UNIVERSIDADE DA BEIRA INTERIOR
Ciências da Saúde

Obtenção de vacina de DNA plasmídico HPV-16 E6/E7 e avaliação da sua imunogenicidade *in vitro* e *in vivo*

Ana Margarida Cardoso Valério de Almeida

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 2014

‘Imagination will often carry us to worlds that never were. But without it we go nowhere’

-Carl Sagan

To the most important person in my life, my mother.

I love you.

Acknowledgements

Firstly, I would like to express my profound gratitude to Professor Doctor Ângela Sousa and Professor Doctor Fani Sousa for all the support and guidance that both were able to provide me throughout this year. Without your expertise, knowledge, helpfulness and faith I would not be able to progress as far as I did.

To Professor Doctor João Queiroz, for being actively involved in this work, scientifically supporting this project and contributing with useful ideas.

To University of Beira Interior, especially to the Health Sciences Research Centre for providing the adequate conditions for the development of this research project, and to Bia Separations company for kindly providing the monolithic support used in this project.

To Patrícia Pereira, whose precious patience and assistance allowed me to overcome the daily hurdles of research. I sincerely thank you for all the support and all the knowledge you shared with me, without you I would not been able to accomplish half of what I did this year.

To Joana Tomás, who patiently guided me through the transfection studies, helping whenever she could. Your availability and help was very important for the development of this work.

To Susana Ferreira and Adriana Afonso, for helping me to better understand the experimental design tool and to interpret the experimental design data.

To all the people from Health Sciences Research Centre of the University of Beira Interior, especially to the Biotechnology and Biomolecular Sciences group, for all the help and advices provided.

To my lab colleagues, thank you for all the funny moments, the long talks and the knowledge we were able to share throughout this year.

To my mother, who was able to support me all my life, believing in me everytime. If it was not for you I would not be able to be here. I love you and I sincerely thank you for all you did for me, from the bottom of my heart.

To all of my friends, specially to Debbie, Maddie, Fani, Seni, Luís, Tixa, João Filipe, Daniela, Andreia, Duarte, Diogo, Henrique, Marília, Luís Daniel, André, Ana Catarina, Catarina Chendo, Catarina Nascimento, Mariana, for your endless friendship and for putting up with me throughout all these years. You are my chosen family, thank you.

To João, for your love and support. Thank you for not letting me fall apart.

Resumo Alargado

A constante evolução da ciência tem permitido uma melhor partilha de conhecimentos na área da tecnologia do DNA recombinante, fornecendo um melhor conhecimento da informação contida nos genes e o impacto que alterações nesses genes poderão ter no organismo. A descodificação do genoma humano aliada ao progresso obtido no desenvolvimento de variados vetores de transporte de informação genética permitiu a evolução de terapias baseadas na entrega de genes terapêuticos, como a terapia génica e as vacinas de DNA. O desenvolvimento destas terapias trouxe uma nova esperança para o tratamento de certas patologias que, até então, permaneciam como intratáveis. Vetores biológicos e não biológicos têm evoluído largamente nos últimos anos, no entanto, a toxicidade demonstrada pela maioria dos vetores biológicos tem levado a um aumento de utilização de vetores não biológicos.

O DNA plasmídico destaca-se entre os diversos vetores genéticos devido à simplicidade da sua produção, obtenção, baixo custo e ausência de toxicidade. As vantagens deste vetor têm levado a que a sua utilização como vacina de DNA tenha aumentado nos últimos anos, tornando-o o vetor de escolha na maioria dos estudos de investigação. As vacinas de DNA têm como modo de atuação a expressão de proteínas antigénicas com o objetivo de induzir uma resposta imunitária direcionada para essas mesmas proteínas, permitindo a prevenção e/ou tratamento de infeções virais e bacterianas. Torna-se imperativo o desenvolvimento de tecnologias que permitam a produção e purificação destes vetores, obtendo a maior percentagem de recuperação e pureza possíveis do plasmídeo na sua forma biologicamente ativa, a isoforma superenrolada (sc). A área da cromatografia tem progredido bastante no desenvolvimento de estratégias eficazes de purificação de plasmídeo, permitindo o aumento de produtividade e obtenção deste vetor e diminuindo eventuais custos associados à sua produção.

O Vírus do Papiloma Humano (HPV) é um vírus sexualmente transmissível que se encontra atualmente associado ao desenvolvimento de massas tumorais devido à produção de duas proteínas oncogénicas, oncoproteínas E6 e E7, capazes de alterar o ciclo de proliferação celular e de provocar o crescimento anormal de células do organismo infetado. A tecnologia de vacinas de DNA apresenta-se assim como uma terapia promissora para infeções provocadas pelo HPV, através da indução de uma resposta imunitária contra as proteínas referidas.

Recentemente, o nosso grupo de investigação conseguiu desenvolver de forma eficaz a produção e purificação da vacina de DNA sc HPV-16 E6/E7 através da utilização de um monolito modificado com ligandos de arginina, tirando partido dos princípios básicos da cromatografia de afinidade. Contudo, a recuperação do plasmídeo não foi a esperada, tendo sido apenas recuperado 39% da molécula alvo.

