatography
siRNA	Small interfering RNA
SV40	Simian virus 40
TAE	Tris, acetic acid, EDTA
Tris	Tris(hydroxymethyl) aminomethane
UV	Ultraviolet
V	volt
VLP	Virus-like particles
W/W	Mass/mass

Chapter I

Introduction

Introduction

Human papillomavirus (HPV) is one of the most common sexually transmitted diseases in the world and has been identified as an etiological factor for several important cancers, including anogenital cancers, and a subset of head and neck cancers [1]. Among HPV-associated cancers, cervical cancers have the most significant morbidity, being the second largest cause of the cancer deaths in women worldwide [2]. Infection with HPV is a key cause of cervical cancer because the HPV DNA has been detected in 99.7% of patients with cervical cancer [1]. More than 200 HPV genotypes have been identified and are classified into low or high-risk types, depending on their propensity to cause cervical cancer, being the HPV-16 and HPV-18 the most high risk types associated with the cervical cancer [2]. The clearly association between infection with high-risk HPV and cervical cancer showed that HPV serves as an ideal target for development of preventive and therapeutic vaccines. Recently, two licensed prophylactic HPV vaccines, Gardasil (bivalent vaccine comprising HPV types 16 and 18) and Cervarix (quadrivalent vaccine comprising HPV types 6, 11, 16 and 18) have been shown to be safe and capable to generating significant protection against specific HPV types [3]. However, prophylactic HPV vaccination does not have therapeutic effect against established HPV infections and HPV-associated lesions [3]. In this way, the development of therapeutic HPV vaccines can be a promising strategy to accelerate the control of cervical cancer and treat currently infected patients [4].

DNA vaccines have emerged as attractive and potentially effective strategies for treatment of inherent and acquired diseases. The most advantage of this novel vaccine technology is the ability to inducing both humoral and cellular immunity [5]. The underlying principle of DNA vaccines consists in the administration of therapeutic genes for specific proteins that induce the immune responses against pathogen [6]. Central to the concept of DNA vaccination is the ability to deliver exogenous nucleic acids to eukaryotic cells of various tissues. However, the success of DNA vaccination is largely dependent on the development of a vector or vehicle that can selectively and efficiently deliver therapeutic genes into these cells. On the other hand therapeutic genes can be transported by viral vectors but this delivery system has presented higher toxicity and immunogenicity degree [7, 8]. Thus, plasmid DNA (pDNA)-based vaccines are more safe, stable, easy production and prolonged action if compared to virus-based vaccines, becoming the most attractive gene-transfer system to be used as biopharmaceutical product [2].

The successful implementation of clinical approaches using plasmid DNA-based strategies requires the continuous improvement of production and purification processes of pDNA [9]. In this way, different chromatographic techniques, including size-exclusion, anion exchange, hydrophobic interaction, reversed phase and affinity chromatography have been applied for supercoiled (sc) pDNA isolation from the remaining host impurities, with more or less success [10]. Recently, a new affinity chromatography approach using some amino acids as specific binding agents have been efficiently applied in the purification of the sc pDNA from the

clarified *Escherichia coli* (*E. coli*) lysate [11-13]. Despite of the selectivity achieved with these amino acids, these conventional matrices present some limitations, such as the low capacity of available supports for pDNA and low diffusivity of pDNA samples [14]. Therefore the monolithic stationary phase has been extensively investigated for pDNA purification. This new support has attracted increasing attention because of its easy preparation, structural and functional properties and high performance compared to conventional columns [14].

Hence, the purpose of this thesis is to develop and implement a new strategy for sc (supercoiled) pDNA HPV-16 E6/E7 isoform purification. The affinity purification strategy is based on the combination of the specificity and selectivity achieved with arginine ligand with the advantages provided by a monolithic support. In addition, the characterization of the sc pDNA sample obtained is also envisioned since it should be free from the host contents in compliance with all regulatory requirements and capable of inducing gene expression with high efficiency.

Thus, the work described in this thesis may be divided in three main stages:
- First comprises the biosynthesis of pDNA HPV-16 E6/E7 in *E. coli* recombinant organism;
- The second stage consists on the implementation of an monolithic affinity purification strategy with immobilized arginine amino acids to separate the sc and oc (open circular) plasmid isoforms;
- The third stage consists on the purification of the sc plasmid isoform from the clarified *E. coli* lysate;

The final research topic consist in analytical characterization of the sc pDNA recovered the arginine-epoxy monolith.

1. Human papillomavirus

Human Papillomavirus infection represents the most common sexually transmitted disease in the world [15]. HPV has been identified as an etiological factor in several important cancers, including a subset of anal, vaginal, vulval, penile, head and neck cancers [1]. Apart from these cancers, cervical cancer is the second largest cause of cancer deaths in women worldwide and HPV infection represents the most important risk factor in its development [16]. HPV DNA has been detected in 99.7% of patients with cervical cancer [15]. It is estimated that half a million of new cases are diagnosed by year and approximately 270.000 of those patients die [4, 17].

Human Papillomavirus are members of the *Papovaviridae* family [18]. More than 200 HPV types have been identified and are classified based on DNA sequence analysis, it being each one associated with infection in specific epithelial sites [2, 19]. Thus, the HPV types can be divided in two main HPV genera, Alpha and Beta papillomavirus [20]. Beta papillomavirus are typically associated with cutaneous infection, whereas Alpha papillomavirus mainly infect the mucosal/genital areas [21]. Alpha HPV are further subdivided into low, intermediate or high-risk depending on their propensity to cause cervical cancer [18], as it is presented in Table 1. However, Alpha HPV also include some HPV types mostly found in cutaneous lesions, such as HPV-2, which cause common warts [22]. More than 30 HPV types are known to infect cervical epithelium, associated with high-grade squamous intraepithelial lesions (HSILs), which are precursors of cervical cancer [4, 20]. These HPV types are classified as high-risk HPV, being 70% of the cervical cancers attributable to HPV-16 and HPV18 [23]. HPV-16 is considered the most significant high-risk HPV type, being responsible for approximately 50% of cervical cancer [24]. On the other hand, the low-risk HPV type is associated with benign genital warts and it is rarely associated with cervical cancer [25]. Since HPV is present in the majority of cervical cancers, the HPV appearing and infection can be controlled through an adequate vaccination.

Table 1. Human papillomavirus grouping, according to their risk to cause cervical cancer (adapted from [18]).

HPV groups	HPV types
High Risk	HPV-16, HPV-18, HPV-45, HPV-56
Intermediate risk	HPV-31, HPV-33, HPV-35, HPV-51, HPV-52, HPV-58
Low risk	HPV-6, HPV-11, HPV-42, HPV-43, HPV-44

1.1. Molecular biology of HPV

Although there are several differences in the DNA sequence of all HPV types, their structure and genomic organization are similar. Human Papillomavirus are small, non-enveloped with a diameter of 55-60 nm and have an icosahedral capsid composed by 72 capsomers [26, 27]. They contain a circular double-strand, covalently closed DNA of around 8000 base pairs (pb) [7904 bp for HPV16 (GenBank® accession number NC_001526)] in the form of chromatin-like complex with cellular histones [18, 20, 28]. The viral genome is organized in three regions (Figure 1): the long control region (LCR) and two protein-encoding regions (early and late gene regions) [26]. All viral genes are encoded on one strand of HPV genomic DNA.
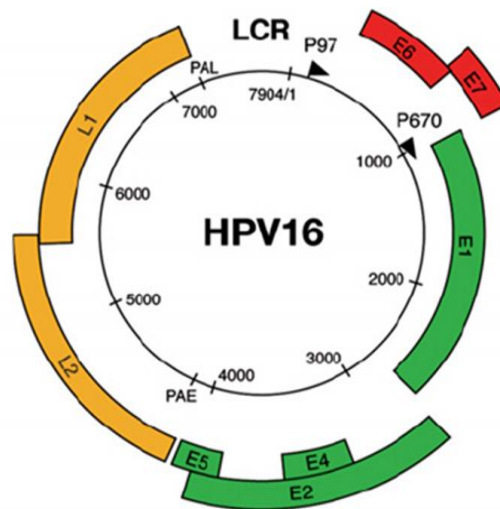
Figure 1. Structure and genome organization of the HPV-16. The HPV-16 genome (7904 bp) is shown as a black circle with the early (p97) and late (p670) promoters marked by arrows. Early ORFs (E1, E2, E4 and E5) are represented in green, while the other early ORFs (E6 and E7) are in red. The late ORFs (L1 and L2) are represented in yellow. All the viral genes are encoded on one strand of the double-stranded circular DNA genome. The long control region is showed between yellow and red regions (adapted from [20]).

Long control region represents approximately 10% of the genome and unlike other two regions does not encode for proteins [26, 29]. Indeed, the LCR contains numerous binding sites for many different transcriptional repressors and transcriptional activators, including activator protein 1 (AP 1), keratinocyte-specific transcriptional factor 1 (KRF 1), nuclear factor (NF-I/CTF) and virally derived transcriptional factors encoded by early and late regions [27, 29]. Thus, LCR is responsible by transcription control from early and late regions and controls the production of viral proteins and infectious particles. LCR also regulates the viral gene expression, suggesting that it may play a critical role in determining the range of hosts for specific HPV types [27].

The late gene region encodes 40% of the HPV genome and contains two separate open reading frames (ORFs), L1 and L2 that are structural components of the viral capsid [2, 26]. The L1

gene encodes the major (55 kDa) capsid protein and forms the main architectural structure of the icosahedral viral capsid [21, 26]. This protein is highly conserved among different papillomavirus species [29]. On the other hand, the L2 gene encodes the minor (70 kDa) capsid protein and it has more sequence variations than L1 protein [26, 27]. Both capsid proteins are only expressed in terminally differentiated squamous epithelial cells [28]. The expression of L1 and L2 proteins is tightly regulated and associated with differentiation of infected epithelial cells [2].

Finally, the early gene region is the major segment of the viral genome encoding 50% of the total genome [26]. This region is downstream of the LCR region and is composed of six ORFs, E1, E2, E4, E5, E6 and E7 [27]. The E genes encode nonstructural proteins that regulate virus transcription and replication and are expressed in nonproductively infected cells and transformed cells. Depending upon the HPV type, the E4, E5 and E7 genes usually encode a single polypeptide, whereas the E1, E2 and E6 genes may be expressed as several related polyproteins through differential splicing [26]. Thereby, early proteins play a critical role in regulation of viral DNA replication (E1 and E2), viral RNA transcription (E2), cytoskeletal reorganization (E4) and cellular transformation (E5, E6 and E7) [2]. In Table 2 are summarized the functions of each viral protein.

The E1 gene encodes polypeptides of 68 and 27 kDa required for extrachromosomal DNA replication and conclusion of the viral life cycle [26, 27]. The 68 kDa protein encodes adenosine triphosphate (ATPase) and helicase activity that binding to a specific sequence in LCR region will initiate DNA replication [26]. The E2 gene also encodes two proteins that together with E1 protein are responsible for regulation of extrachromosomal DNA replication. One of the E1 and E2 proteins inhibits the transcription of the early region whereas the other protein increases the transcription of the early region [29]. By acting as transcriptional activators and repressors due to the binding to a specific sequence in LCR region, E2 proteins regulate virus transcription and genome replication [30].

Such as L1 and L2 capsid protein, the E4 protein is expressed late in virus replication when complete virions are being assembled [31]. The E4 protein is associated with cellular membranes and accumulates in the cell cytoplasm [32]. Despite this protein does not seem to possess transforming proprieties, it has an important role for the virus maturation and replication [32].

The functional role of E5 protein is less known because this protein is not expressed in most HPV-positive cancers, suggesting that the E5 protein is not essential in maintaining the malignant transformation of the host cells [29]. When present, the E5 protein interacts with cell membrane receptors, such as epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) and this may stimulate cell proliferation in HPV-infected cells [27]. Therefore, the E5 protein possesses a weak transforming activity in HPV-infected cells [27].

Table 2. A brief description of the functions of human papillomavirus open-reading frames (adapted from [26])

Genes	Function
L1	Major capsid protein
L2	Minor capsid protein
E1	Viral DNA synthesis
E2	Transcription regulatory protein
E4	Disrupts cytokeratins, late protein
E5	Interacts with growth factor receptors
E6	Transforming protein, binds and initiates p53 degradation
E7	Binds retinoblastoma gene family members (Rb1, p107, p103)

1.1.1. Human papillomavirus E6 oncoprotein

Papillomavirus E6 protein consists in 150 amino acids with approximately 18 kDa and it is localized in the nuclear matrix and non-nuclear membrane fraction [21, 33]. At the structural level, the E6 protein is composed by two zinc finger domains and in the base of each zinc finger are two motifs containing two cysteines (Cys-X-X-Cys, where x is any amino acid), which are conserved throughout all HPV types [34]. In the second zinc finger domain exists an LXXLL motif, which is required to mediate interactions between the E6 protein and other proteins, namely the cellular E3 ubiquitin-ligase known as E6-associated protein (E6AP) [21]. E6 protein can form a complex with E6AP that is required for the association of E6 and p53 proteins, resulting in loss of p53 activity within cells [35]. In this case, the p53 degradation occurs through a ubiquitin-dependent mechanism, in contrast to the large T antigen of simian virus 40 (SV40) and E1B protein of adenovirus 5, which will inactive the p53 protein by sequestration into a complex [36]. This mechanism has been reported for the high-risk HPV-16 and 18, while in the case of low-risk HPV, E6 also binds p53 but with reduced efficiency, it being not capable of inducing p53 degradation [35].

P53 tumor suppressor can sense damage or potential damage in the cellular DNA and invoke a protective response (by blocking the cell cycle or inducing apoptosis in the affected cell), it being extremely important to maintain the genome integrity [37]. When the p53 protein detects damage in the cellular DNA, it activates the WAF1/Cip1 gene, which is an inhibitor of cyclin-dependent kinases. High cellular levels of WAF1/Cip1 inhibit cyclin-dependent kinases, which are essential for progression through the cell cycle. Thus, the cell cycle stops in the G_1 phase in order to repair the damage in the DNA and maintain genome integrity [38]. Meanwhile, inactivation or mutation of p53 function leads to continued replication of cells with damaged DNA promoting mutation, chromosomal instability and carcinogenesis of the host genome (as it is showed in Figure 2) [26].
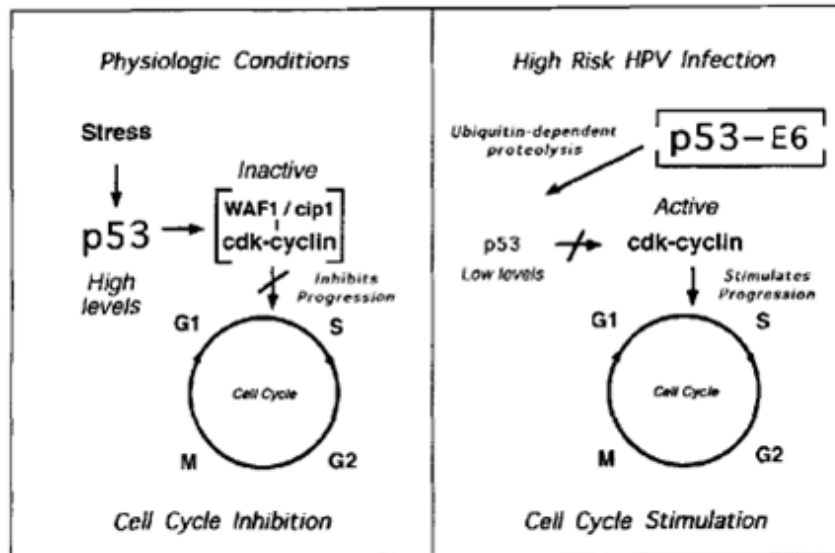
Figure 2. Effect of the interaction between E6 protein and p53 tumor suppressor. (Left) In normal physiologic conditions when p53 detects damage in the cellular DNA, it activates WAF1/Cip1, which inhibits cyclin-dependent kinases and consequently stops the cell cycle in G_1 phase. (Right) When an infection occurs with a high-oncogenic risk HPV type, the presence of E6-p53 complexes promotes the p53 inactivation, leading to uncontrolled cell replication (adapted from [27]).

1.1.2. Human papillomavirus E7 oncoprotein

Papillomavirus E7 protein consists in 98 amino acids that were localized to the nuclear matrix [34]. This protein is composed by three functional domains designated by Conserved Regions (CR1, 2 and 3), knowing that CR1 and CR2 domains are structural and functionally related to the E1A protein of adenovirus and the large T-antigen of simian virus 40 [39]. The CR1 domain plays a crucial role in cellular transformation independent of pRb-binding, whereas the CR2 domain exits Leu-X-Cys-X-Glu (LXCXE) motif act as a pRb-binding site [21, 40]. Thus, E7 proteins form complexes with the retinoblastoma (Rb) protein, which is related with proteins p107 and p130, being negative regulators of cell growth [26].

In physiological conditions, the Rb protein forms complexes with transcription factors, such as E2F, which regulate the expression of cellular genes that function in DNA synthesis. The Rb protein binds to E2F transcription factor, becoming phosphorylated by one or more cyclin-dependent kinases, and together with the hypophosphorylated Rb are responsible to block cell cycle progression in G_0 and G_1 phases [41]. Other Rb-related proteins, p107 and p130, also form complexes with E2F and regulate the cell cycle progression [27]. In the case of progression from G_1 into S phase, E2F-Rb complexes are dissociate and the free E2F stimulates the transcription of the genes required for transition into S phase [41]. In contrast, the E7 protein binds to Rb, p107 and p130 proteins and releases free E2F transcription factors in the cells, which can stimulate the transcription of the E2F-dependent genes necessary for DNA replication and cellular cycle progression [27, 42] (as it is show in Figure 3). The ability

of E7 protein for disrupt E2F-Rb complexes has been reported to high-risk HPV types, while the low-risk HPV-E7 protein have demonstrated a low Rb, binding affinity [27, 42].
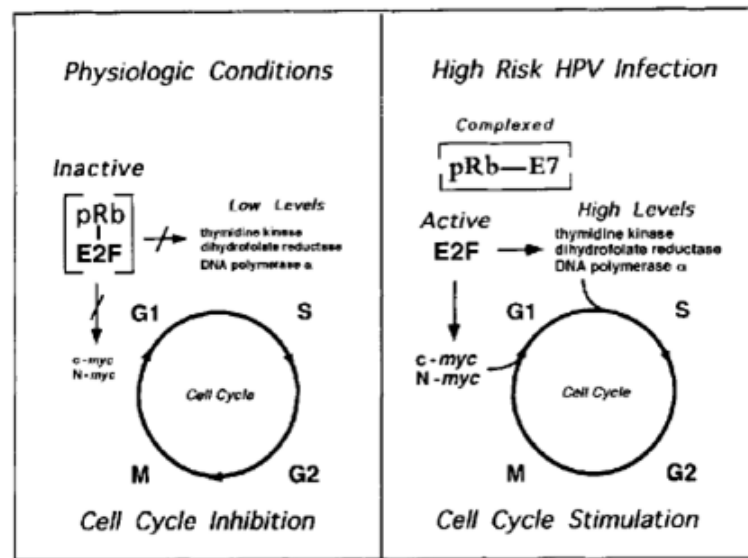
Figure 3. Effect of the interaction between the E7 protein and pRB protein (Left) In normal physiologic conditions, the pRB protein forms complexes with E2F by phosphorylation of the pRB with cyclin-dependent kinases and consequently the Rb stays hypophosphorylated leading to the cell cycle stop in G_0 or G_1 phases. (Right) When the infection with a high-oncogenic risk HPV type occurs, the presence of E7-pRB complexes releases free E2F, which stimulate the transcription of the genes responsible for DNA replication and cell cycle progression (adapted from [27]).

1.2. Preventive HPV vaccines

As the HPV viral DNA has been detected in more of 99% patients with a cervical cancer, the vaccination with fragments of HPV genomic DNA can be a potential solution for this pathology, it being essential a basic understanding of the HPV biology. Two vaccination types can be used against the cervical cancer. Preventive vaccines prevent the infection by generating neutralizing antibodies to block HPV viral infection and therapeutic vaccines eliminate the infection by inducing a virus-specific T cells-mediated response [2].

Recently, two HPV preventive vaccines (quadrivalent and bivalent) were developed and have been licensed for use in many countries. The quadrivalent vaccine of HPV L1 virus-like particles (VLP) was called Gardasil, developed by Merck (NJ, USA) and approved by the USA Food and Drug Administration (FDA). This vaccine is produced in *Saccharomyces cerevisiae* (*S. cerevisiae*) and protects against four of the most medically relevant HPV genotypes: HPV-16 and HPV-18 for cervical cancer, and HPV-6 and HPV-11 for benign genital warts [1, 24]. The bivalent vaccine HPV L1 VLP was called Cervarix and developed by GlaxoSmithKline (GlaxoSmithKline Biologicals, Rixensart, Belgium). This vaccine uses an insect cell expression system and protects against two HPV genotypes, HPV-16 and HPV-18 [1, 24].

Both Gardasil and Cervarix vaccines are based on the use of HPV VLP to generate neutralizing antibodies against the major capsid protein, L1. Despite of L1 is not expressed in the infected basal cells, this protein has been largely studied as target for the preventive vaccine development. These vaccines only induced the humoral response, by promoting the production of specific antibodies. Moreover, the ability of HPV-16 L1 VLP vaccination to generate memory B cells has been reported. In this way, memory B cells provide a rapid response of antibodies upon secondary exposures [43]. Although Gardasil and Cervarix vaccines have been applied with some success as preventive vaccines, they also present some limitations, such as high manufacturing cost, type-restricted protection and no present therapeutic effects which are displayed in Table 3.

Table 3. Disadvantages of current licensed human papillomavirus vaccines – Cervarix® and Gardasil® (adapted from [44]).

Limitation	Impact
Cost of manufacturing, trials and intellectual property leads to high vaccine price	Developing countries that do not have the economic means to afford the vaccine continue to bear the burden of cervical cancer
Type restriction	Vaccine does not protect against all HPV types causing cervical cancer, thus screening must continue for HPV-vaccinated women
No therapeutic value	No effect on individuals with pre-existing HPV infection

1.3. Therapeutic HPV vaccines

Because of limitations from current HPV vaccines previously mentioned, the development of therapeutic vaccines to facilitate the control of cervical cancer and others HPV-associated malignances becomes increasingly necessary. The choice of target antigen is the key of the therapeutic vaccines design. HPV early proteins are potential target antigens because they are involved in the regulation of virus transcription and replication and can be expressed in both nonproductively infected cells and in transformed cells. Particularly, HPV E6 and E7 viral genes are mainly responsible for malignant transformation [4].

2. Gene Therapy and DNA Vaccines

Millions of people die each year from hereditary, degenerative and infectious diseases, or cancer [45]. These diseases are considered untreatable by conventional clinical methods. In last decades, the determination of relationships between genes and the maintenance and regulation of the organism, and the decoding of the Human Genome in 2000, led to the conclusion of the existence of wide range of diseases depending on the gene [46, 47]. Thus, it emerged the opportunity for the development of therapeutic strategies based on the gene intervention such as gene therapy and DNA vaccination [48].

The Human Gene Therapy is defined as the introduction of genetic material that encodes a desired gene into human target cell with the aim of correcting phenotypic or genotypic abnormalities or providing new functions to the cells [49, 50]. This therapeutic strategy may be divided into somatic and germ gene therapies [49]. However, only somatic cells are targeted for treatment because genetic modifications in somatic cells affect the individual patient, while alterations in germ cells have potential to affect the future generations [49].

DNA vaccination is other innovative human DNA therapeutic strategy based on the gene intervention for the treatment of inherited and acquired diseases [51]. DNA vaccination is based on the administration of therapeutic genes in human target cells nucleus [52]. When administrated, therapeutic gene is transcribed and translated in the cytoplasm of the cell and the resultant protein will induce the immune response [52, 53].

2.1. DNA vaccination

As previously mentioned, millions of people die at each year due to incurable diseases that often provide drug resistance. In this way, the research of more stable and efficient vaccines has been increasingly important to treat some infectious pathologies. DNA vaccines are new therapeutic strategies that consist in the synthesis of DNA vector-encoded therapeutic genes for specific proteins, which induce the immune responses against the pathogen [6]. DNA vaccination might provide several important advantages over conventional vaccines, such as facilitated development and production (more advantages are presented in Table 4) [54].

Table 4. Advantages of DNA vaccines (adapted from [55]).

Advantages of using DNA vaccines compared with conventional protein vaccines	
Immunogenicity	◊ Can induce both humoral and cellular immune responses
	◊ Low effective dosages (micrograms) in animal models
Safety	◊ Unable to revert into virulence unlike live vaccines
	◊ Efficacy does not require the use of toxic treatments unlike some killed vaccines
Engineering	◊ Plasmid vectors are simple to manipulate and can be rapidly tested
	◊ Combination approaches are easily adapted
Manufacture	Low manufacturing cost
	◊ Produced at high frequency in bacteria and easily isolated
Stability	◊ More temperature-stable than conventional vaccines
	◊ Long shelf-life
Mobility	◊ Ease of storage and transport
	◊ Likely not to require a cold chain

DNA vaccination has become an attractive immunization strategy due to the ability to induce strong and long-lasting cellular and humoral immune responses with high efficiency [5]. The DNA vaccine encoding therapeutic genes enters the cell nucleus, where the gene initiates the transcription, followed by protein production in the cytoplasm [6]. These proteins generate peptides that have been processed and presented by professional antigen-presenting cells (APCs), like macrophages and dendritic cells (DCs), in the context of major histocompatibility complex (MHC) class II molecules that then are exhibited on the surface of the cell [56]. Specific helper T cells (CD4$^+$ T cells) recognize this antigen peptide/MHC Class II molecule complex and are activated [45]. CD4$^+$ T cells play a central role in immune response. CD4$^+$ T cells secrete cytokines that have a myriad activities including, depending up on the cytokine, promoting B cell survival and antibody production and helping cytolytic T lymphocyte (CTL) responses [56]. In order to induce cytotoxic T lymphocyte response, CD8$^+$ T cells recognize peptides derived from endogenous proteins presented in the context of MHC class I molecules [54]. The presence of co-stimulatory molecules produced by activated CD4$^+$ T cells is another requirement for the activation of CD8$^+$ T cells, resulting in a potent antitumor effect against antigen-expressing tumors [55, 57]. On the other hand, DNA vaccines are capable to induce humoral immune response. In this way, B cells recognize antigens that are either present extracellularly, or exposed extracellularly by being transmembrane proteins inducing a strong antibody response to the respective proteins [56]. Figure 4 illustrate, the simplified form, the generation of both cytotoxic and helper T-cells responses and antibody generation by the DNA vaccines.

On the other hand, DNA vaccines are highly effective at inducing long-lived memory responses, so that upon subsequent exposure to the pathogen a rapid and specific immune response is mounted to prevent and/or clear infection [58].
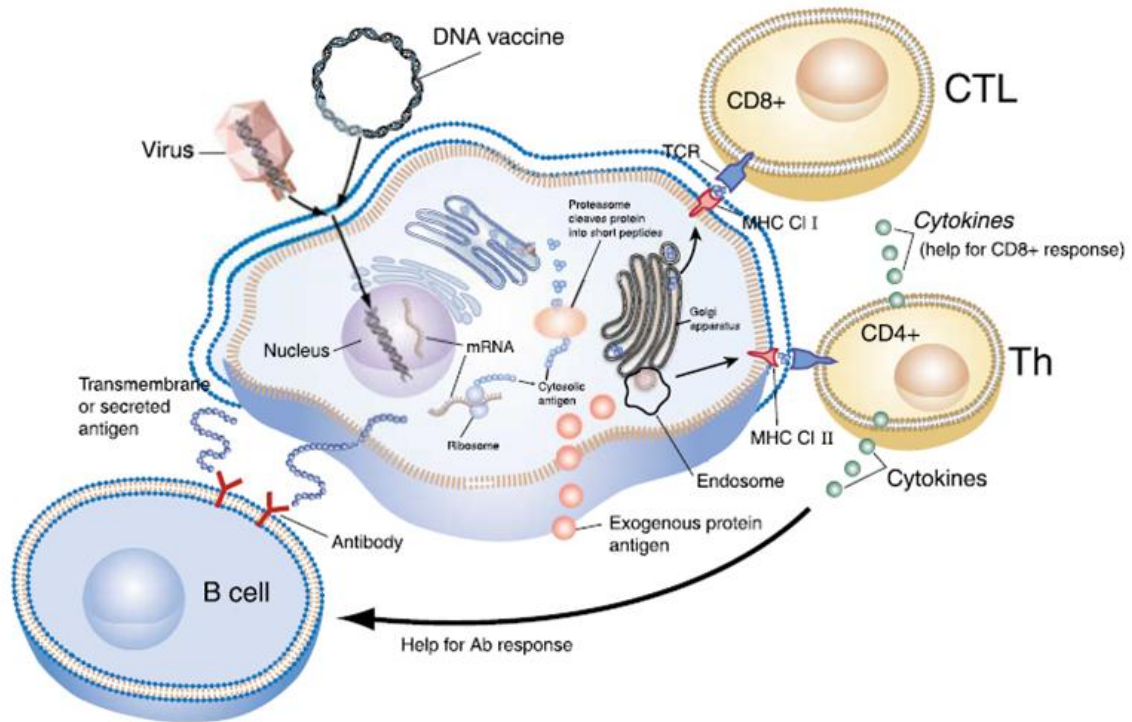
Figure 4. DNA vaccines stimulate the induction of both cellular and humoral responses. Professional antigen presenting cells take up an exogenous antigen into its endolysosomal degradation pathway, which convert proteins in peptides. These peptides are present in association with MHC class II molecules. After, CD4+ T cells recognize the peptide/MHC Class II molecule complex and are activated to produce cytokines. These cytokines help B cells activate into antibody producing cells, and help cytolytic T lymphocyte responses. Despite the action of this cytokines in cytolytic T lymphocyte responses, the recognition of the peptide/MHC class I molecules by the CD8+ T cells is necessary. For antibody responses, B cells recognize and respond to antigens that can present in transmembrane of the APC or present extracellularly (adapted from [45]).

2.2. Viral and Non-viral vectors

Gene medicines or nucleic acid drugs can be categorized on the basis of their therapeutic relevance as gene inhibitors or gene inductors [59]. Gene inhibitors (i.e. oligonucleotides, ribozymes, DNAzymes, aptamers, and small interfering RNAs (siRNAs)) are potent drugs that interfere with mRNA leading to silencing of the defective genes [60]. On the other hand, gene inductors (cDNA (complementary DNA) and plasmids containing transgenes) can be divided into gene vaccines and gene substitutes [59]. Gene vaccines are antigens of specific pathogens encoding either the genes or RNA that have the ability to induce T cell-mediated response and humoral immune response, as well as production of antibodies [61]. Gene substitutes are transcriptionally fully competent genes introduced into cells to reimburse deficiency of a specific protein or its insufficient protein production [60].

13

Central to the concept of gene therapy is the ability to deliver exogenous nucleic acids to wide variety of cells, tissues, and organs [50]. However, one of the problems associated to DNA vaccination is to develop the delivery vehicle of therapeutics genes that brings together 3 major criteria: (1) it should protect the gene against degradation by nucleases in intercellular matrices, (2) it should bring the gene across the plasma membrane and into the nucleus of target cells, and (3) it should have no detrimental effects [62]. Two major groups of vehicles for delivering genetic material exist, the viral and non-viral vectors [63].

2.2.1. Viral delivery systems

Viruses have numerous biological proprieties that allow them to recognize and enter in eukaryotic cells, expressing their genes and consequently infecting the host cell. These proprieties led to the use of viruses as gene delivery vectors [64]. Viral vectors encapsulate the viral sequences that are required for the assembly of viral particles, the packaging of the viral genome into the particles and the therapeutic gene [46]. Table 5 represents the main viral vectors (Retroviruses, Lentiviruses, Adenoviruses and Adeno-associated virus) and their advantages and disadvantages.

Table 5. Advantages and disadvantages of main viral vector groups (Adapted from [47]).

Vector	Advantages	Disadvantages
Adenovirus	◊ High transfection efficiency ◊ Transfects proliferating and nonproliferating cells ◊ Substantial clinical experience	◊ Strong immune responses ◊ Insert size limit of 7.5 kbp (kilo base pairs) ◊ Difficult to manufacture and quality control ◊ Poor storage characteristics ◊ Short duration of expression
Retrovirus	◊ Fairly prolonged expression ◊ High transfection efficiency ◊ Substantial clinical experience ◊ Low immunogenicity	◊ Low transfection efficiency in-vivo ◊ Insert size limit of 8 kbp *ex-vivo* ◊ Transfects only proliferating cells ◊ Difficult manufacture and quality control ◊ Safety concerns (mutagenesis)
Lentivirus	◊ Transfects proliferating and non-cells ◊ Transfects haematopoietic stem cells	◊ Very difficult manufacture and quality control ◊ Poor storage characteristics ◊ Insert size limit of 8 kbp ◊ No clinical experience ◊ Safety concerns (origins in HIV)
Adeno-associated virus	◊ Efficient transfection of wide variety of cell types *in-vivo* ◊ Prolonged expression ◊ Low immunogenicity	◊ Difficult manufacture and quality control ◊ Insert size limit of 4.5 kbp ◊ Limited clinical experience ◊ Safety concerns (mutagenesis)

In general, viral vectors are very efficient gene delivery vectors, however this delivery system also presents some disadvantages. These include safety concerns (mutagenesis and carcinogenesis), induction of immune responses (which abolishes the therapeutic gene, if consecutive administrations are made), the low DNA amount that can be loaded and the high commercial cost [63, 65]. Considering these limitations, the use of non-viral DNA-based therapy has emerged as a convincing approach in medical science.

2.2.2. Non-viral delivery system

Non-viral vectors are being developed to overcome problems associated with viral gene delivery, such as capacity for insertional mutagenesis [66]. In fact, non-viral vectors are less pathogenic and may have reduced toxicity by comparing to the existing viral vector, being safer and easier to manufacture under low cost. On the other hand, non-viral vectors also include as advantage the capacity of delivering larger genetic units without limited size.

Non-viral delivery systems for gene transfer can be classified into physical (carrier-free gene delivery) and chemical categories (synthetic vector-based gene delivery). Physical approaches employ a physical force that permeates the cell membrane and facilitates intracellular gene transfer. On the other hand, chemical approaches use synthetic or naturally occurring compounds as carriers to deliver nucleic acids into the target cells [60]. Although a significant progress has been made in the basic science and applications of various non-viral gene delivery systems, the non-viral approaches also present some limitations. In Table 6, briefly discuss the advantages and limitations of the non-viral gene delivery systems

Table 6. Advantages and limitations of several non-viral gene delivery systems (adapted from [62]).

Method	Advantages	Disadvantages
Physical methods		
Needle injection	Simplicity and safety	Low efficiency
Gene gun	Good efficiency	Tissue damage in some applications
Electroporation	High efficiency	Limited working range; need for surgical procedure for nontopical applications
Hydrodynamic delivery	High efficiency, simplicity, effectiveness for liver gene delivery	Extremely effective in small animals; surgical procedure may be needed for localized gene delivery
Ultrasound	Good potential for site specific gene delivery	Low efficiency *in vivo*
Chemical methods		
Cationic lipids	High efficiency *in vitro*; low to medium high for local and systemic gene delivery	Acute immune responses; limited activity *in vivo*
Cationic polymers	Highly effective *in vitro*; low to medium high for local and systemic gene delivery	Toxicity to cells; acute immune responses
Lipid/polymer hybrids	Low to medium-high efficiency *in vitro* and *in vivo*; low toxicity	Low activity *in vivo*

3. Plasmid DNA as non-viral vector

In therapeutic applications, plasmid DNA is used as vector or vehicle that can selectively and efficiently deliver a gene to target cells and induce the production of relevant proteins. In the last years, the use of plasmid-DNA-based delivery vectors has gained popularity (14% of all trials in 2004 and 18% in 2007), being the most popular non-viral system used in clinical trials [67]. This fact is probably due to the low toxicity and immunogenicity when compared with viral vectors [68].

Plasmids are circular, double stranded DNA molecules with a size range of 0.8 to 120 kilo base pairs, being considered very large biomolecules when compared to proteins. Plasmid DNA structure is divided into hydrophilic backbone (sugar and phosphate group) and hydrophobic interior of double helix (planar bases stacked on each other) [10]. For pH higher than 4, the phosphate groups are negatively charged, resulting in a negative DNA molecule [10, 69].

The supercoiled plasmid isoform is considered the most appropriate and active form for therapeutic applications due to its functional, compact and undamaged structure. However, other isoforms can be generated from the sc pDNA isoform by either single strand nick (open circular isoform) or double strand nick (linear isoform) [70]. The existence of different pDNA isoforms depends on factors such as DNA sequence, supercoiling stress or unfavorable environment conditions (extreme pH or high temperature) [10]. High temperatures promote the extended thermal motion, which leaded a gradual unwinding of the double DNA helix [10, 69]. In this way, several studies of transfection efficiency have been performed to investigate the biological effects of the pDNA topology and have revealed higher transfection efficiency with sc plasmid isoform [13, 71]. On the other hand, the application of plasmid DNA naked for transfection of eukaryotic cells results in lower transfection efficiency than comparing to viral vectors. Estimates indicate that only one in every thousand plasmid molecules presented to the cells can perform an efficient transfection [68]. The use of chemical or physical methods has significantly improved the gene delivery efficiency [72]. Hence, the best option is to develop adequate strategies to produce high copy number of sc pDNA, to obtain this plasmid isoform from the host avoiding its degradation, to eliminate the host impurities by an efficient purification method and to combine with a good physical or chemical-gene delivery system.

3.1. Plasmid DNA manufacturing

The simplicity of production methods is also one of the advantages of non-viral pDNA vectors for clinical applications instead of viral vectors. Over the years, several effective techniques for pDNA manufacturing have been developed, although the major challenge remains to achieve scalable and economical means of producing large quantities of sc pDNA fulfilling the requirements of regulatory agencies such as the Food and Drug Administration and the

European Agency for the Evaluation of Medical Products (EAEMP). In general, the manufacturing process starts by the plasmid vector construction and selection of appropriate bacterial host, followed by the choice and optimization of the fermentation conditions and subsequent isolation and purification steps [68], as it is presented in Figure 5.
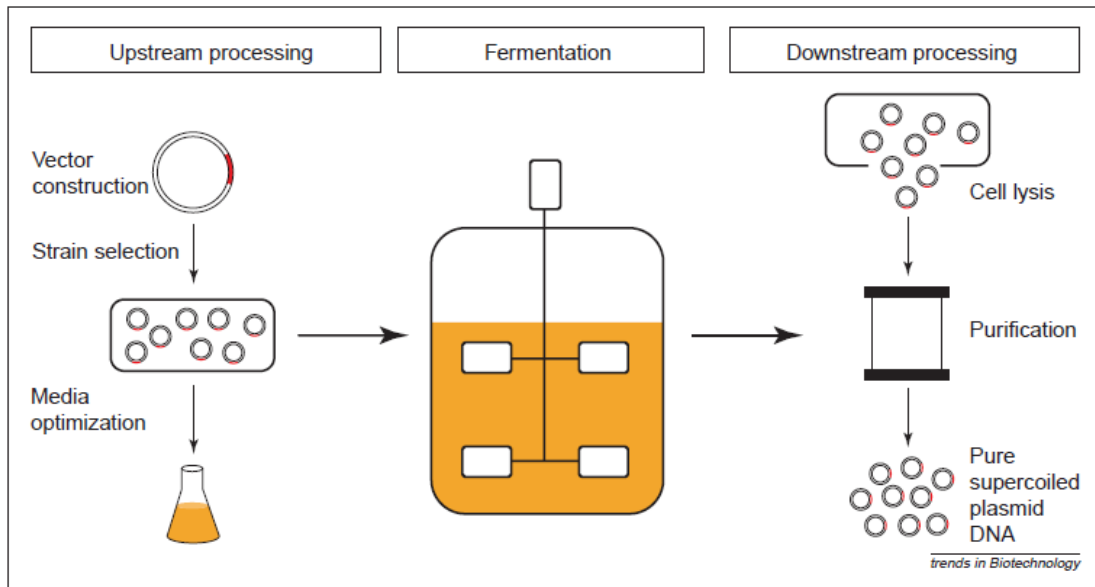
Figure 5. Three necessary stages to obtain a pure supercoiled plasmid DNA (adapted from [68]).

The design and engineering of pDNA is an important step to ensure the successfully pDNA manufacturing and to obtain the maximum degree of transfection. The basic structure of the most plasmid DNA applied in gene therapy and DNA vaccination includes: (1) an origin of replication for efficient propagation in an adequate host such as *Escherichia coli*, (2) an antibiotic resistance gene for growth selection, (3) a strong promoter to drive expression in eukaryotic cells, (4) a polyadenylation termination sequence, and (5) a gene insert coding for the antigen of interest [58].

Posteriorly, the choice of bacterial strain and selection of fermentation conditions are important factors to enable the production of large quantities of sc pDNA under stable conditions [68].

3.2. Downstream process

After fermentation, the specific plasmid content is only about 3% W/W of the *E. coli* extract [73]. Thus, a sequence of unit operations is essential for the release of plasmids from the host cells and their separation from other cellular constituents [58]. Although this downstream

process is critical for impurities elimination, also it is considered the critical step of the pDNA manufacturing process [74].

The first step in the pDNA downstream processing consists in *E. coli* cells recovery from the broth by centrifugation or microfiltration [74]. Thereafter, plasmid molecules can be released from host cells by several disruption techniques, although the alkaline lysis is the most applied technique [69]. Bearing this in mind, the alkaline lysis is based on cell disruption by using a buffer containing sodium hydroxide (NaOH) as alkali-promoted hydrogen bond disruption, followed by the release of all intracellular components (plasmid DNA, RNA, genomic DNA (gDNA), endotoxins and proteins). NaOH with high pH levels promotes the denaturation of genomic DNA, cell wall material and most of the cellular proteins. Peculiarly, sc pDNA isoform also unwinds as a consequence of the NaOH, but if the pH is lower than 12.5 prevents the separation of the complementary strands [69]. Following the lysis step, a precipitate containing cell debris, denatured gDNA and proteins is formed by adding the neutralizing solution with potassium acetate and finally the precipitate is removed by centrifugation [68].

Although most gDNA and proteins have been denatured and precipitated during alkaline lysis, the clarified sample resultant from alkaline lysate still contains proteins, RNA, gDNA, endotoxins and less than 1% (w/w) of pDNA. Concentration and clarification steps should be also performed to remove proteins, endotoxins and host nucleic acids and increase the plasmid mass fraction [69, 74]. The reduction of RNA and proteins are usually achieved by precipitation with a chaotropic salt, such as ammonium sulfate [69].

3.3. Plasmid DNA Purification

The purification of pDNA for research or clinical applications requires efficient technologies to obtain a final product constituted only by the biologically active supercoiled isoform (higher than 97%), according to guidelines established by regulatory agencies [68, 75]. Liquid chromatography is the central operation to separate the sc pDNA isoform from the others plasmid isoforms such as open circular (oc), linear (ln) and denatured isoforms together with the removal of host remaining impurities [75]. In addition, the physical and chemical similarities between impurities and pDNA, such as negative charge (RNA, gDNA and endotoxins), hydrophobicity (endotoxins) and molecular mass (gDNA and endotoxins) represent one of the major bottlenecks in the sc pDNA obtaining process [69]. Several chromatographic methods based on properties such as size, charge, hydrophobicity or affinity has been developed with more or less success in sc pDNA purification [10].

3.3.1. Size-exclusion chromatography (SEC)

Size-exclusion chromatography allows the plasmid fractionation and purification from other molecules present in a clarified lysate based on differences of their molecular size. Smaller molecules, such as RNA and endotoxins, have greater ability to penetrate inside the pore space, retarding the movement through the column. While, larger molecules ,such as gDNA and different forms of pDNA, travel around the particles of the packing material and are first eluted [76]. Thus, SEC can be used with high productivity in separation of pDNA from RNA and the other smaller impurities. However, the capacity of SEC to differentiate plasmid isoforms and isolate sc pDNA in one single step is limited [10].

3.3.2. Anion-exchange chromatography (AEC)

The simplest explanation of anion-exchange chromatography of nucleic acids is based on the attraction of oppositely charged molecules. Indeed, nucleic acids are considered polyanionic molecules that will interact with positively charged ligands immobilized in the stationary phase [68]. Variations in the salt concentration have been explored for binding/elution of different nucleic acids. According to general principles of anion-exchange, the binding should preferably occur at low salt concentration. After binding, the different nucleic acids should elute by increasing order of overall net charge, which in turn is dependent on the chain length and conformation of several molecules [69]. With an optimized gradient, the separation of supercoiled and open circular forms of pDNA in an anion-exchange column can be possible, because sc isoform is more compact and has a higher charge density than the open circular pDNA form. However, due to the existence of physical and chemical similarities between impurities and pDNA, like chemical composition and structure (gDNA and RNA) or charge (endotoxins), their separation can in some cases be insufficient [10].

3.3.3. Hydrophobic interaction chromatography (HIC)

Hydrophobicity is another physiochemical property that can also be explored in the purification of the sc pDNA isoform. This technique recognizes differences in hydrophobicity between double-strand and single-strand nucleic acid and endotoxins. In this way, the sample binding is achieved by column loading at a high salt concentration. The elution of different species is achieved by increasing order of hydrophobicity and by reducing the salt concentration, which weakens hydrophobic interactions. HIC technique showed to be efficient in the separation of pDNA from the RNA and endotoxins, nonetheless the ability to separate pDNA isoforms is weak. Moreover, the need for application of high salt concentrations is also view as a disadvantage [10].

3.3.4. Affinity chromatography (AC)

Affinity chromatography is the unique technique that uses specific binging agent to recognize and purify biomolecules based on their biological function or chemical structure [75]. Affinity chromatography has become a popular method for biomolecules purification due to advantages of eliminating additional steps, increasing yields and improving process economics. However this chromatographic method presents some limitations mainly in regard to the biological origin of ligands that tend to be fragile and associated to low binding capacity [10]. In this way, synthetic ligands, combining the selectivity of natural ligands with high capacity and durability of synthetics systems, have been developed [10].

Affinity chromatography separates biomolecules on the basis of reversible interactions between the target biomolecule and its specific ligand immobilized in the stationary phase (as it is showed in Figure 6). Thus, a sample containing the target biomolecule is injected onto the stationary phase, with appropriate pH and ionic strength in the buffer, allowing the binding between the target biomolecule and the specific ligand [77]. After binding, the elution steps can be performed either specifically, with competitive ligand, or non-specifically by changing buffer conditions, such as pH, ionic strength or polarity depending on the characteristics of the biomolecules [10].
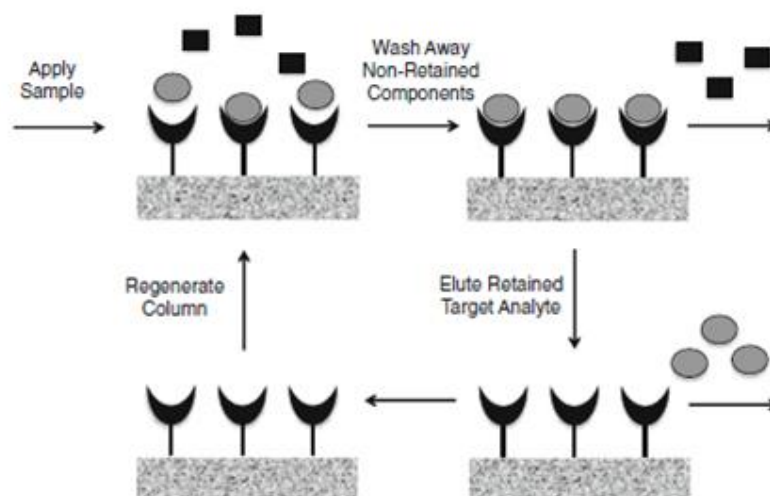
Figure 6. Principle of affinity chromatography. Firstly, a sample containing the target molecule is injected onto the column. After binding of molecules with specificity for immobilized ligands, a wash was applied for elution of the unbound species. Finally, alterations in elution buffer lead to elution of bounded species (adapted from [77]).

Several affinity chromatography types have been utilized in the sc pDNA purification with more or less success, among which, immobilized metal-ion affinity chromatography, triple-helix affinity chromatography, polymyxin B affinity chromatography, protein-DNA affinity

chromatography and more recently amino acid-DNA affinity chromatography [10]. Table 7 summarizes advantages and limitations of these affinity chromatography types.

Table 7. Affinity chromatography methods for plasmid DNA purification (adapted from [10]).

Affinity type	Principle	Specific binding	Advantages	Limitations
Immobilized metal-ion	Chelating ligands charged with divalent metal ions specifically interact with aromatic nitrogen atoms through π-d orbital overlap	Single-stranded nucleic acids (particularly purine bases)	Efficient resolution of RNA from gDNA and pDNA; High endotoxin removal; Separation of denatured pDNA	pDNA in the flow through; Incomplete RNA capture in complex mixtures; Co-elution of all DNA forms
Triple-helix	Specific sequences present on DNA are recognized by an immobilized oligonucleotide, forming a triple-helix	Double-stranded DNA	Discrimination of different plasmids based on their sequence; sc pDNA isolation in one chromatographic step; Reduction of RNA, gDNA and endotoxin contamination levels; Possibility for scale-up	Loss of pDNA during wash step; Low yields; Slow kinetics of triple-helix formation; Long chromatographic run times
Polymyxin B	Immobilized polymyxin B specifically recognizes the lipid structure of endotoxins	Endotoxins	Elimination of endotoxin contamination from pDNA preparations	Non-specific interaction of ligands with pDNA; Poor yields; Toxicity of polymyxin B
Protein-DNA	A protein or protein complex immobilized on the matrix specifically recognizes a DNA motif	pDNA	Discrimination of different plasmids based on their sequence; pDNA isolation from clarified lysates; Elimination of proteins and RNA from preparation	Relatively low yields; Contamination with gDNA
Amino acid-DNA	Multiple interactions occur between immobilized amino acids and nucleic acids	sc pDNA	sc pDNA purification in a single chromatographic step; Efficient elimination of RNA, gDNA, proteins and endotoxins	Elution with high salt concentration and relatively low yields (for histidine)

In the last years, amino acids have been used in biotechnological applications like solvent additives for protein purification due to their natural and safe composition [78]. On the other hand, atomic evidences suggested the existence of favorable interactions between some amino acids and nucleic acid bases as well as the natural occurrence of protein-DNA complexes in biological systems [79, 80]. Based on these facts, the use of amino acids as affinity ligands has proven to be a promising strategy to efficiently isolate pDNA with high yield and purity [10]. Thus, our research group has applied a new affinity chromatographic methodology by using some amino acids immobilized on conventional agarose matrix for the isolation and purification of supercoiled pDNA isoform. In fact, histidine and lysine amino acids have been explored as affinity ligands and the sc pDNA isoform was specifically isolated from a clarified *E. coli* lysate, however they also presented relatively low yield [11, 12]. On the other hand, arginine amino acid revealed the involvement of specific interactions with plasmid molecules, namely sc pDNA isoform, being purified from a clarified *E. coli* lysate with high yield and purity degree [13]. Although the amino acid ligands have increased the specificity and recognition by the sc pDNA isoform, these conventional stationary phases present some limitations, related to the low capacity and diffusivity of pDNA samples [14]. Furthermore, conventional supports can lead to extended chromatographic residence time and back pressure that affects the plasmid structural integrity, recovery and peak resolution [81].

3.4. Monolithic Technology

Monolithic chromatographic supports are considered satisfactory alternatives to conventional supports due to their advantageous structural characteristics. Monolithic support is a continuous bed prepared by polymerization of different materials, resulting in a single piece of highly porous solid material, whose pore size depends on the polymerization process [81]. This structural appearance of the monolithic support eliminates some problems related to the scale-up and scale-down variations in packing quality and the need to repack a column due to the inadvertent introduction of air bubbles [81]. As the monolithic material can be directly polymerized into the column forming a three-dimensional network, all mobile phase is forced to flow through the channels via convection, resulting in a very fast mass transfer [71]. The large channels of monolithic supports allow the easily accommodate of large pDNA biomolecules in their internal surface area, under reduced pressure drops even at high flow rates [81]. The association between excellent mass transfer and easy accommodation of the pDNA in large porous of monolith justifies the high binding capacity for pDNA [14]. Thus, the monolithic technology brings great improvements of the speed, resolution and capacity to the chromatographic performance in the sc pDNA isoform purification.

Chapter II

Material and Methods

2.1. Materials

All experiments were carried out in the 0.34 mL (3 mm thick and 12 mm diameter) modified monolith packed into a CIM^{TM} disk housing, kindly prepared and provided by BIA Separations (Ljubljana, Slovenia). Arginine–Sepharose 4B gel was obtained from Amersham Biosciences (Uppsala, Sweden). The Qiagen Plasmid Purification Maxi Kit was from Qiagen (Hilden, Germany). Sodium chloride and ammonium sulfate were purchased from Panreac (Barcelona, Spain) and tris(hydroxymethyl) aminomethane (Tris) from Merck (Darmstadt, Germany). All solutions used in chromatographic experiments were freshly prepared using deionized water ultra-pure grade, purified with a Milli-Q system from Millipore (Billerica, MA, USA) and analytical grade reagents. The elution buffers were filtered through a 0.20 μm pore size membrane (Schleicher Schuell, Dassel, Germany) and degassed ultrasonically. The iQ SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) was used for genomic DNA quantification.

2.1.1. Plasmid DNA

The 8.702 kbp HPV-16 E6/E7, Addgene plasmid 8641 [82] was obtained from Addgene (Cambridge, USA). The vector encodes for the human HPV 16 E6/E7 proteins. The vector contains the human beta-actin mammalian expression promoter and the ampicillin resistance gene (Figure 7).
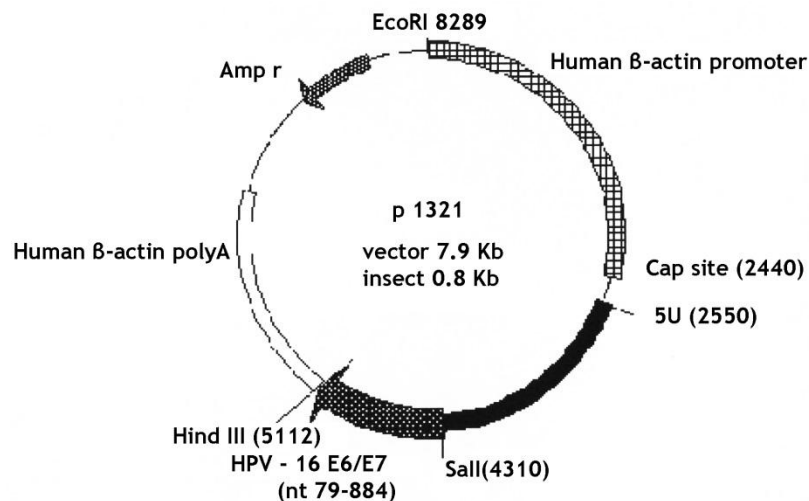
Figure 7. Plasmid HPV-16 E6/E7 backbone.

The 6.05-kbp plasmid pVAX1-*LacZ* was obtained from Invitrogen (Carlsband, CA, USA).

2.2. Methods

2.2.1. Plasmids and bacterial growth conditions

The amplification of HPV-16 E6/E7 and pVAX1-*LacZ* plasmids was obtained by *E. coli* DH5α fermentation. Growth was carried out 37 ℃ under 250 rpm shaking, using Terrific broth medium (20 g/L tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 0.017 M KH_2PO_4 and 0.072 M K_2HPO_4) supplemented with 100 μg ampicillin/mL for cells with HPV-16 E6/E7 and 30 μg kanamycin/mL for cells with pVAX1-*LacZ*. Growth was suspended at the late log phase [OD_{600}~11] and cells were recovered by centrifugation, at 4500 g for 10 min at 4 ℃, and pellets were stored at -20 ℃.

2.2.2. Alkaline lysis with Qiagen Kit

The plasmid DNA was obtained by alkaline lysis using the Qiagen plasmid maxi kit according to the manufacturer's instructions. Initially, the protocol was based on a modified alkaline lysis procedure for cellular debris elimination. Then, the binding of pDNA to Qiagen Anion-Exchange columns was promoted under appropriate low-salt and pH conditions. RNA, proteins and low-molecular-weight impurities were removed by a medium-salt wash while the pDNA was eluted in a high-salt buffer and then concentrated by isopropanol precipitation. The final plasmid DNA sample was applied in preparative chromatography and dynamic binding capacity studies.

2.2.3. Alkaline lysis and primary isolation of plasmids

After the bacterial growth, cells were lysed through of a modification on the alkaline method [83] as described by Diogo et al., 2000 [84]. Bacterial pellets resulting from centrifugation were resuspended in 20 mL of solution I (50 mM glucose, 25 mM Tris-HCl and 10 mM ethylene-diamine tetraacetic acid (EDTA) at pH 8.0). Alkaline lysis was performed by adding 20 mL of solution II (200 mM NaOH and 1% (w/v) sodium dodecylsulfate (SDS)) and 5 minutes incubation at room temperature. The solution II was neutralized by adding 16 mL of solution III (3 M potassium acetate at pH 5.0) and followed by 20 minutes incubation in ice. Cellular debris, gDNA and proteins were eliminated by centrifuging twice at 20 000 g (30 minutes, 4 ℃) with a Allegra^{TM} 25R centrifuge (Beckman Coulter).

The concentration and reduction of impurity content was performed according to a previously published method [84]. Briefly, DNA present in supernatant was precipitated by adding 0.7 volumes of isopropanol. After 30 minutes incubation in ice, the precipitated was recovered by

centrifugation at 16 000 g (30 minutes, 4 ℃). The pellet was then redissolved in 4 mL of 10 mM Tris-HCl and 10 mM EDTA buffer, pH 8.0. Proteins and RNA were precipitated by adding ammonium sulfate at pDNA solution up to a final concentration of 2.5 M. After 15 minutes incubation in ice, impurities were removed by centrifugation at 16 000 g (20 minutes, 4 ℃). The supernatant was desalted by passing through PD-10 desalting columns according the manufacturer's instructions, using 10 mM Tris-HCl and 10 mM EDTA (pH 8.0) solution as the elution buffer.

2.2.4. Affinity chromatography

Chromatographic experiments were performed using an AKTA Purifier system (GE Healthcare Biosciences Uppsala, Sweden) consisting of a compact separation unit and a personal computer with Unicorn control system version 5.11. A CIM™ arginine-epoxy disk with the column bed volume of 0.34 mL and average pore sizes of 1500 nm was prepared from a CIM™ epoxy disk. The immobilization procedure consisted of pumping the arginine solution through the monolithic disk (2 g arginine dissolved in 8 mL of HCl) followed by 72 h thermostating of the column at 60 ℃. Finally, the monolithic disk was washed with water and the remaining epoxy groups were hydrolyzed by soaking with 0.5 M H_2SO_4 for 30 min.

In the isoforms separation step of the HPV-16 E6/E7 plasmid, the monolithic column was equilibrated with 560 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at a flow rate of 1 mL/min. Plasmid sample resulting from alkaline lysis with Qiagen Kit was injected into the column using a 200 µL loop in same flow conditions. After elution of unbound species, the ionic strength was increased stepwise to 2 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) for elute the bound species. Fractions were pooled, according to the chromatograms obtained, concentrated and desalted with visvapin concentrators (Vivascience) for further electrophoresis analysis.

In the sc HPV-16 E6/E7 pDNA purification from a complex *E. coli* lysate sample, the monolithic column was equilibrated with 600 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at 1 mL/min. Complex lysate sample resulting from alkaline lysis was injected into the column using a 500 µL loop. After elution of unbound species the ionic strength was increased stepwise to 740 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), to 795 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) and then to 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) in order to gradually elute bound species. Fractions were pooled, according to the chromatograms obtained, concentrated and desalted with visvapin concentrators. Fractions were also used for electrophoresis analysis and impurities (gDNA, proteins and endotoxins) quantification.

All experiments were accomplished at room temperature and the absorbance was continuously monitored at 260 nm. After chromatography runs the column and AKTA Purifier system were cleaned with MiliQ water.

2.2.5. Agarose gel electrophoresis

Pooled fractions from each chromatographic experiment were analyzed by horizontal electrophoresis using 15-cm-long 0.8% agarose gels (Hoefer, San Francisco, CA, USA) stained with greensafe (1 µg/mL). Electrophoresis was carried out 110 V, for 30 minutes with TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA, pH 8.0). The gel was visualized under UV light in a Vilber Lourmat system (ILC Lda, Lisbon, Portugal). Hyper Ladder I (Bioline, London, UK) was used as a DNA molecular weight marker.

2.2.6. Dynamic binding capacity

In order to compare the dynamic binding capacity between a conventional matrix and a monolith, these two supports were prepared and connected to an AKTA Purifier system. A 10 × 7 mm glass column (GE Healthcare Biosciences, Uppsala, Sweden) was packed with the commercial L-arginine-agarose gel (Sigma, St Louis, MO, USA) giving a total bed volume of 500 µL. A CIM^{TM} epoxy monolith was modified by immobilization of arginine ligands. These experiments were performed at the same flow rate and pDNA concentration (1 mL/min and 50 µg/mL respectively). The HPV-16 E6/E7 plasmid solution was obtained by several alkaline lysis processes using the Qiagen Kit following by homogenization in 10 mM Tris-HCl and 10 mM EDTA buffer (around 130 mL). Each column was equilibrated with 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) and thereafter, the columns were overloaded with the plasmid solution under the same equilibrium conditions. Determination of dynamic binding capacity was carried out by recording breakthrough curves and calculating the amount of bound pDNA per mL support at 10%, 50% and 100% breakthrough curve. Dynamic binding capacity was calculating according to equation (1) [85]

$$q = \frac{(V_{breakthrough} - V_d)C_0}{V_c}$$

(1)

where $V_{breakthrough}$ is the volume corresponding to the feedstock loaded at 10% or 50% breakthrough based on recorded absorbance (mL), V_d is the void volume determination from bypass run in milliliter (dead volume), C_0 is the pDNA concentration in the feedstock (mg/mL) and finally, V_c represents the monolith bed volume (mL).

Afterwards, the elution of the bound plasmid was achieved by incre asing the ionic strength to 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer. Finally, both supports were cleaned and regenerated with a 0.1 M NaOH solution.

2.2.7. Plasmid quantification

The pDNA concentration and purity of the pooled fractions from the monolithic chromatographic experiment and samples collected throughout of the alkaline lysis process were evaluated, according to an adaptation of the analytical method based on hydrophobic interaction chromatography previously described [86]. A 4.6/100 mm HIC source 15 PHE PE column (Amersham Bioscience) was connected to an AKTA Purifier system (GE Healthcare Bioscience) and equilibrated with 1.5 M $(NH_4)_2SO_4$ in 10 mM Tris-HCl buffer (pH 8.0). Then, 20 µL of the sample suitably diluted in the equilibrium buffer was injected and eluted at 1 mL/min. After 4 minutes, the elution buffer was instantaneously changed to 10 mM Tris-HCl buffer pH 8.0 without ammonium sulfate. This elution condition was maintained for 6 minutes until all retained species have been eluted. Finally, the column was re-equilibrated for 10 minutes with the equilibrium buffer to prepare the column for the next run. The absorbance of the eluate was record at 260 nm. The pDNA concentration in each sample was calculated by using a calibration curve constructed with pDNA standards of 1 to 100 µg/mL (as it is presented in Figure 8), purified with a commercial Qiagen kit (Hilden, Germany). The purity degree was defined as the percentage of the pDNA peak area related with the total area of all chromatographic peaks.

y = 0.1079x - 0.2888
R^2 = 0.9985
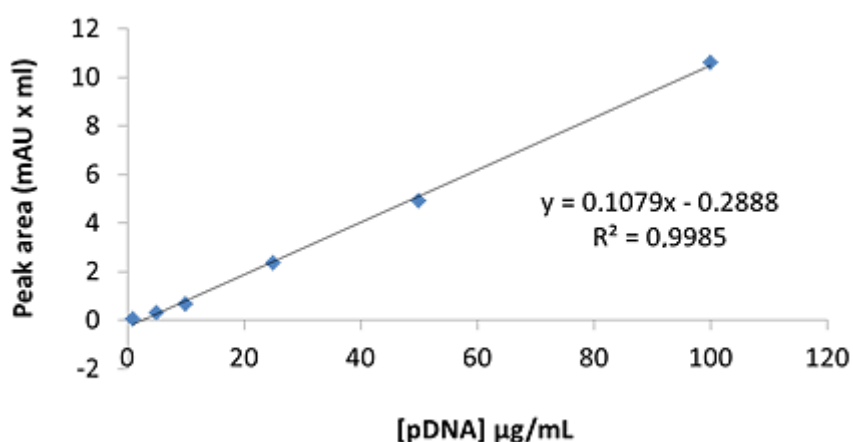
Figure 8. Reference curve plasmid DNA standards (1-100 µg/ml).

2.2.8. Protein quantification

Protein content was assessed by the micro-BCA (bicinchoninic acid) protein assay kit from Pierce (Rockford, USA). Brief, 50 µL of each sample was added to 200 µL of BCA reagent in a microplate. After 30 minutes of incubation ate 60 °C, absorbance was recorded at 595 nm in microplate reader. The calibration curve (Figure 9) was constructed with Bovine Serum Albumin (BSA) as a standard protein (0.01-0.1 mg/ml).
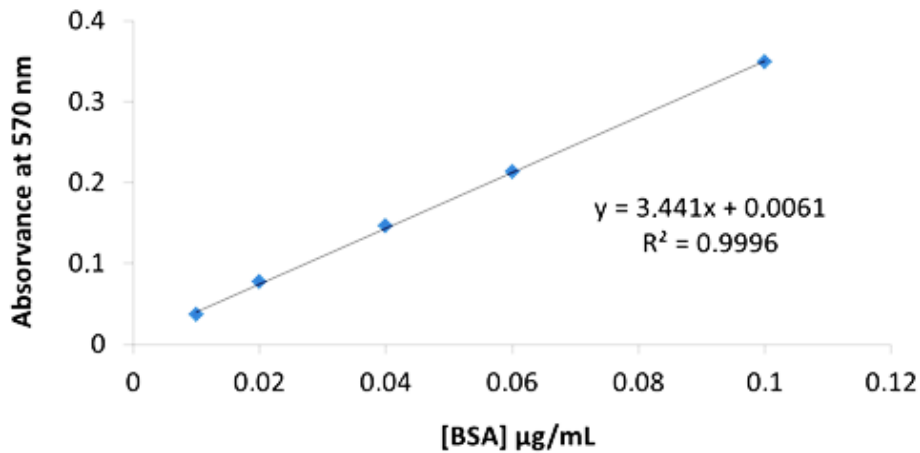
Figure 9. Reference curve Bovine Serum Albumin standards (0.01-0.1 mg/ml).

2.2.9. Genomic DNA quantification

Genomic DNA measurements were obtained by real-time polymerase chain reaction (PCR) in a iQS Multicolor Real-Time PCR Detection System (BioRad), such as described previously [87]. Sense (5'-ACACGGTCCAGAACTCCTACG-3') and antisense (5'-CCGGTGCTTCTTCTGCGGGTAACGTCA-3') primers were used to amplify a 181-bp fragment of the 16S rRNA gene. PCR amplicons were quantified by following changes in fluorescence of the DNA binding dye Syber Green I. The calibration curve to achieve the gDNA concentration was constructed by a serial dilutions of the *E. coli* DH5α gDNA sample (purified with the Wizard gDNA purification kit; Promega) in the range of 0.005 to 50 ng/µL, as it is presented in Figure 10.
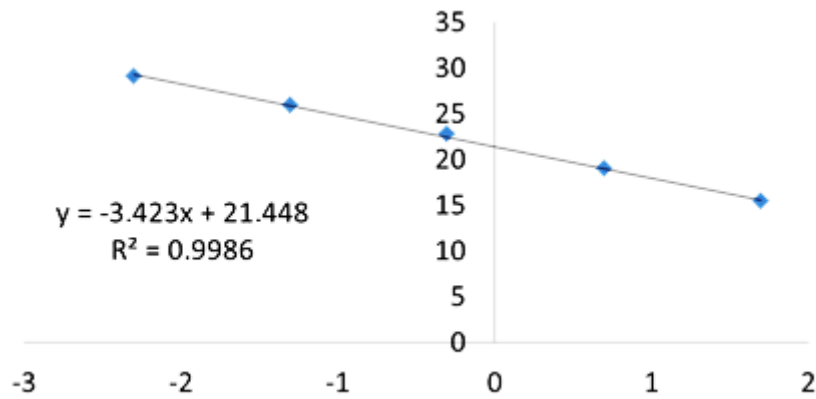
Figure 10. Reference curve *E. coli* DH5α genomic DNA standards (0.005 - 50 ng/µL).

2.2.10. Endotoxin quantification

Endotoxin contamination was evaluated by using the ToxinSensor™ Chromogenic *Limulus* amoebocyte lysate assay kit from GenScript (USA, Inc.) according to manufacturer's instructions. The calibration curve (Figure 11) was constructed with 10 EU/mL stock solution provided with the kit (0.01-0.1 EU/mL). All tubs and tips or diluents used to perform this quantification must be endotoxin-free. In this way, entire procedure was performed inside of a laminar flow cabinet. Samples to analyze and samples from the kit were diluted, or dissolved respectively, with non-pyrogenic water, which was also used as the blank.
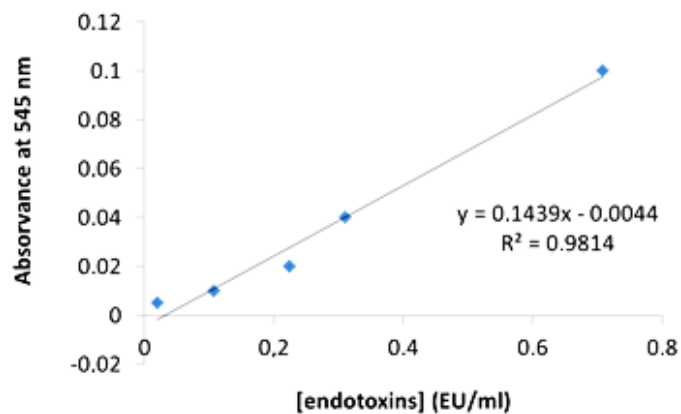
Figure 11. Reference curve endotoxins standards (0.01-0.1 EU/mL).

31

Chapter III

Results and discussion

In order to improve the use of therapeutics based on plasmid DNA, the development of new approaches capable of efficiently produce and purify the supercoiled plasmid isoform is increasingly required. Several chromatographic processes have been exploited in the pDNA purification, making affinity chromatography an attractive purification strategy due to advantages of eliminating additional steps, increasing yields and improving processes economics [10]. Recently, the use a novel approach based on interactions between amino acid and nucleic acids has showed a specific biorecognition by the sc plasmid isoform, allowing the elimination of impurities and achieving the desired purity degree [13]. However, this strategy only ensures the specific recognition of the sc plasmid isoform [14]. In addition, other factors like dynamic binding capacity, fast isolation and high resolution between sc pDNA sample and host impurities should be considered. Thereby, monoliths have emerged as an excellent solution for pDNA purification, ensuring the fast and efficient separation process with low back pressure, high binding capacity and maintaining the stability and productivity for plasmids [81].

In this way, the main goal of this thesis is to combine the specificity and selectivity achieved by the arginine amino acid as affinity ligand with the versatility of the epoxy-based monolithic support to attain large amounts of homogeneous preparations of sc HPV-16 E6/E7 DNA vaccine. A posterior characterization of the modified monolith was performed in order to determine available binding sites for the pDNA and to compare with a conventional arginine-agarose matrix. In addition, analytical characterization of the sc pDNA recovered from the arginine-epoxy monolith will be performed in order to confirm that the sample is suitable for immunogenic therapies against HPV infection (according to guidelines from FDA and other regulatory agencies).

3.1. Plasmid amplification in recombinant *E. coli* and electrophoresis optimization

The production and recovery of high copy number of supercoiled plasmid DNA is fundamental for therapeutic applications. Hence, factors as the host organism, the recombinant plasmid vector and the growth environment must be taken into account to obtain an efficient manufacturing process, which can provide the plasmid production in large quantities for *in vitro* and *in vivo* transgene expression [88]. *Escherichia coli* strains have been considered the efficient host in the pDNA production, not only yielding high amounts of the vector but also maintaining its integrity [89]. The plasmid HPV-16 E6/E7 amplification was obtained by *E. coli* DH5α batch-fermentation, and approximately 20 h of growth were required to attain a high cell density.

The pDNA amplification process in the *E.coli* host yielded a higher content of the sc pDNA isoform in comparison to the oc form. This fact is very important because the oc isoform is less therapeutically effective in transfection and heterologous expression than supercoiled

plasmid DNA. Sousa F. and coworkers demonstrated that the sc pDNA was able to transfect 62% of eukaryotic cells, a higher efficiency when compared to the one achieved with the oc isoform (16%) [13]. Nevertheless, despite sc pDNA be produced in *E.coli* further purification of this isoform is critical for gene therapeutic applications.

GreenSafe is a nucleic acid stain that can be used as a safer alternative to the traditional ethidium bromide stain for detecting nucleic acids in agarose gels, due to its advantages such as low toxicity, mutagenicity and carcinogenicity. Although the GreenSafe is a larger biomolecule than ethidium bromide, the sensitive is the same and can be used exactly in the same way in agarose gel electrophoresis. GreenSafe stain emits fluorescence when bound to DNA or RNA and radiated by UV light.

In the first step, a HPV-16 E6/E7 plasmid sample (oc and sc isoforms) obtained with the Qiagen kit was analyzed by horizontal electrophoresis using 1% agarose gel stained with 20 μL of GreenSafe. The conditions used in the gel preparation were previously optimized for detecting other nucleic acids, namely the pVAX1-*LacZ* plasmid that has lower molecular size than the HPV-16 E6/E7. The agarose gel electrophoresis for this sample is depicted in Figure 12.

Figure 12. Agarose gel electrophoresis analysis of HPV-16 E6/E7 plasmid sample obtained with the Qiagen Kit. The agarose gel was stained with 0.25 μg/mL of GeenSafe.

In fact, the first agarose gel electrophoresis analysis showed that only one band was present in the plasmid sample. However, this result is in disagreement with other studies, which refer that a plasmid sample resulting from the alkaline lysis normally has the presence of two bands, corresponding to the oc and sc plasmid isoforms [51, 90]. This discrepancy can be due to two different causes: either there is absence of the oc isoform in the final plasmid sample, or there can be issues related to GreenSafe staining.

Unfavorable conditions such as extreme pH or high temperatures induce pDNA conformational changes. Freitas S. and coworkers demonstrated that an increase in temperature leads to an increased unwinding of the sc pDNA isoform [91]. This decrease in sc isoform content was accompanied by an increase in the oc isoform [91].

In this way, the HPV-16 E6/E7 pDNA sample was left at room temperature during four days. After that, the sample was again analyzed by horizontal electrophoresis at the same gel conditions (1% agarose and 0.25 μg/mL of GreenSafe). The result of this new analysis showed once more the presence of one only band (Figure 13).
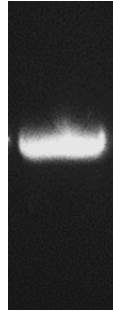
Figure 13. Agarose gel electrophoresis analysis of HPV-16 E6/E7 plasmid sample obtained with the Qiagen Kit, after four days at room temperature. The agarose gel was stained with 0.25 µg/mL of GeenSafe.

This experiment showed that the problem is not the absence of the oc isoform, and probably can be related with issues in GreenSafe staining. Thus, in order to confirm the presence of oc isoform in the sample, agarose gel electrophoresis was performed using 1% agarose stained with 0.5 µg/mL of ethidium bromide (Figure 14).
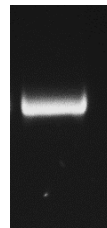

Figure 14. Agarose gel electrophoresis analysis of HPV-16 E6/E7 plasmid sample obtained with the Qiagen Kit. The agarose gel was stained with 0.5 µg/mL of ethidium bromide.

Figure 14 revealed the presence of two bands corresponding to the oc and sc plasmid isoforms. This result confirmed that the initial problem was associated to the use of GreenSafe as stain for detecting nucleic acids in agarose gels. As the GreenSafe biomolecule is larger than the ethidium bromide, its association with the HPV-16 E6/E7 plasmid results in a complex of high molecular size. Thereby, the difference in both electrophoresis types can be related with the size of this complex, which hampers the plasmid migration through the agarose pores, and consequently avoids the differential migration of both plasmid isoforms. Bering this in mind, an optimization of the agarose gel formulation stained with GreenSafe is required and a compromise between the agarose content and GreenSafe concentration should be achieved.

Therefore, several experiments were performed to achieve the best conditions and obtain the representation of the two bands relatively to the oc and sc isoforms, by using different agarose percentages and GreenSafe volumes (Figure 15).
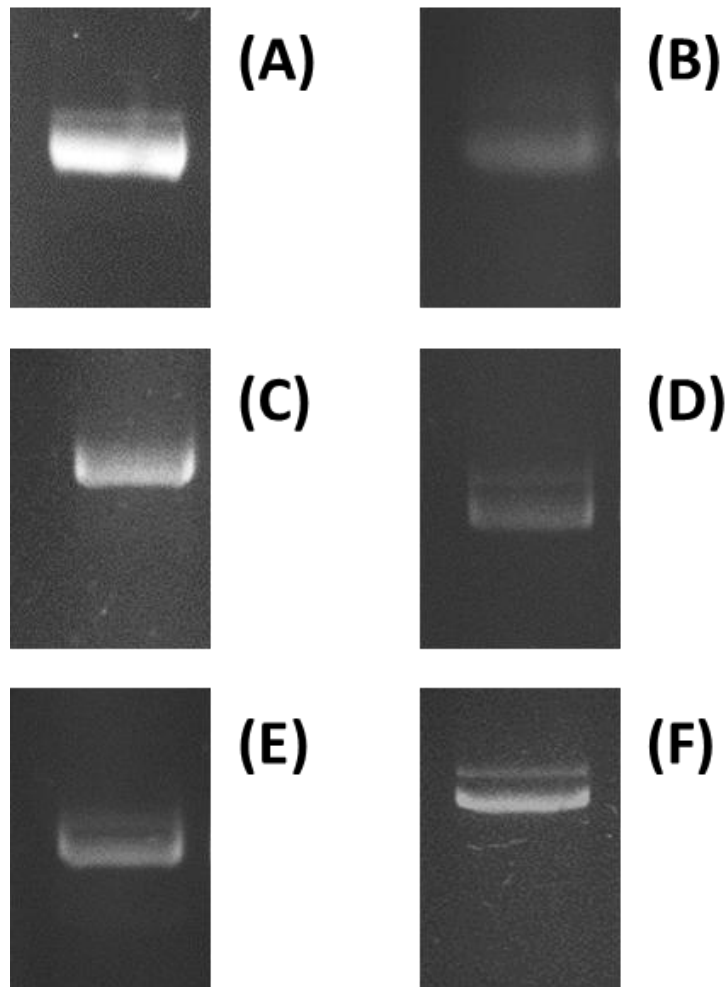
Figure 15. Agarose gel electrophoresis analysis of plasmid sample obtained with the Qiagen Kit at different agarose percentages and Greensafe volumes (mother solution with 0.5 µg/mL): (A) 0.6% agarose and 10 µL of GreenSafe; (B) 0.6% agarose and 50 µL of GreenSafe; (C) 0.8% agarose and 30 µL of GreenSafe; (D) 0.8% agarose and 75 µL of GreenSafe; (E) 1% agarose and 15 µL of GreenSafe; (F) 1% agarose and 50 µL of GreenSafe.

The electrophoresis gel results showed that 0.6% agarose and 10 µL of GreenSafe (Figure 16 (A)) revealed two bands although they are not clearly separated, whereas using 50 µL of GreenSafe (Figure 16 (B)) only revealed one band. In experiments using 0.8% of agarose and 30 µL of GreenSafe (Figure 16 (C)) was verified the presence of one band, whereas using 75 µL of GreenSafe (Figure 16 (D)) two bands were observed, although also not clearly visualized. These results indicated that the increase of the agarose content could improve the plasmid isoforms separation. Thus, the best result was obtained for 1% of agarose and 50 µL of GreenSafe, where the plasmid isoforms were completely and clearly separated (Figure 16 (F)).

3.2. Separation of sc pDNA from oc pDNA

The first screen was performed to verify the modification of the arginine-epoxy monolith by comparing to a non-modified epoxy monolith. Thereby, the non-modified epoxy column was equilibrated with 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0). After injection of 200 µL of native plasmid DNA (oc and sc isoforms) resulting from alkaline lysis with a Qiagen kit, a first and unique peak was rapidly attained in the flowthrough due to the elution of species with lower affinity to the epoxy matrix (Figure 16). Afterward, a NaCl linear gradient was performed from 0 M to 2 M in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) but no specie was eluted (Figure 16).
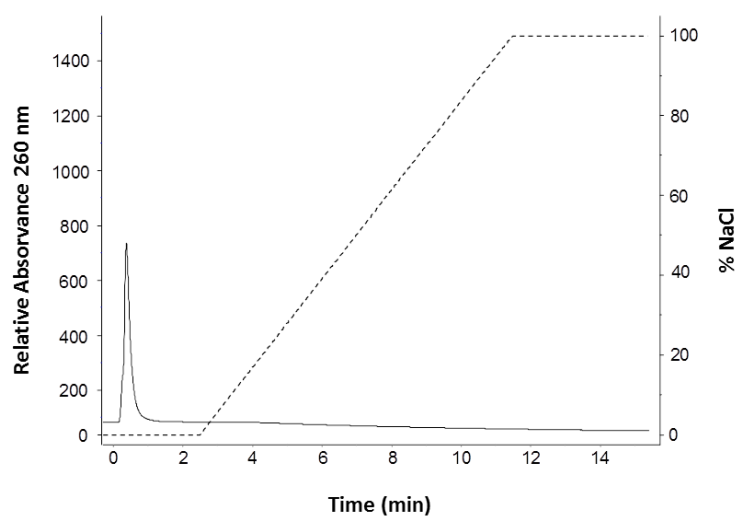
Figure 16. Evaluation of the non-immobilized epoxy monolith role in the oc and sc plasmid isoforms separation. A pre-purified pDNA sample (oc + sc) of 200 µL was injected and a linear gradient was performed at 1 mL/min by increasing the NaCl concentration in the eluent from 0 M to 2 M in 10 mM Tris-HCl.

In fact, the chromatographic profile of the first screen showed the existence of only one peak obtained in the flowthrough, which represents the elution of species with lower affinity to the matrix. These results suggest that epoxy groups of the non-immobilized monolith did not interact with pDNA molecules under ionic conditions.

Given the satisfactory results achieved with conventional arginine-agarose matrices in the efficient separation of sc and oc pDNA isoforms [90], the same ligand was immobilized in an epoxy monolith. In this way, the same chromatographic experiment applied for the non-immobilized epoxy monolith was performed in epoxy monolith modified with arginine amino acid (Figure 17).
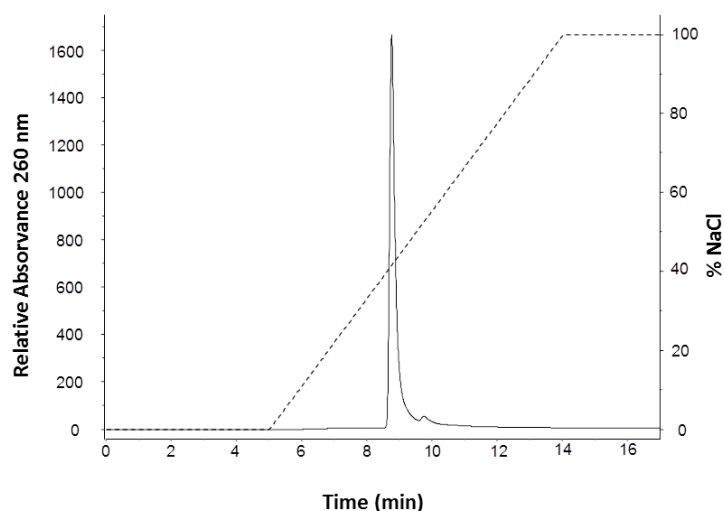
Figure 17. Evaluation of epoxy-monolith modification by immobilization of arginine amino acid. (A) A pre-purified pDNA sample (oc + sc) of 200 μL was injected and a linear gradient was performed at 1 mL/min by increasing the NaCl concentration in the eluent from 0 M to 2 M in 10 mM Tris-HCl, by using the arginine-epoxy monolith.

In this chromatographic assay, the total pDNA (oc and sc isoforms) retention was obtained at 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) and the total elution was verified in a single peak during a linear gradient by increasing the NaCl concentration in buffer from 0 M to 2 M (Figure 17). This result confirmed that arginine ligands are immobilized on the epoxy monolith and are responsible for pDNA sample binding to the column. In this way, several binding/elution studies were essential in order to determinate the ideal conditions for total isoforms separation.

Initial experiments were performed to choose the best salt concentration for binding/elution conditions in order to determine the concentration range where total isoforms separation is achieved. Arginine-epoxy monolithic column was first equilibrated with 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at room temperature and a flow rate of 1mL/min. After plasmid sample injection, several increasing NaCl concentrations were applied in order to evaluate the retention and elution of each plasmid isoform, as schematized in Table 8.

Table 8. Summary of binding and elution of oc and sc plasmid isoforms in arginine-epoxy monolithic chromatography.

	Elution step [NaCl] (mM)		
pDNA Sample loaded	400 mM	600 mM	2000 mM
oc pDNA	Retention	Elution	-
sc pDNA	Retention	Retention/Elution	Elution

The data present in Table 8 show that the total (oc + sc) pDNA retention was obtained at 400 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), and partial elution was verified at of 600 mM NaCl buffer (pH 8.0).

Experiments performed to establish binding/elution conditions showed a relation between plasmid isoforms and the ionic strength of the used buffer. In fact, when the pDNA native sample was loaded into the arginine-epoxy monolithic column with 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), both oc and sc isoforms were retained in the column. On the other hand, the increase of ionic strength led to elution of the sample, indicating that the interaction of pDNA with the matrix was weakened by competition. This phenomenon can be explained by the presence of NaCl in mobile phase that reduces the electrostatic attraction exerted by the matrix surface and affects the hydrodynamic size of plasmid isoforms [92]. Bearing these criteria into account, plasmid isoforms separation was a consequence of size compaction and electronegativity reduction of plasmid molecules with different conformations by increase of the NaCl concentration.

The diversity of interactions associated to arginine ligands (e.g., electrostatic, hydrophobic, affinity) suggest the occurrence of multiple contacts either with pDNA bases or the pDNA backbone. As previously mentioned, electrostatic interactions are easily predicted and favorable between the negative character of phosphate groups in the pDNA backbone and the positive character of arginine ligands at pH 8.0. These phenomena are common in many protein–DNA complexes, providing stability rather than specificity [80]. The involvement of electrostatic interactions is also confirmed by the salt concentration needed to elute plasmid isoforms from arginine-epoxy monolith, comparing to other anion-exchange chromatographic supports (higher than 500 mM NaCl) [93].

According to results in Table 8, new experiments were performed to find the exact concentrations for total separation of the pDNA isoforms. Taking in consideration that both plasmid isoforms are totally retained at 400 mM, new assays were performed at different conditions than the previous ones. Instead of start the assay with 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) to retain both pDNA isoforms, these new assays were initiated with a NaCl concentration higher than 400 mM in order to promote the immediate elution of the oc isoform in the flowthrough, retaining to the column only the sc pDNA isoform (Figure 18). This strategy could be advantageous if the pDNA isoforms are separated in only two chromatographic steps.

Figure 18. (A) Separation of pDNA isoforms with arginine-epoxy monolith. Elution was performed at 1 mL/min by stepwise gradient of 600 mM and 2M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), as represented by the dashed line. UV detection at 260 nm. Injection volume was 200 µL. (B) Agarose gel electrophoresis analysis of each peak represented in the chromatogram. Lane M: molecular weight marker; Lane A: pDNA sample injected onto the column (oc + sc); lane 1: oc mixed with sc; lane 2: sc.

Thus, arginine-epoxy monolithic column was equilibrated with 600 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0). The pDNA sample resulting from alkaline lysis with the Qiagen kit was applied onto the column in the same buffer conditions. After plasmid injection, a first peak was eluted at 600 mM NaCl followed by a second peak when the ionic strength was increased to 2M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), as it is presented in Figure 18 (A). The agarose gel electrophoresis was performed to identify the species eluted in each peak of the chromatogram. The results from agarose gel electrophoresis showed that the oc pDNA was eluted in the first peak (Figure 18 (B), lane 1) and the sc pDNA was distributed in both first and second peaks (Figure 18 (B), lane 2).

On the possibility of differential interactions between oc and sc plasmid isoforms and the arginine ligand and knowing the effect of the NaCl concentration in binding/elution behavior of the both isoforms, new experiments have performed in order to achieve the total separation of plasmid isoforms. Thus, these new assays are based on the same strategy of previous experiments, only decreasing the NaCl concentration. The new chromatographic profile is schematized in the Figure 19.
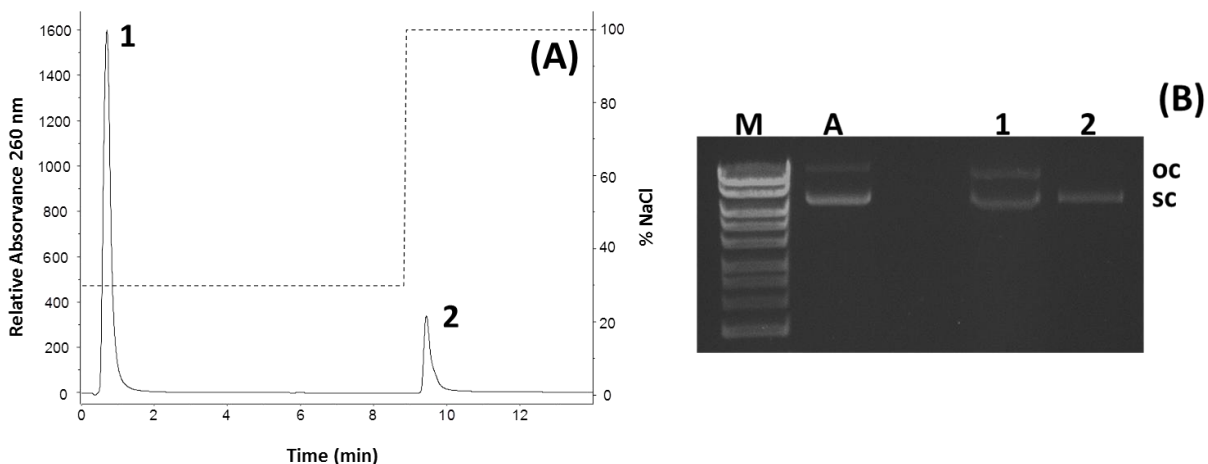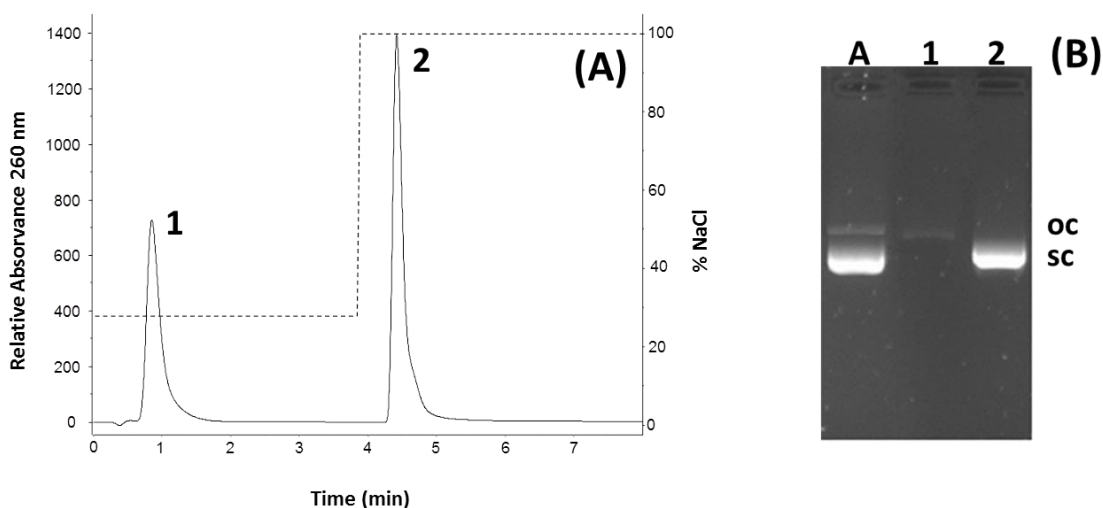
Figure 19. (A) Separation of pDNA isoforms with arginine-epoxy monolith. Elution was performed at 1 mL/min by stepwise gradient of 560 mM and 2 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), as represented by the dashed line. UV detection at 260 nm. Injection volume was 200 µL. (B) Agarose gel electrophoresis analysis of each peak represented in the chromatogram. Lane A: pDNA sample injected onto the column (oc + sc); lane 1: oc; lane 2: sc.

Initially, the monolithic column was equilibrated with 560 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at flow rate of 1 mL/min. After plasmid sample injection (200 µL), a first peak of unbound species was obtained under flowthrough, and then the ionic strength was increased to 2 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) for eluting the bound species in a second peak. The agarose gel electrophoresis was performed to confirm the species present in each peak according to the chromatographic profile (Figure 19 (A)). The results from agarose gel electrophoresis showed that the oc pDNA was achieved in the first peak (lane 1), it being totally separated from the sc pDNA that was attained in the second peak (line 2) (Figure 19 (B)).

The results presented in the Table 8 and Figures 18 and 19 revealed the proneness of arginine ligands to distinguish and differentially interact with both isoforms, further suggesting a specific recognition of the sc pDNA isoform, since it remains bound to the column when the oc form is totally eluted. Thus, the analysis of the sc pDNA interaction to the chromatographic support in a molecular level is required to understand the differences in isoform recognition.

Seeman and coworkers proposed that certain nucleotides could be recognized by particular amino acids side chain and reasoned that the greater specificity of interactions could occur with the major groove (particularly at positions where hydrogen bonding (H-bond) is promoted) rather than the minor groove [94]. Moreover, relative preferences of amino acids for bases are basically dependent on the donor and acceptor patterns of bases, like guanine, which shows the highest degree of participation in H-bonds [95]. Arginine is characterized by the (i) ability to interact in different conformations (single, bidentate and complex); (ii) the

length of its side chain; and (iii) ability to promote good hydrogen bonding geometries that support the specific recognition mechanism [94]. Although Luscombe and coworkers have demonstrated that van der Wall forces and water mediated bonds are present in amino acid – DNA interactions, the arginine specificity could be predominantly due to hydrogen bonds. Indeed, hydrogen bonds are favored between arginine amino groups and nucleic acid bases by bidentate and complex interactions, preferentially with the guanine base [94]. In bidentate interactions two or more hydrogen bounds are made with a base pair [94]. Comparatively to complex interactions the amino acid is capable of interacting with more than one base step simultaneously [94].

Given that sc pDNA isoform has a higher degree of base exposure than oc isoform, resultant from the supercoiling phenomenon that is a consequence of deformations induced by the torsional strain [90], hydrogen bonds may play an important role in the sc pDNA recognition. Thus, the multitude of interactions associated to the arginine amino acid is responsible for anticipated elution of the oc isoform whereas the sc isoform elutes later.

Overall, this purification strategy combines the structural versatilities provided by the epoxy monolithic support with the specificity and selectivity of the arginine ligand to recognize the sc pDNA isoform. Additionally, under the optimized binding/elution conditions, oc plasmid isoform was eluted at 560 mM NaCl while sc pDNA isoform eluted was at 2M NaCl. The high resolution obtained and the low salt concentration required for the sc pDNA purification becomes this technique more appropriate and economic to apply in the biopharmaceutical process to purify the sc HPV-16 E6/E7 DNA vaccine at large scale.

3.3. Dynamic binding capacity

The determination of dynamic binding capacity of a stationary phase is a critical factor of the chromatography performance. Conventional chromatographic supports have been reported as having low binding capacity for large molecules such as plasmid DNA. Effectively, these chromatographic supports have a structure based on small size pore, which are not engineered to handle large molecules [81]. Thus, the binding of plasmid molecules occur only at the external surface of the matrix [74].

In a recent study, Tiainen and coworkers demonstrated that the superporous agarose anion exchangers show four to five times higher dynamic binding capacity for plasmid DNA than the corresponding non-porous agarose particles [96]. Similar values around 2.4 mg/mL were found with tentacular support also for plasmid DNA [97]. On the other hand, continuous chromatographic supports with large pores, such as membrane or monolith-technology, have demonstrated good capacity values [81]. Bencina and coworkers demonstrated that the methacrylate monolith has a plasmid binding capacity of 15.78 mg/mL, whereas Endres and coworkers showed a plasmid capacity of 15 mg/mL achieved by adsorptive membranes [98-100].

Thereby, for a complete characterization of the arginine-epoxy monolith, the dynamic binding capacity was evaluated and was compared with the dynamic binding capacity of the conventional arginine-agarose matrix. Breakthrough experiments were performed with 0.05 mg/mL HPV-16 E6/E7 pDNA solution obtained by alkaline lysis with the Qiagen kit. Each column was equilibrated with 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at 1 mL/min and thereafter, the columns were overloaded with the plasmid solution under the same equilibrium conditions. Determination of dynamic binding capacity was carried out by recording breakthrough curves and calculating the amount of bound pDNA per mL support at 10%; 50% and 100% breakthrough curve. The breakthrough curve profile obtained for both chromatographic supports (conventional agarose support and epoxy-based monolith) is present in Figure 20.
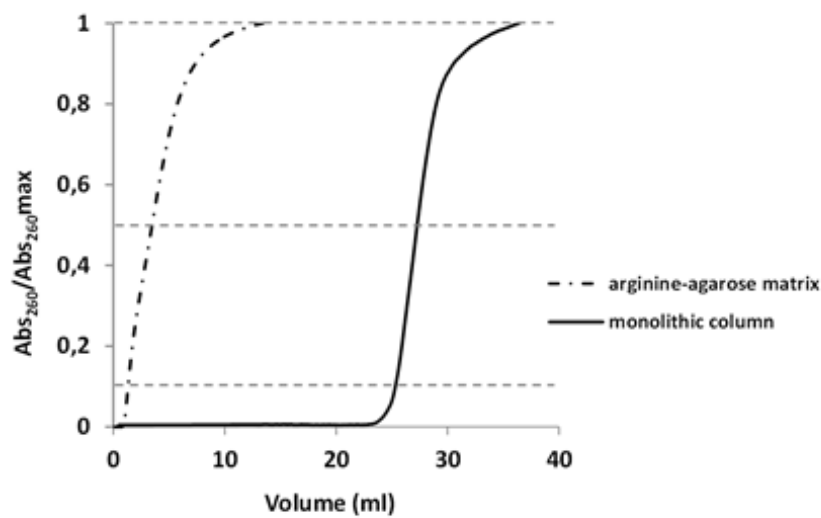
Figure 20. Breakthrough experiments with arginine-epoxy monolith and conventional arginine-agarose matrix. Flow rate: 1 mL/min; feedstock: HPV-16 E6/E7 plasmid solution, at 0.05 mg/mL, was prepared in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0.

The results present in Figure 20 and Table 9 shows a lower capacity for conventional agarose-based matrix compare with an epoxy-based monolith. The lower capacity obtained for conventional agarose-based matrix is the result of the interaction between pDNA molecules with functional groups only at the surface of the bead. These results are in accordance to the results previously reported by Sousa F. and coworkers, where the dynamic binding capacity achieved by the histidine-agarose matrix was 0.2804 mg/mL at same plasmid concentration, 0.05 mg/mL [101] (this value is relatively higher than the value present in Table 9 because the capacity was estimated at 50% breakthrough).

43

Table 9. Comparison of the dynamic binding capacity between arginine-epoxy monolith and conventional arginine-agarose matrix. The breakthrough experiments were performed with 0.05 mg/mL of HPV-16 E6/E7 plasmid solution.

Column type	DBC (mg/ml)		
	10%	50%	Total
Arginine-agarose matrix	0.133	0.357	1.112
Monolithic column	3.5514	3.8382	5.1824

On the other hand, monolithic supports are considered the material of choice for the purification and analysis of pDNA due to their appropriate structure capable of accommodate large pDNA biomolecules. Thus, monoliths offer a very high binding capacity for pDNA when compared with conventional agarose-based matrices [71, 81]. The results presented in Table 9 show that monolith capacity was 3.5514 mg/mL at 10% breakthrough curve for a pDNA concentration value of 0.05 mg/mL. Indeed, the capacity obtained for this arginine-epoxy monolith was twenty six fold higher than the capacity achieved on arginine agarose-based support at 10% breakthrough.

These results also present higher values of the pDNA binding capacity when compared to other results previously reported by Sousa A. and coworkers, where the non-grafted CDI monolith revealed a capacity of 3.640 mg/mL for a pDNA concentration value of 0.05 mg/mL. A possible explanation for the difference between the values of the binding capacity in the both monolithic supports can be associated with the difference of non-grafted and grafted monoliths. Since, the experiment performed with epoxy-based monolithic support without arginine immobilization did not reveal interactions between plasmid DNA sample and matrix. These results show that arginine ligands are essentials for pDNA binding and consequently high capacity value. Thus, these results compared with the results obtained for the non-grafted CDI monolith also suggest that the ligand immobilization could be increased the dynamic binding capacity of monolithic supports [14]. These results are in concordance with other study that showed the improvement on capacity, around 17 mg/mL, when it was used the grafted DEAD weak anion-exchange methacrylate monolith when compared with non-grafted monolith [102]. In this way, the arginine-epoxy monolith prepared in this work can be considered as a good affinity chromatographic support.

3.4. sc pDNA purification from the clarified *E. coli* lysate

In the first study it was proved that the immobilized arginine in the epoxy-based monolith recognized specifically the sc pDNA active conformation. This fact encouraged the investigation to proceed with the study of the arginine-epoxy monolith in the recognition of the sc pDNA isoform from the clarified *E. coli* lysate. From this stand point, the first preliminary experiment was performed in the same conditions previously used where isoforms separation was achieved. The aim of this first screen was to verify the interaction behavior

between both lysate sample, namely plasmid isoforms, and arginine ligands. Furthermore, an additional study in the pDNA isoform separation was performed by using 1 M NaCl in the second elution step, it being observed that the sc pDNA was totally eluted under this salt concentration. Thus, the arginine-epoxy monolithic column was first equilibrated with 560 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at room temperature and 1mL/min. After *E. coli* lysate injection, the ionic strength was increased to 1M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0). The correspondent chromatographic profile from this assay is presented in Figure 21.

Figure 21. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 560 mM to 1M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 µL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chromatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1 and 2 are shown in lanes 1 and 2 respectively.

Analyzing the chromatographic profile obtained under the described conditions, the existence of two resolved peaks were eluted at 560 mM (peak 1) and 1 M (peak 2) is evident, as demonstrated in the Figure 21 (A). In order to identify the eluted species, peak fractions were recovered during elution and further analyzed by agarose gel electrophoresis. The Figure 21 (B) revealed that the first peak corresponds to oc and sc pDNA isoforms and RNA, while in the second peak were present the sc pDNA isoform and RNA molecules.

The chromatographic profile obtained under the described conditions showed that nucleic acids present in clarified *E. coli* lysate have different binding/elution behavior, which also affected the oc and sc isoforms separation. Thus, the attempt to get the sc pDNA isolation from the clarified E. coli lysate with the same strategy is not feasible and the elution gradient should be optimized. Once that lower ionic strength promotes strong interactions between nucleic acids and arginine ligands, an additional step with 10 mM Tris-HCl and 10 mM EDTA

buffer (pH 8.0) for the equilibrium stage can improve the interactions and selectivity between the several nucleic acids present in the *E. coli* lysate, namely the sc pDNA. A chromatographic run was performed with the addition of this new equilibrium step, as it is presented in Figure 22 (A).
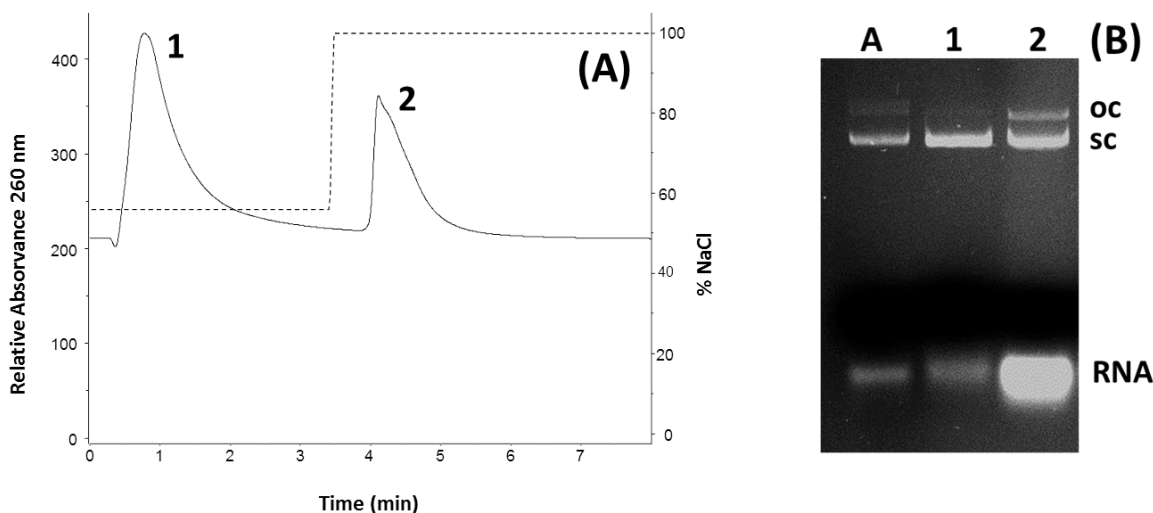

Figure 22. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 0 M to 560 mM and finally to 1 M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 µL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chromatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1 and 2 are shown in lanes 1 and 2 respectively.

By analysis of the chromatographic profile obtained, after injection of the pDNA-containing clarified lysate, the total sample retention was observed in the presence Tris-HCl buffer. However, two peaks were also eluted for the same steps previously used. In order to establish a correlation between peaks in chromatographic profile and several species present in clarified *E. coli* lysate an agarose gel electrophoresis was performed. The results showed that the first peak corresponds only to the RNA, whereas the second peak shows three bands corresponding to the oc and sc pDNA isoforms and also RNA (Figure 22 (B)).

Curiously, the first eluted specie was RNA, although a small fraction of RNA was also eluted in the second peak, and the elution behavior of pDNA isoforms was changed because both isoforms were retained at 560 mM NaCl concentration. These results suggest that the RNA molecules have weak interactions with arginine ligands and probably a small increase of NaCl concentration in the second step can allow the total RNA elution in the first peak. In fact, RNA molecules are negatively charged and single-stranded but also are characterized as low molecular weight biomolecules [100]. Thus, a possible explanation for the RNA elution in the first peak is the smaller size of these molecules comparing to plasmid DNA isoforms.

Further experiments were performed to optimize the salt concentration in order to eliminate the *E. coli* RNA in the first peak and thereafter separate the pDNA isoforms. These tests were conducted as described in above screening experiments, increasing the NaCl concentration of the second chromatographic step. The chromatographic profile obtained in these circumstances is presented in Figure 23 (A).
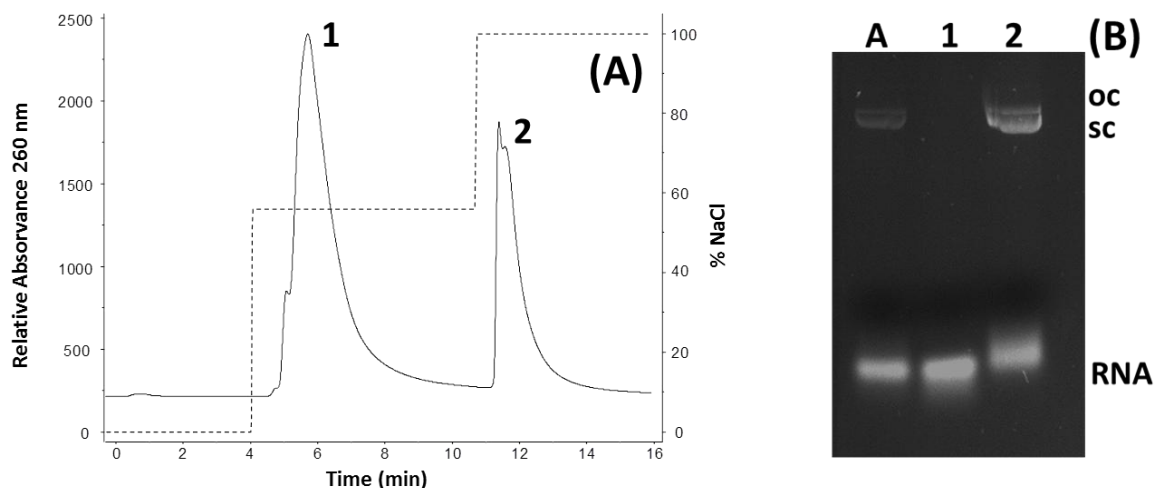

Figure 23. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 0M to 700mM and finally to 1M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 µL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chramatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1 and 2 are shown in lanes 1 and 2 respectively.

The resuts in Figure 23 (A) showed that the first chromatographic peak obtained at 700 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) was significantly increased, as it was expected. The agarose gel electrophoresis analysis showed that the first peak only contains RNA, and the second peak contains plasmid isoforms and a small fraction of RNA. These results revealed that 700 mM NaCl in the second chromatographic step was not enough to isolate the *E. coli* RNA from pDNA isoforms. Therefore, following previous experiments, a further increase to 730 mM in the NaCl concentration was performed, as it is depicted in Figure 24 (A).
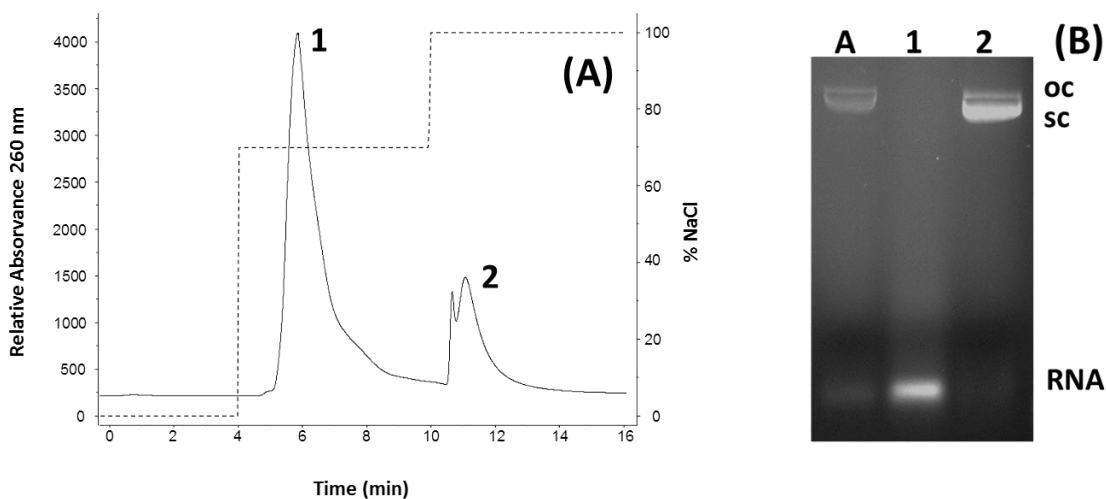
Figure 24. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 0 M to 730 mM and finally to 1 M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 µL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chromatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1 and 2 are shown in lanes 1 and 2 respectively.

Two peaks were obtained in the chromatographic assay (Figure 24 (A)) and the agarose gel electrophoresis analysis (Figure 24 (B)) showed that the oc pDNA was presented in second peak and sc pDNA was presented in both first and second peaks. The RNA molecule was mostly eluted in the first peak, however a small fraction continued eluting in the second peak.

The results obtained for the purification of the sc pDNA from the clarified *E. coli* lysate have revealed the involvement of electrostatic interactions between other specific interactions with the several nucleic acids present in the sample. The electrostatics interactions between nucleic acid phosphate groups and the arginine amino acid positively charged may confer stability to the biomolecule retention [80]. In fact, molecular modeling studies have demonstrated that although Van der Waals forces and water mediated bonds are present in amino acid – DNA interactions, H-bonds are the predominant forces that dictate their contact [80]. Luscombe and coworkers also described favored interactions between arginine amino acid and guanine bases from nucleic acids, predominantly due to bidentate and complex interactions via hydrogen bonding [80]. This biorecognition from the nucleic acid bases can be responsible for the ability of arginine matrix to specifically recognize the different nucleic acids present in clarified *E. coli* lysate.

According to the result present in the Figure 24, a further increase in NaCl concentration of the first peak is not sufficient to promote the sc pDNA purification. Thus, new

48

chromatography runs with an additional step was performed in order to promote the isolation of the sc pDNA isoform from the oc pDNA and the RNA. The chromatographic profile of this new experiment is present in Figure 25 (A). After injection of the pDNA-containing clarified lysate sample, three resolved peaks were obtained.
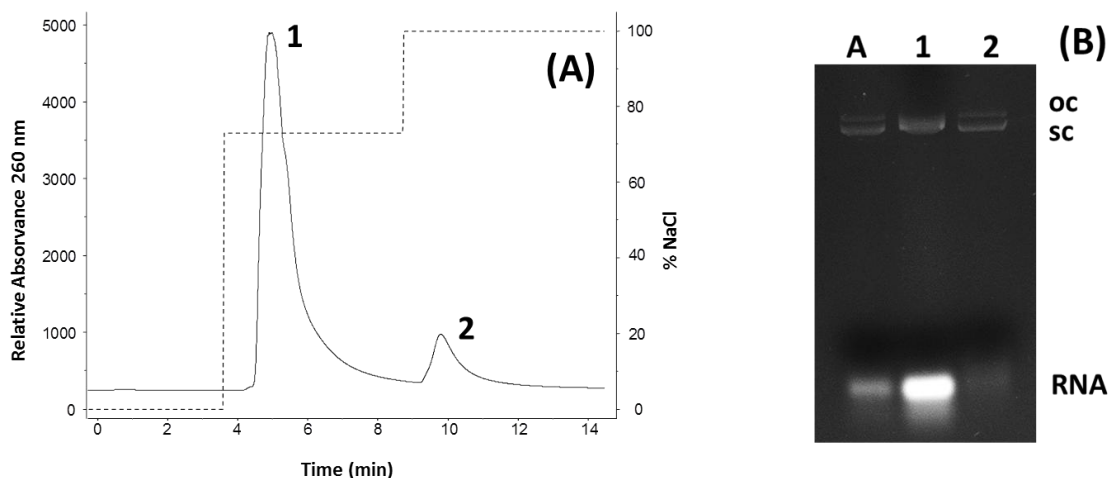

Figure 25. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 0 M to 690 mM, following to 770 mM and finally to 1 M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 μL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chromatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1, 2 and 3 are shown in lanes 1, 2 and 3 respectively.

The chromatographic run was initiated at 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0) at room temperature and a flow rate of 1mL/min for the equilibrium of the column. The first peak was achieved by increasing the ionic strength of the buffer to 690 mM NaCl. After, the ionic strength of elution buffer was increased to 770 mM of NaCl in order to partially elute specific molecules in a second peak. The chromatographic run was concluded through a final elution step using 1 M of NaCl to elute the remaining bound nucleic acids. The electrophoretic analysis showed that RNA molecules were majority eluted in the first peak, although a small fraction of them was also eluted at other salt concentration (second and third peaks). Relatively to the oc and sc pDNA isoforms, both were achieved in the second and third elution peaks.

This result can be associated with the binding step with Tris-HCl buffer. According to Sousa F. and coworkers, the arginine ligand is very effective in binding of plasmid isoforms at low ionic strength, confirming the stronger interactions of the pDNA isoforms by the arginine ligands [103]. In fact, electrostatic interactions are involved between these phosphate groups and the arginine ligands [95]. However, these electrostatic interactions may confer stability to biomolecule retention, but not specificity. Nevertheless, electrostatic interactions do not

totally explain the selectivity between oc and sc pDNA isoforms. A possible explanation may arise from the theory that van der Waals energy fluctuations as well as twisting fluctuations create intercrossed or semi-relaxed intermediate and interact with the affinity support in different modes [104].

In this way, a new strategy without the equilibrium step at 10 mM of Tris buffer was applied in order to improve the sc plasmid isolation from the others species present in the clarified *E. coli* lysate. These experiments were carried out by an equilibrium step at 600 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0), at room temperature, and a flow rate of 1mL/min. Under these conditions, a first peak was eluted in the flowthrough containing unbound species. After, the ionic strength was increased to 740 mM NaCl, which resulted in a second peak. And finally, the biomolecules that were retained at previous ionic strength were only recovered by increasing the NaCl concentration in the buffer to 1 M (Figure 26 (A)).
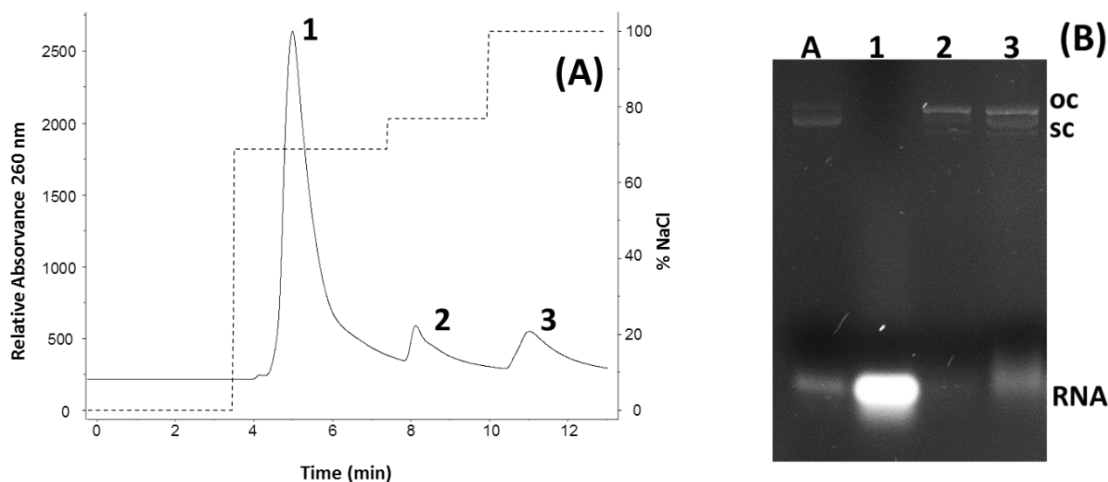

Figure 26. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 600 mM to 740 mM and finally to 1 M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 µL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chromatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1, 2 and 3 are shown in lanes 1, 2 and 3 respectively.

Under the conditions used, three peaks were eluted at 600 mM NaCl (peak 1), 740 mM NaCl (peak 2) and 1M NaCl (peak 3). The fractions of each peak were pooled and analyzed by agarose gel electrophoresis in order to establish the correlation between several nucleic acids present in clarified *E. coli* lysate and the chromatographic peaks. Figure 27 (A) confirms that the first and second peaks mostly correspond to the RNA molecules that did not interact with the support, while the sc plasmid isoform was identified in the third peak. Despite the third peak correspond mostly to the sc plasmid isoform, a small fraction of the RNA molecules is also present.

In fact, this chromatographic assay showed that the sc pDNA isoform was almost isolated. However, a strategy able to eliminate the small fraction of RNA that is contaminating the sample is necessary.

RNA molecules were differentially eluted with several ionic strength conditions, by suggesting the presence of several RNA species in the clarified *E. coli* lysate. Because RNA molecules are single-stranded nucleic acids, the elution behavior on arginine-epoxy monolith could be associated to the base exposure degree [13]. Bioinformatics predictions refer that messenger RNA (mRNA) and transfer RNA bases are more available to interact with amino acids because they exhibit an less compact secondary structure; while ribosomal RNA (rRNA) forms more compact secondary structures, exhibiting a lower base exposure degree [105]. Therefore, the present results can be in accordance to these predictions.

According to this fact, the last chromatographic run was performed with and additional step in order to eliminate the vestigial RNA present in the peak of the sc pDNA isoform. First, the arginine-epoxy monolithic column was equilibrated with 600 mM NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer (pH 8.0). After injection of pDNA-containing clarified lysate in same salt concentration of the equilibrium buffer, a first peak was obtained resulting from the elution of the unbounded species. Next, the ionic strength of elution buffer was increased to 740 mM NaCl, following to 795 mM NaCl in order to partially elute different species in the second and third peaks, respectively. The chromatographic run was concluded through a final elution step at 1M NaCl to obtain the strongly bound species. The resultant chromatographic profile is present in the Figure 27 (A).
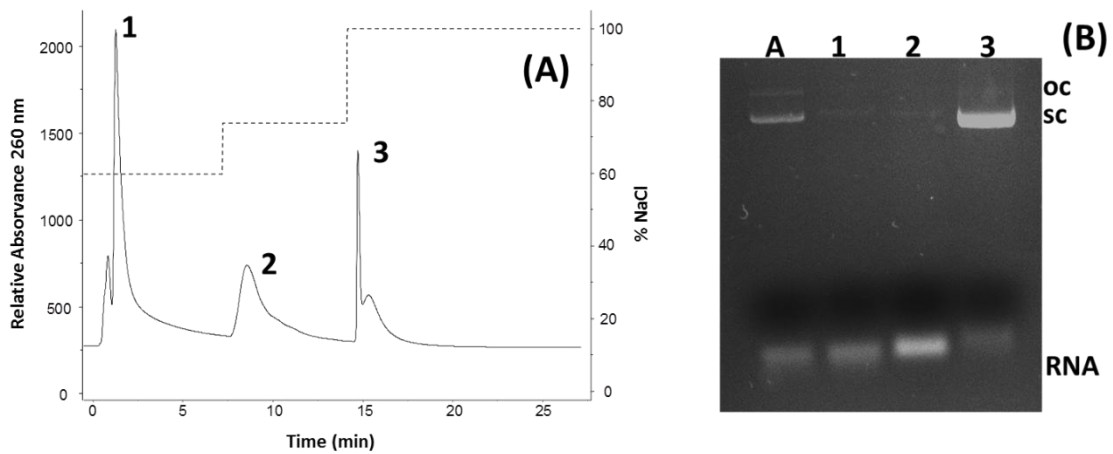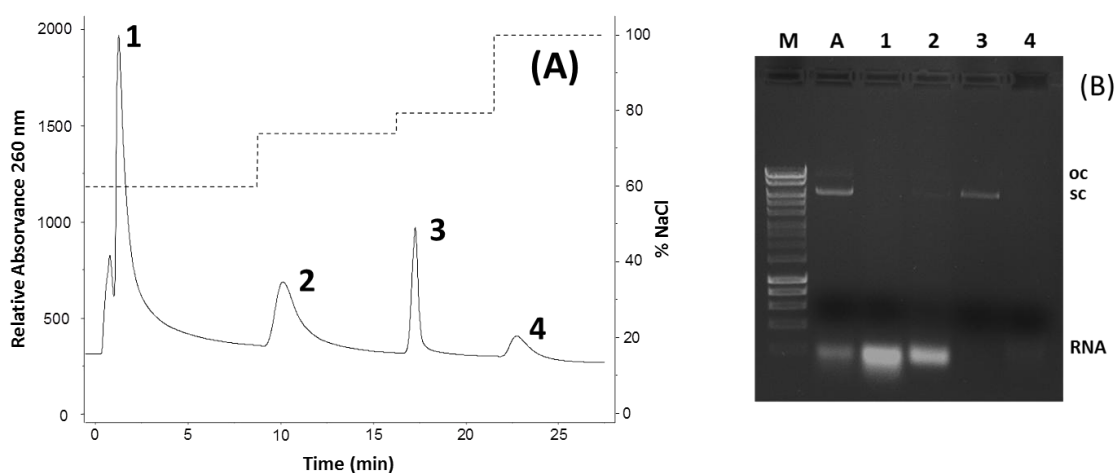

Figure 27. (A) Chromatographic profile of the *E. coli* lysate sample injection in the arginine-epoxy monolith. Mobile phase – buffer A: 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0; buffer B: 1 M NaCl in 10 mM Tris-HCl and 10 mM EDTA buffer pH 8.0. Elution was performed at 1 mL/min by stepwise gradient increasing NaCl concentration in the eluent from 600 mM to 740 mM, following to 795 mM and finally to 1 M, as represented by the dashed line. UV detection at 260 nm. Injection volume was 500 µL. (B) Agarose gel electrophoresis analysis of samples collected from each peak of the chromatographic assay. Lane A: feed sample injected in the column; Fractions corresponding to peaks 1, 2, 3 and 4 are shown in lanes 1, 2, 3 and 4 respectively.

The presence of several peaks in the chromatogram suggests that the arginine-epoxy monolith interacts differently with sc and oc plasmid isoforms and also with several RNA species present in a clarified *E. coli* lysate. In order to identify the species eluted in each chromatographic peak, an agarose gel electrophoresis was performed. The results showed that the elution of oc pDNA isoform occurs in the second peak (Figure 27 (B), lane 2) whereas the elution of sc pDNA isoform occurs in third peak (Figure 27 (B), lane 3) when the ionic strength was increased. On the other hand, RNA species were eluted in the first, second and fourth peaks (Figure 27 (B), lane 1, 2 and 4 respectively). These results confirm that arginine-epoxy monolith has the ability for specifically recognize and isolate the sc pDNA isoform from the clarified *E. coli* lysate.

As shown in the Figure 27, nucleic acids are isolated after 25 minutes. This result suggests that the chromatographic run in the modified-epoxy monolith is accomplished in less time than using the conventional arginine-agarose matrix for the same goal, since in the last case may take 5 fold more [13]. One of the advantages present by the monolithic supports consists in fast separations, avoiding the prolonged contact time of the sample with the support, which is in accordance to the results obtained [81].

In accordance to these results, the CIM^{TM} arginine-epoxy disk has demonstrated to be a tool with the ability to develop fast and efficient separation processes using low back pressure, with high capacity and maintaining the stability and productivity for plasmids. On the other hand, the affinity chromatography based on arginine as specific ligand arises as a good strategy for to obtain the specific biorecognition for the sc plasmid isoform. Thus the combination of structural versatilities provided by the monolithic supports with the specificity and selectivity of amino acid ligands rises as a promising strategy in the sc pDNA purification for future therapeutic applications.

3.5. Analytical parameters for plasmid quality assessment

The sc pDNA resultant from the purification with arginine-epoxy monolith should be free from host impurities such as proteins, RNA, genomic DNA and endotoxins as well as from other plasmid conformations (oc and ln). In this way, the regulatory agencies like FDA recommend that the sample homogeneity must be higher than 97% of sc isoform, proteins and RNA must be undetectable by micro-BCA method and 0.8% agarose gel respectively, the gDNA and endotoxin levels must be lower than 0.01 µg gDNA/µg pDNA by real-time analysis and 0.1 EU/µg pDNA by LAL assay respectively [68, 106]. Thus, a strict analysis to the samples resultant from the sc pDNA isolation and purification processes should be made in order to determine the impurity level in each sample (gDNA, RNA, endotoxins and proteins).

Firstly, the quantification and purity of sc pDNA sample purified by the arginine-epoxy monolithic support were performed by high-performance hydrophobic interaction

chromatography (HPLC) according to the method previously developed [86]. Briefly, this method determines pDNA concentration in crude *E. coli* extracts and other samples resultant from the sc pDNA isolation and purification processes. The purity degree, purification factor and yield (Table 10) of sc pDNA sample obtained with the arginine-epoxy monolith can also be evaluated [86].

The HPLC analysis of the crude *E. coli* lysate revealed a multi-peak chromatogram corresponding to different biomolecules. The large quantity and diversity of biomolecules present in this stage was due to the plasmid-containing extract has not been undergo any purification process. The analytical chromatogram shown in Figure 28 (A) revealed one characteristic peak at 0.7 min of the total pDNA [86], followed by several peaks corresponding to different impurities. Comparing this chromatographic profile with the obtained profile in the recovered sample from the concentration and clarification steps, a significant impurity reduction was verified (Figure 28 (B) and (C) respectively). This considerable reduction in the impurity content was due to the isopropanol concentration and ammonium sulfate precipitation steps, performed prior to arginine affinity chromatography [107]. The analysis of sample collected in the third step of the monolithic purification strategy, also revealed a characteristic peak at 0.7 min referent to pDNA (Figure 28 (D)). Besides this peak, the analytical chromatogram showed other peak which supposes the presence of impurities. However, the analytical chromatogram obtained with the injection of binding buffer as a negative control, clearly showed the presence of one peak at the same time. Thus, the interference peak present after the elution of the sc pDNA is due to the column phenomena that also occur without sample injection, as it shown in Figure 28 (E).
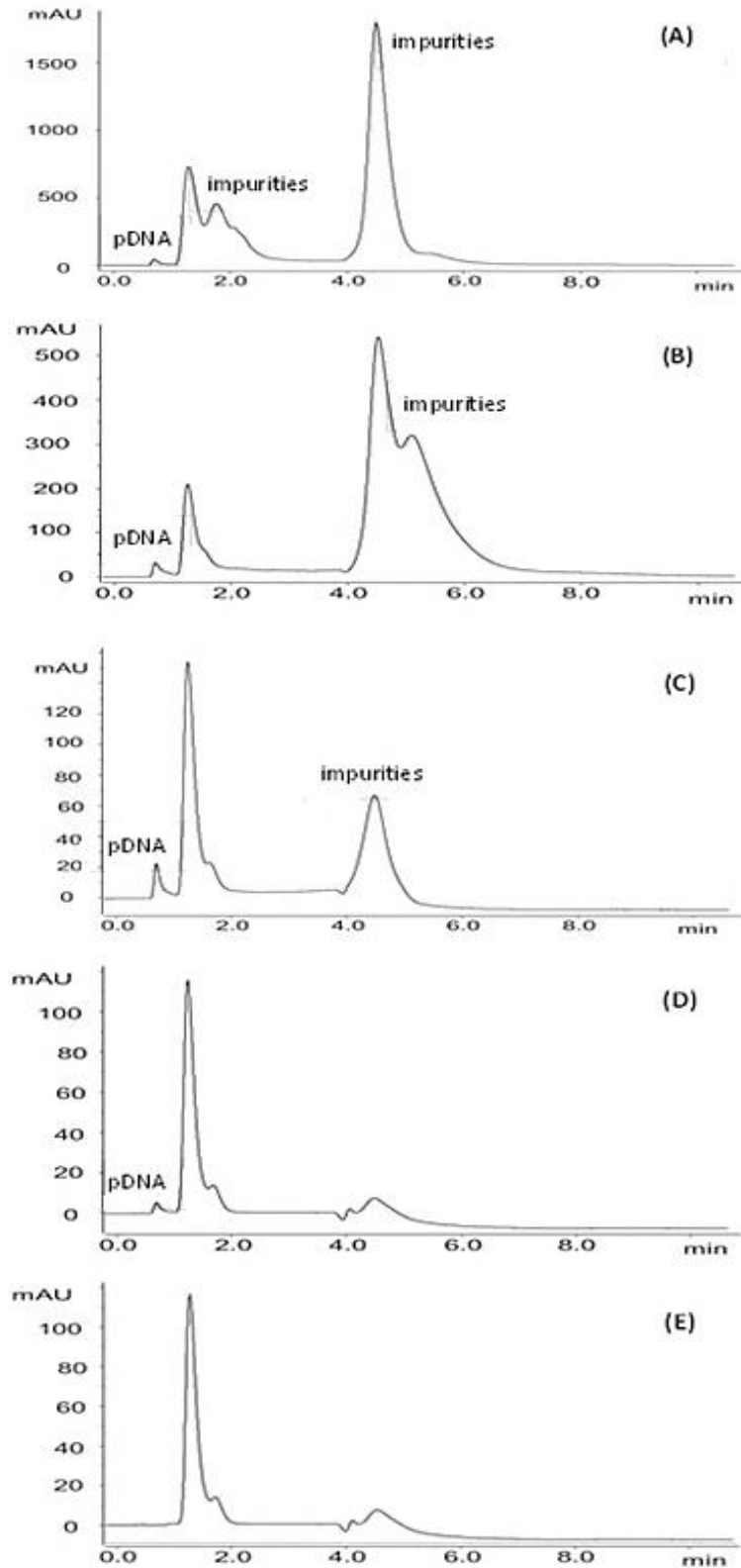
Figure 28. Analytical chromatographic profiles of different pDNA-containing samples recovered throughout the purification process: (A) *E. coli* lysate; (B) plasmid-containing sample after isopropanol concentration; (C) plasmid-containing sample after ammonium sulfate clarification (feed sample injected onto the column); (D) sc isoform fraction and (E) zero (injection of binding buffer).

To quantify the pDNA purity degree of each analyzed sample, the ratio between plasmid peak area and total peak areas of the chromatogram was measured. Thus, the plasmid purity for the feed sample injected in the monolithic column was about 10% (Table 10). The purity degree of the sc plasmid sample attained in third peak of monolithic chromatography was around 100% (Table 10). These results are in agreement with previous electrophoretic analysis and confirm that the monolith modification with arginine ligands can be useful in a sc DNA purification process and effective in impurities reduction.

Table 10. HPLC analysis of concentration, purity and recovery yield of sc HPV-16 E6/E7 pDNA isolated by arginine-epoxy monolith.

Process step	Volume	pDNA		HPLC purity	Purification factor	Step yield
	(mL)	(µg/mL)	(µg)	(%)		(%)
Feed sample	6	42.57	255.42	9.60	-	-
sc-peak 3 (12 runs)	5.88	17.02	100.08	100	10.42	39.18

Moreover, the yield of purification step can be determined by the ratio between the mass of plasmid present in the recovered sample and mass of plasmid present in the feed sample. As showed in the Table 11, from 255.42 µg of pDNA contained in the feed sample, only 100.08 µg of the sc plasmid was recovered and totally isolated, corresponding to the yield of 39% (Table 11). This result indicated a loss of pDNA in the purification process with the arginine-epoxy monolith.

In order to clarify if this loss is related to the monolith modification or to the HPV-16 E6/E7 therapeutic plasmid, an additional study was performed to quantify the concentration, purity and yield of sc pVAX1-*LacZ* plasmid from the *E. coli* lysate isolated also with the arginine-epoxy monolith. Consequently, the feed sample that was injected in the monolithic column contained 1129.44 µg of pVAX1-*LacZ*, whereas the total amount of plasmid recovered after the purification step was 954.93 µg (oc and sc isoforms), corresponding to 86% yield. Therefore, this yield is significantly higher than other studies based on affinity chromatography, such as 62% yield achieved in triplex-affinity chromatography [108, 109] and 45% sc pDNA yield achieved with the conventional histidine-affinity matrix [12]. Comparatively with the recent application of non-immobilized CDI monolith, which achieved 74% yield of total plasmid [14] and the anion exchange monolithic disk with 75% yield [110], the value obtained by arginine-epoxy monolithic support (86%) was satisfactory. However, the 39% sc HPV-16 E6/E7 plasmid yield achieved with arginine-epoxy monolith in the present work is not a consequence of the modified support, but it is probably due to HPV-16 E6/E7 plasmid characteristics, such as the higher molecular size or the nucleotide bases constitution that

can lead to a stronger retention to the monolith resulting in its loss during the purification step.

The presence of proteins in a final sc plasmid sample used in a therapeutic study can lead to an allergic response such as anaphylactic shock, or chronic, such as autoimmune disease [111]. On the other hand, the contamination by proteins can take to biological reactions due to production of antibodies, hormones and/or cytokines [106, 111]. Thus, regulatory agencies established that proteins should not be detected in the purified sc pDNA sample. The protein quantification was measured using the micro-BCA test. According to the data presented in the Table 11, residual concentrations of 3.51 mg/mL, 0.71 mg/mL and 0.03 mg/mL were determined on the alkaline lysis, isopropanol precipitation and ammonium sulfate precipitation fractions respectively. These results suggest a significant reduction of proteins during plasmid isolation process, namely in isopropanol concentration and ammonium sulfate precipitation steps. However, the feed sample injected in the modified monolithic column revealed a small amount of proteins that were eliminated during the purification process (Table 11).

Table 11. Protein, gDNA and endotoxins quantification in several steps of the isolation and purification processes to obtain the sc pDNA from the clarified *E. coli* lysate.

Sample	Protein	Endotoxin		Genomic DNA	
	(mg/mL)	(EU/mL)	(EU/µg of pDNA)	(ng/mL)	(ng/µg of pDNA)
Primary isolation					
Alkaline lysis	3.51	111.52	2.42	27800	602.28
Isopropanol precipitation	0.71	477.28	1.16	200900	477.99
$(NH_4)_2SO_4$ precipitation	0.03	94.96	0.64	29100	194.73
Purification step					
RNA-peak1	Undetectable	-	-	-	-
RNA-peak2	Undetectable	-	-	-	-
sc-peak3	Undetectable	0.86	0.05	140	8.23
RNA-peak4	Undetectable	-	-	-	-

A common problem with the use of the *E. coli* microorganism as a host strain is the release of endotoxins into the medium upon processing, which are difficult to remove during the subsequent purification steps [112]. Endotoxins are highly negatively charged lipopolysaccharide (LPS) and a major component of the Gram-negative bacterial cell wall such as *E. coli* [113]. These molecules are an extremely potent stimulator of the mammalian

systems, including nonspecific activation of the immune system, stimulation of cytokine overproduction and induction of the characteristic shock syndrome associated with bacterial infections [113, 114]. High endotoxin levels can affect the eukaryotic cell proliferation and transfection efficiency [114]. In this way, the correct elimination of endotoxins is extremely important in order to avoid the undesirable effects during a therapeutic application of the sc HPV-16 E6/E7 pDNA sample [115]. The results of Table 11 show that the endotoxin content in the final of alkaline lysis step was 2.42 EU/µg of pDNA, decreasing to 0.64 EU/µg of pDNA in the feed sample probably due to the isopropanol and ammonium sulfate precipitation. Nevertheless, the endotoxin content in feed sample was considerably reduced in the chromatography step, since the final sc pDNA fraction had 0.05 EU/µg of pDNA, which respects the guidelines of regulatory agencies (Table 11).

Finally, the presence of genomic DNA in a final sc plasmid sample can lead to the oncogene appearance that are encoded by gDNA fragments, and their possible activation or deactivation once in recipient cells, can lead to the tumor formation [106]. The gDNA quantification was performed by real-time PCR. Such as for the other impurities, the reduction of gDNA was more pronounced in the alkaline lysis process, from 602.28 ng/µg of pDNA in the initial lysate to 194.73 ng/µg of pDNA in the feed sample (Table 11). The efficiency of arginine-epoxy monolithic support to eliminate gDNA was also significant because the residual amount of gDNA in the sc plasmid sample was 8.23 ng/µg of pDNA (Table 11). This value is lower than the 0.01 µg/µg of pDNA recommended by regulatory agencies [68, 106].

Chapter IV

Conclusions and futures perspectives

Nowadays, millions of people die each year from incurable diseases, because often patients are resistant to drugs. This situation leads to a great need for new stable vaccines. The use of DNA vaccines based on plasmid DNA for the treatment of inherited and acquired diseases offer a safer alternative for conventional vaccines. In this way, efficient purification processes should be explored to obtain the sc pDNA with the purity degree required for pharmaceutical applications. The implementation of some amino acids as immobilized ligands has been producing interesting results with respect to plasmids purification. Hence, the purpose of this thesis was to develop and implement novel affinity chromatography approach for sc plasmid purification from clarified *E. coli* lysate. This strategy is based on the combination of structural versatilities provided by the epoxy-monolithic supports with the specificity and selectivity of arginine immobilized ligands. In addition, this thesis described the analytical characterization of sc pDNA sample recovered from the arginine-epoxy monolith that should respect requirements from regulatory agencies.

Firstly, the epoxy monolith was modified by arginine ligand immobilization and was tested with a pDNA HPV-16 E6/E7 sample pre-purified with a commercial kit. Comparing to resultant chromatographic assays of the modified monolith with the non-modified monolith, it was concluded that the epoxy monolith was correctly modified, since the pDNA only interacted with the monolith that had arginine ligands.

For a complete characterization study of the arginine-epoxy monolith, the dynamic binding capacity was evaluated and compared with the dynamic binding capacity of the conventional arginine-agarose matrix. This characterization revealed that the monolith binding capacity at 10% breakthrough was 3.5514 mg/mL, whereas for the conventional arginine-agarose support was 0.133 mg/mL for a plasmid concentration of 0.05 mg/mL. Therefore, the binding capacity of modified monolith was twenty six fold higher than the conventional support.

Additionally,the sc plasmid isoform purification was achieved by optimizing the ionic chromatographic conditions that promoted the binding/elution of the plasmid isoforms in the arginine-epoxy monolith. This study contributes to the development of new affinity chromatography strategy which specifically recognizes the sc pDNA isoform.

Therefore, the application of the arginine-epoxy monolith for the sc pDNA isolation from the other nucleic acids present in a clarified *E. coli* lysate was proposed. Moreover, a strategy for sc plasmid purification from clarified *E. coli* lysate with arginine-epoxy monolith was developed, by using a stepwise gradient increasing NaCl concentration in the eluent (600 mM, 740 mM, 795 mM and 1M). Furthermore, the purification was achieved in 25 minutes which is a very fast separation comparatively with the 120 minutes obtained with conventional arginine-agarose support, ensuring structural stability of the target molecule.

The quality and purity of the sc plasmid resultant from the purification step with the arginine-epoxy monolith was evaluated by using the methodologies required by regulatory agencies such as FDA. Quality control tests indicated that the plasmid sample presented a purity degree approximately 100% and an homogeneity higher than 97% of sc isoform, with an extremely reduced level of impurities (RNA, proteins, genomic DNA and endotoxins).

The efficient purification of the pDNA HPV-16 E6/E7 provides a promising strategy for the development of the DNA vaccines against Human Papillomavirus. Further *in vitro* studies would be crucial to assess the *in vitro* expression of E6 and E7 proteins encoded in the plasmid. In this way, the dendritic cells can be a good choice for these assays because they are good antigen presenting cells and plays a key role in generation of antigen-specific antiviral and antitumor T-cells immune responses.

In conclusion, this modified monolithic support represents an advantageous alternative to conventional supports for sc pDNA purification due to fast separation and consequent short contact time, favoring the structural stability of the target molecule, showing good capacity and selectivity to separate the nucleic acids present in clarified *E. coli* lysate and efficiently isolate the sc plasmid isoform.

Chapter V

Bibliography

1.   Monie, A., Tsen, S. D., Hung, C. and Wu, T. *Therapeutic HPV DNA vaccines*. Expert review of vaccines, 2009. **8**(9): p. 1221-1235.

2.   Hung, C., Monie, A., Alvarez, R. D.and Wu, T. *DNA vaccines for cervical cancer: from bench to bedside*. Experimental & molecular medicine, 2007. **39**(6): p. 679.

3.   Lehtinen, M. and Dillner, J. *Clinical trials of human papillomavirus vaccines and beyond*. Nature Reviews Clinical Oncology, 2013.

4.   Hung, C., Ma, B., Monie, A., Tsen, S. and Wu, T. *Therapeutic human papillomavirus vaccines: current clinical trials and future directions*. Expert Opinion on Biological Therapy, 2008. **8**(4): p. 421–439.

5.   Akbari, O., Panjwani, N., Garcia, S., Tascon, R., Lowrie, D., *et al. DNA vaccination: transfection and activation of dendritic cells as key events for immunity*. The Journal of experimental medicine, 1999. **189**(1): p. 169-178.

6.   Donnelly, J. J., Wahren, B. and Liu, M. A. *DNA vaccines: progress and challenges*. The Journal of Immunology, 2005. **175**(2): p. 633-639.

7.   Davis, H. L., Michel, M. and Whalen, R. G. *Use of plasmid DNA for direct gene transfer and immunization*. Annals of the New York Academy of Sciences, 1995. **772**(1): p. 21-29.

8.   Luo, D. and Saltzman, W. M. *Synthetic DNA delivery systems*. Nature biotechnology, 2000. **18**(1): p. 33-37.

9.   Stadler, J., Lemmens, R. and Nyhammar, T. *Plasmid DNA purification*. The journal of gene medicine, 2004. **6**(1): p. S54-S66.

10.   Sousa, F., Prazeres, D. M. F. and Queiroz, J. A. *Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA*. Trends in biotechnology, 2008. **26**(9): p. 518-525.

11.   Sousa, A., Sousa, F. and Queiroz, J. A. *Impact of lysine-affinity chromatography on supercoiled plasmid DNA purification*. Journal of Chromatography B, 2011. **879**(30): p. 3507-3515.

12.   Sousa, F., Freitas, S., Azzoni, A. R., Prazeres, D. M. F. and Queiroz, J. A. *Selective purification of supercoiled plasmid DNA from clarified cell lysates with a single histidine-agarose chromatography step*. Biotechnology and applied biochemistry, 2006. **45**(3): p. 131-140.

13.   Sousa, F., Prazeres, D. M. F. and Queiroz, J. A. *Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography*. The journal of gene medicine, 2009. **11**(1): p. 79-88.

14.   Sousa, A., Bicho, D., Tomaz, C. T., Sousa, F. and Queiroz, J. A. *Performance of a non-grafted monolithic support for purification of supercoiled plasmid DNA*. Journal of Chromatography A, 2011. **1218**(13): p. 1701-1706.

15.   Yoshikawa, H., Ebihara, K., Tanaka, W. and Noda, K. *Efficacy of quadrivalent human papillomavirus (types 6, 11, 16 and 18) vaccine (GARDASIL) in Japanese women aged 18-26 years.* Cancer science, 2013. **104**(4): p.465-472.

16.   Wu, C., Monie, A., Pang, X., Hung, C. and Wu, T. *Improving therapeutic HPV peptide-based vaccine potency by enhancing CD4+ T help and dendritic cell activation.* J Biomedical Science, 2010. **17**(1): p. 88.

17.   Álvarez-Argüelles, M. E., Melón, S., Junquera, M. L., Boga, J. A., Villa, L., *et al. Human Papillomavirus Infection in a Male Population Attending a Sexually Transmitted Infection Service.* PloS one, 2013. **8**(1): p. e54375.

18.   Sanclemente, G. and Gill, D. *Human papillomavirus molecular biology and pathogenesis.* Journal of the European Academy of Dermatology and Venereology, 2002. **16**(3): p. 231-240.

19.   Chan, P. K. S., Picconi, M. A., Cheung, T. H., Giovannelli, L. and Park, J. S. *Laboratory and clinical aspects of human papillomavirus testing.* Critical Reviews in Clinical Laboratory Sciences, 2012. **49**(4): p. 117-136.

20.   Doorbar, J. *Molecular biology of human papillomavirus infection and cervical cancer.* Clinical Science, 2006. **110**: p. 525-541.

21.   Pang, C. L. and Thierry, F. *Human papillomavirus proteins as prospective therapeutic targets.* Microbial pathogenesis, 2012. **58**: p. 55-65.

22.   De Villiers, E., Fauquet, A., Broker, T. R., Bernard, H. and Hausen, H. Z. *Classification of papillomaviruses.* Virology, 2004. **324**(1): p. 17-27.

23.   Malagón, T., Drolet, M., Boily, M., Franco, E. L.,Jit, M., *et al. Cross-protective efficacy of two human papillomavirus vaccines: a systematic review and meta-analysis.* The Lancet Infectious Diseases, 2012.**12**(10): p.781-789.

24.   Kim, H. J., Lim, S. J., Kwag, H. and Kim, H. *The choice of resin-bound ligand affects the structure and immunogenicity of column-purified human papillomavirus type 16 virus-like particles.* PloS one, 2012. **7**(4): p. e35893.

25.   Münger, K. *The molecular biology of cervical cancer.* Journal of Cellular Biochemistry, 1995. **59**(23): p. 55-60.

26.   McKaig, R. G., Baric, R. S. and Olshan, A. F. *Human papillomavirus and head and neck cancer: epidemiology and molecular biology.* Head & neck, 1998. **20**(3): p. 250-265.

27.   Park, T., Fujiwara, H. and Wright, T. C. *Molecular biology of cervical cancer and its precursors.* Cancer, 1995. **76**(10): p. 1902-1913.

28.   Szentirmay, Z., Pólus, K., Tamás, L., Szentkuti, G., Kurcsics, J., *et al. Human papillomavirus in head and neck cancer: molecular biology and clinicopathological correlations.* Cancer and Metastasis Reviews, 2005. **24**(1): p. 19-34.

29.   Motoyama, S., Ladines-Llave, C. A., Villanueva, S. L. and Maruo, T. *The role of human papilloma virus in the molecular biology of cervical carcinogenesis.* Kobe Jounal of Medical Sciences, 2004. **50**(1-2): p. 9-19.

30.   McBride, A. A., Romanczuk, H. and Howley, P. M. *The papillomavirus E2 regulatory proteins.* Journal of Biological Chemistry, 1991. **266**(1841): p. 1-18414.

31.   Doorbar, J., Ely, S., Sterling, J., Mclean, M. and Crawford, L. *Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network.* Nature, 1991. **352**: p. 824-827.

32.   Brown, D. R., Fan, L., Jones, J. and Bryan, J. *Colocalization of human papillomavirus type 11 E1 [symbol: see text] E4 and L1 proteins in human foreskin implants grown in athymic mice.* Virology, 1994. **201**(1): p. 46.

33.   Tommasino, M. and Crawford, L. *Human papillomavirus E6 and E7: proteins which deregulate the cell cycle.* Bioessays, 1995. **17**(6): p. 509-518.

34.   Androphy, E. J. *Molecular biology of human papillomavirus infection and oncogenesis.* Journal of investigative dermatology, 1994. **103**(2): p. 248-256.

35.   Huibregtse, J. M., Scheffner, M. and Howley, P. M. *A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18.* The EMBO Journal, 1991. **10**(13): p. 4129.

36.   Werness, B. A., Levine, A. J. and Howley, P. M. *Association of human papillomavirus types 16 and 18 E6 proteins with p53.* Science, 1990. **248**(4951): p. 76-79.

37.   Lane, D. P. *P53 and human cancers.* British medical bulletin, 1994. **50**(3): p. 582-599.

38.   El-Deiry, W. S., Tokino, T., Velculescu, V. E., Levy, D. B., Parsons, R., *et al.* *< i> WAF1</i>, a potential mediator of p53 tumor suppression.* Cell, 1993. **75**(4): p. 817-825.

39.   Phelps, W. C., Bagchi, S., Barnes, J. A., Raychaudhuri, P., Kraus, V., *et al.* *Analysis of trans activation by human papillomavirus type 16 E7 and adenovirus 12S E1A suggests a common mechanism.* Journal of virology, 1991. **65**(12): p. 6922-6930.

40.   Phelps, W. C., Munger, K., Yee, C. L., Barnes, J. A. and Howley, P. M. *Structure-function analysis of the human papillomavirus type 16 E7 oncoprotein.* Journal of virology, 1992. **66**(4): p. 2418-2427.

41.   Hartwell, L. H. and Kastan, M. B. *Cell cycle control and cancer.* Science, 1994. **266**(5192): p. 1821-1828.

42.   Klingelhutz, A. J. and Roman, A. *Cellular transformation by human papillomaviruses: lessons learned by comparing high-and low-risk viruses.* Virology, 2012. **424**(2): p. 77.

43.   Dauner, J. G., Pan, Y., Hildesheim, A., Harro, C. and Pinto, L. A. *Characterization of the HPV-specific memory B cell and systemic antibody responses in women receiving an unadjuvanted HPV16 L1 VLP vaccine.* Vaccine, 2010. **28**(33): p. 5407-5413.

44.   Wang, J. W. and Roden, R. B. S. *Virus-like particles for the prevention of human papillomavirus-associated malignancies.* Expert review of vaccines, 2013. **12**(2): p. 129-141.

45.   Liu, M. A. *DNA vaccines: a review.* Journal of internal medicine, 2003. **253**(4): p. 402-410.

46.   Pfeifer, A. and Verma, I. M. *Gene therapy: promises and problems*. Annual review of genomics and human genetics, 2001. **2**(1): p. 177-211.

47.   Phillips, A. J. *The challenge of gene therapy and DNA delivery*. Journal of Pharmacy and Pharmacology, 2001. **53**(9): p. 1169-1174.

48.   Denèfle, P.P., *Introduction to gene therapy: a clinical aftermath*, in *Viral Vectors for Gene Therapy*2011, Springer. p. 27-44.

49.   Smith, K. R. *Gene therapy: theoretical and bioethical concepts*. Archives of medical research, 2003. **34**(4): p. 247-268.

50.   Verma, I. M. and Weitzman, M. D. *Gene therapy: twenty-first century medicine*. Annual Review of Biochemistry, 2005. **74**: p. 711-738.

51.   Sousa, A., Sousa F. and Queiroz, J. A. *Biorecognition of supercoiled plasmid DNA isoform in lysine-affinity chromatography*. Journal of chromatography B, 2009. **877**(27): p. 3257-3260.

52.   Wahren, B. *Gene vaccines*. Immunotechnology, 1996. **2**(2): p. 77-83.

53.   Iezzi, M., Quaglino, E., Amici, A., Lollini, P. and Cavallo, F. *DNA vaccination against oncoantigens: A promise*. Oncoimmunology, 2012. **1**(3): p. 316-325.

54.   Tighe, H., Corr, M., Roman, M. and Raz, E. *Gene vaccination: plasmid DNA is more than just a blueprint*. Immunology today, 1998. **19**(2): p. 89-97.

55.   Shedlock, D. J. and Weiner, D. B. *DNA vaccination: antigen presentation and the induction of immunity*. Journal of leukocyte biology, 2000. **68**(6): p. 793-806.

56.   Gurunathan, S., Klinman, D. M. and Seder, R. A. *DNA vaccines: immunology, application, and optimization**. Annual review of immunology, 2000. **18**(1): p. 927-974.

57.   Kim, T. W., Hung, C., Ling, M., Juang, J., He, L., *et al. Enhancing DNA vaccine potency by coadministration of DNA encoding antiapoptotic proteins*. Journal of Clinical Investigation, 2003. **112**(1): p. 109-117.

58.   Liu, M .A. and Ulmer, J. B. *Human clinical trials of plasmid DNA vaccines*. Advances in genetics, 2005. **55**: p. 25-40.

59.   Patil, S. D., Rhodes, D. G., and Burgess, D. J. *DNA-based therapeutics and DNA delivery systems: a comprehensive review*. The AAPS journal, 2005. **7**(1): p. E61-E77.

60.   Pathak, A., Patnaik, S. and Gupta, K. C. *Recent trends in non-viral vector-mediated gene delivery*. Biotechnology journal, 2009. **4**(11): p. 1559-1572.

61.   Srivastava, I. K. and Liu, M. A. *Gene Vaccines [dhelix]*. Annals of internal medicine, 2003. **138**(7): p. 550-559.

62.   Gao, X., Kim, K. and Liu, D. *Nonviral gene delivery: what we know and what is next*. The AAPS journal, 2007. **9**(1): p. E92-E104.

63.   Liu, F. and Huang, L. *Development of non-viral vectors for systemic gene delivery*. Journal of controlled release, 2002. **78**(1): p. 259-266.

64.   Rubanyi, G.M. *The future of human gene therapy*. Molecular aspects of medicine, 2001. **22**(3): p. 113-142.

65.   Ragusa, A., García, I. and Penadés, S. *Nanoparticles as nonviral gene delivery vectors*. NanoBioscience, IEEE Transactions on, 2007. **6**(4): p. 319-330.

66.   Glover, D. J., Lipps, H. J. and Jans, D. A. *Towards safe, non-viral therapeutic gene expression in humans*. Nature Reviews Genetics, 2005. **6**(4): p. 299-310.

67.   Edelstein, M. L., Abedi, M. R. and Wixon, J. *Gene therapy clinical trials worldwide to 2007—an update*. The journal of gene medicine, 2007. **9**(10): p. 833-842.

68.   Ferreira, G. N. M., Monteiro, G. A., Prazeres, D. M. F. and Cabral, J. M. S. *Downstream processing of plasmid DNA for gene therapy and DNA vaccine applications*. Trends in biotechnology, 2000. **18**(9): p. 380-388.

69.   Diogo, M. M., Queiroz, J. A. and Prazeres, D. M. F. *Chromatography of plasmid DNA*. Journal of Chromatography A, 2005. **1069**(1): p. 3-22.

70.   Urthaler, J., Buchinger, W. and Necina, R. *Improved downstream process for the production of plasmid DNA for gene therapy*. Acta Biochimica Polonica, 2005. **52**(3): p. 703.

71.   Sousa, A., Tomaz, C. T., Sousa, F. and Queiroz, J. A. *Successful application of monolithic innovative technology using a carbonyldiimidazole disk to purify supercoiled plasmid DNA suitable for pharmaceutical applications*. Journal of Chromatography A, 2011. **1218**(46): p. 8333-8343.

72.   Wells, D. J. *Gene therapy progress and prospects: electroporation and other physical methods*. Gene therapy, 2004. **11**(18): p. 1363-1369.

73.   Sousa, A., Sousa, F. and Queiroz, J. A. *Differential interactions of plasmid DNA, RNA and genomic DNA with amino acid-based affinity matrices*. Journal of separation science, 2010. **33**(17-18): p. 2610-2618.

74.   Prazeres, D. M. F., Ferreira, G. N. M., Monteiro, G. A., Cooney, C. L. and Cabral, J. M. S. *Large-scale production of pharmaceutical-grade plasmid DNA for gene therapy: problems and bottlenecks*. Trends in biotechnology, 1999. **17**(4): p. 169-174.

75.   Ghanem, A., Healey, R., and Adly, F. G. *Current trends in separation of plasmid DNA vaccines: A review*. Analytica chimica acta, 2012. **760**: p. 1-15.

76.   Kazakevich, Y. and R. LoBrutto, *Size-Exclusion Chromatography*. HPLC for Pharmaceutical Scientists: p. 263-279.

77.   Pfaunmiller, E. L., Paulemond, M. L., Dupper, C. M. and Hage, D. S. *Affinity monolith chromatography: a review of principles and recent analytical applications*. Analytical and bioanalytical chemistry, 2013. **405**(7): p. 2133-2145.

78.   Arakawa, T., Tsumoto, K., Kita, Y., Chang, B. and Ejima, D. *Biotechnology applications of amino acids in protein purification and formulations*. Amino Acids, 2007. **33**(4): p. 587-605.

79.   Hoffman, M. M., Khrapov, M. A., Cox, J. C., Yao, J., Tong, L., *et al. AANT: The amino acid-nucleotide interaction database*. Nucleic acids research, 2004. **32**(suppl 1): p. D174-D181.

80.   Luscombe, N. M., Laskowski, R. A. and Thornton, J. M. *Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level*. Nucleic acids research, 2001. **29**(13): p. 2860-2874.

81.   Sousa, Â., Sousa, F. and Queiroz, J. A. *Advances in chromatographic supports for pharmaceutical-grade plasmid DNA purification*. Journal of separation science, 2012. **35**(22): p. 3046-3058.

82.   Münger, K., Phelps, W. C., Bubb, V., Howley, P. M. and Schlegel, R. *The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes*. Journal of virology, 1989. **63**(10): p. 4417-4421.

83.   Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning*. Vol. 2. 1989: Cold spring harbor laboratory press New York.

84.   Diogo, M. M., Queiroz, J. A., Monteiro, G. A., Martins, S. A. M., Ferreira, G. N. M., *et al. Purification of a cystic fibrosis plasmid vector for gene therapy using hydrophobic interaction chromatography*. Biotechnology and bioengineering, 2000. **68**(5): p. 576-583.

85.   Ljunglöf, A., Bergvall, P., Bhikhabhai, R. and Hjorth, R. *Direct visualisation of plasmid DNA in individual chromatography adsorbent particles by confocal scanning laser microscopy*. Journal of Chromatography A, 1999. **844**(1): p. 129-135.

86.   Diogo, M. M., Queiroz, J. A. and Prazeres, D. M. F. *Assessment of purity and quantification of plasmid DNA in process solutions using high-performance hydrophobic interaction chromatography*. Journal of Chromatography A, 2003. **998**(1): p. 109-117.

87.   Martins, S. A. M., Prazeres, D. M. F., Cabral, J. M. S. and Monteiro, G. A. *Comparison of real-time polymerase chain reaction and hybridization assays for the detection of Escherichia coli genomic DNA in process samples and pharmaceutical-grade plasmid DNA products*. Analytical biochemistry, 2003. **322**(1): p. 127-129.

88.   O'Kennedy, R. D., Ward, J. M. and Keshavarz-Moore, E. *Effects of fermentation strategy on the characteristics of plasmid DNA production*. Biotechnology and applied biochemistry, 2003. **37**(1): p. 83-90.

89.   Carnes, A. E., Hodgson, C. P. and Williams, J. A. *Inducible Escherichia coli fermentation for increased plasmid DNA production*. Biotechnology and applied biochemistry, 2006. **45**(3): p. 155-166.

90.   Sousa, F., Matos T., Prazeres, D. M. F. and Queiroz, J. A. *Specific recognition of supercoiled plasmid DNA in arginine affinity chromatography*. Analytical biochemistry, 2008. **374**(2): p. 432-434.

91.   Freitas, S. S., Azzoni, A. R., Santos, J. A. L., Monteiro, G. A. and Prazeres, D. M. F. *On the stability of plasmid DNA vectors during cell culture and purification*. Molecular biotechnology, 2007. **36**(2): p. 151-158.

92.   Ongkudon, C. M. and Danquah, M. K. *Process optimisation for anion exchange monolithic chromatography of 4.2 kbp plasmid vaccine (pcDNA3F).* Journal of Chromatography B, 2010. **878**(28): p. 2719-2725.

93.   Prazeres, D. M. F., Schluep, T. and Cooney, C. *Preparative purification of supercoiled plasmid DNA using anion-exchange chromatography.* Journal of Chromatography A, 1998. **806**(1): p. 31-45.

94.   Seeman, N. C., Rosenberg, J. M. and Rich, A. *Sequence-specific recognition of double helical nucleic acids by proteins.* Proceedings of the National Academy of Sciences, 1976. **73**(3): p. 804-808.

95.   Sousa, F., Cruz, C. and Queiroz, J. A. *Amino acids-nucleotides biomolecular recognition: from biological occurrence to affinity chromatography.* Journal of Molecular Recognition, 2010. **23**(6): p. 505-518.

96.   Tiainen, P., Gustavsson, P., Ljunglof, A. and Larsson, P. *Superporous agarose anion exchangers for plasmid isolation.* Journal of Chromatography A, 2007. **1138**(1): p. 84-94.

97.   Eon-Duval, A. and Burke, G. *Purification of pharmaceutical-grade plasmid DNA by anion-exchange chromatography in an RNase-free process.* Journal of chromatography B, 2004. **804**(2): p. 327-335.

98.   Benčina, M., Podgornik, A. and Štrancar, A. *Characterization of methacrylate monoliths for purification of DNA molecules.* Journal of separation science, 2004. **27**(10-11): p. 801-810.

99.   Endres, H. N., Johnson, J. A. C., Ross, C. A., Welp, J. K. and Etzel, M. R. *Evaluation of an ion-exchange membrane for the purification of plasmid DNA.* Biotechnology and applied biochemistry, 2003. **37**(3): p. 259-266.

100.   Ongkudon, C. M. and Danquah, K. M. *Anion exchange chromatography of 4.2 kbp plasmid based vaccine (pcDNA3F) from alkaline lysed< i> E. coli</i> lysate using amino functionalised polymethacrylate conical monolith.* Separation and Purification Technology, 2011. **78**(3): p. 303-310.

101.   Sousa, F., Prazeres, D. M. F. and Queiroz, J. A. *Dynamic binding capacity of plasmid DNA in histidine-agarose chromatography.* Biomedical Chromatography, 2007. **21**(9): p. 993-998.

102.   Frankovič, V., Podgornik, A., Krajnc, N. L., Smrekar, F., Krajnc, P., *et al.* *Characterisation of grafted weak anion-exchange methacrylate monoliths.* Journal of Chromatography A, 2008. **1207**(1): p. 84-93.

103.   Sousa, F., Prazeres, D. M. F. and Queiroz, J. A. *Binding and elution strategy for improved performance of arginine affinity chromatography in supercoiled plasmid DNA purification.* Biomedical Chromatography, 2009. **23**(2): p. 160-165.

104.   Schlick, T., Olson, W. K., Westcott, T. and Greenberg, J. P. *On higher buckling transitions in supercoiled DNA.* Biopolymers, 1994. **34**(5): p. 565-597.

105. Ellis, J. J., Broom, M. and Jones, S. *Protein-RNA interactions: structural analysis and functional classes.* Proteins: Structure, Function, and Bioinformatics, 2007. **66**(4): p. 903-911.

106. Ferreira, G. N. M. *Chromatographic approaches in the purification of plasmid DNA for therapy and vaccination.* Chemical engineering & technology, 2005. **28**(11): p. 1285-1294.

107. Diogo, M. M., Ribeiro, S. C., Queiroz, J. A., Monteiro, G. A., Tordo, N., *et al. Production, purification and analysis of an experimental DNA vaccine against rabies.* The journal of gene medicine, 2001. **3**(6): p. 577-584.

108. Schluep, T. and Cooney, C. L. *Purification of plasmids by triplex affinity interaction.* Nucleic acids research, 1998. **26**(19): p. 4524-4528.

109. Wils, P., Escriou, V., Warnery, A., Lacroix, F., Lagneaux, D., *et al. Efficient purification of plasmid DNA for gene transfer using triple-helix affinity chromatography.* Gene Therapy, 1997. **4**(4): p. 323.

110. Smrekar, F., Podgornik, A., Ciringer, M., Kontrec, S., Raspor, P., *et al. Preparation of pharmaceutical-grade plasmid DNA using methacrylate monolithic columns.* Vaccine, 2010. **28**(8): p. 2039-2045.

111. Briggs, J. and Panfili, P. R. *Quantitation of DNA and protein impurities in biopharmaceuticals.* Analytical chemistry, 1991. **63**(9): p. 850-859.

112. Mattiasson, B., Teeparuksapun, K. and Hedström, M. *Immunochemical binding assays for detection and quantification of trace impurities in biotechnological production.* Trends in biotechnology, 2010. **28**(1): p. 20-27.

113. Cotten, M., Baker, A., Saltik, M., Wagner, E. and Buschle, M. *Lipopolysaccharide is a frequent contaminant of plasmid DNA preparations and can be toxic to primary cells in the presence of adenovirus.* Gene therapy, 1994. **1**: p. 239-246.

114. Butash, K. A., Natarajan, P., Young, A. and Fox, D. K. *Reexamination of the effect of endotoxin on cell proliferation and transfection efficiency.* BioTechniques, 2000. **29**(3): p. 610-619.

115. Wei, Z., Huang, W., Li, J., Hou, G., Fang, J., *et al.* *Studies on endotoxin removal mechanism of adsorbents with amino acid ligands.* Journal of Chromatography B, 2007. **852**(1): p. 288-292.