O Desenho experimental é uma ferramenta estatística que, através da escolha correta dos fatores a serem avaliados, bem como os seus intervalos em estudo, permite a otimização de respostas de um sistema experimental. Deste modo, através do design experimental foi feita uma otimização ao sistema de purificação da vacina de DNA sc HPV-16 E6/E7 de modo a garantir um aumento de recuperação da molécula, mantendo o elevado nível de pureza. Com esse intuito, após uma avaliação inicial dos fatores e dos intervalos a serem usados, o design 'Central Composite Face' (CCF) foi utilizado para delinear um conjunto de experiências cromatográficas de modo a encontrar o ponto ótimo para a percentagem de recuperação do plasmídeo ser maximizada. A otimização foi bem-sucedida, permitindo a obtenção de uma percentagem de recuperação de cerca de 83%, mantendo-se a percentagem de 100% para a pureza.

Após a otimização da estratégia de purificação, estudos de transfeção *in vitro* foram realizados de modo a avaliar a capacidade de transfeção celular e consequente expressão da proteína codificada pelo gene-alvo contido na vacina de DNA. Células CHO-1, isoladas a partir de tecido ovárico de rato chinês, foram cultivadas e transfetadas com a isoforma sc purificada através da estratégia otimizada com o monolito de arginina, bem como com a isoforma circular aberta (oc) e DNA plasmídico obtido através de um kit comercial, de modo a avaliar qual a melhor estratégia para transfeção. Através das técnicas de western blot e imunocitoquímica foi possível verificar que a entrada do pDNA nas células eucarióticas ocorreu com sucesso (processo de transfeção), observando-se um aumento significativo de expressão génica das proteínas E6 e E7 em comparação ao grupo de controlo (células não transfetadas). A avaliação da expressão génica da proteína E6 dos diferentes tipos de plasmídeos utilizados permitiu verificar que o aumento de expressão desta proteína foi mais significativo com a amostra de plasmídeo sc purificado pelo monolito de arginina, concluindo-se que de facto a isoforma sc induz uma maior eficiência de transfeção.

Palavras-chave

Desenho experimental; Cromatografia de afinidade; transfeção; plasmídeo HPV16-E6/E7 superenrolado; vacinas de DNA.

Abstract

The infection by Human Papilloma Virus (HPV) is associated with the development of different tumours, in particular the cervical cancer. Oncoproteins E6 and E7, produced by this virus, are responsible for the disturbance of the cell cycle, through interaction with several proto-oncogenes, leading to uncontrolled proliferation of the infected host cells. Therefore, the development of a suitable therapy against HPV infection with these oncoproteins is a promising strategy. DNA vaccines arise as a potential therapeutic solution in cancer treatment, being able to trigger a strong immune response against the target antigen, normally expressed by the infected cells. The purification of supercoiled (sc) plasmid HPV16 E6/E7 DNA vaccine with the arginine monolith was recently developed by our research group. In spite of achieving 100% purity, only 39% of the target molecule was recovered. Experimental design is a new tool able to project several experiments, by evaluating and combining different factors, with the intent of improving and optimizing a given experiment. Through the use of Composite Central Face design and the choice of three factors to be evaluated, such as binding step, washing step and pH, different experiments were performed in order to achieve the optimal range for the sc HPV16 E6/E7 purity and recovery. The aim was successfully achieved with 83% of recovery and 100% of purity. Thereafter, transfection studies were performed in order to evaluate the plasmid DNA (pDNA) vaccine efficiency. Several plasmid samples obtained from different purification methods were tested: plasmid purified by a commercial kit, open circular isoform (oc) and sc isoform purified by our optimized strategy with the arginine monolith. After 72 hours of transfection, the expression of E6 protein in CHO-1 cells was evaluated through immunocytochemistry. Through immunofluorescence comparison, higher E6 protein expression was detected by sc pDNA, showing a significant increase, when compared to control group. On the other hand, pDNA purified with the commercial kit and oc pDNA had no significant immunofluorescence different in comparison with control group. These data suggest that the sc pDNA obtained by our optimized purification strategy is able to efficiently transfect cells and express the target proteins, encouraging us to proceed to *in vivo* studies in order to evaluate the immunogenicity of this DNA vaccine.

Keywords

Affinity chromatography; Experimental design; supercoiled HPV16 E6/E7 plasmid; DNA vaccines; Transfection.

Table of Contents

CHAPTER 1 - INTRODUCTION 1

1.1 DNA-BASED THERAPY 1

1.1.1 Gene therapy 1

1.1.2 DNA vaccines 2

1.1.3 Biological vs non-biological DNA therapy approaches 4

1.1.3.1 Biological approaches 4

1.1.3.2 Non-biological approaches 5

1.2 HUMAN PAPILLOMAVIRUS 9

1.2.1 Human papillomavirus molecular biology 10

1.2.2 Papillomavirus E6 oncoprotein 11

1.2.3 Papillomavirus E7 oncoprotein 12

1.2.4 Preventive and therapeutic vaccination 13

1.3 PLASMID DNA TECHNOLOGY 14

1.3.1 Construction, production and primary isolation of pDNA 14

1.3.2 Plasmid DNA purification 16

1.3.3.1 Monoliths: a new chromatographic support technology 19

1.3.4 Design of experiments 20

1.3.5 Plasmid application 21

CHAPTER 2 – GLOBAL AIMS 25

CHAPTER 3 - MATERIALS AND METHODS.....27

3.1 PRODUCTION.....27

3.2 PURIFICATION.....27

3.2.1 Alkaline lysis with NZYTech kit 27

3.2.2 Modified alkaline lysis 27

3.2.3 Affinity chromatography 28

3.2.3.1 Agarose gel electrophoresis 29

3.2.4 Design of experiments.....29

3.2.4.1 Supercoiled pDNA quantification 29

3.3 CELL CULTURE.....30

3.3.1 Transfection.....30

3.3.2 Protein extraction.....31

3.4 WESTERN BLOT 31

3.4.1 Protein quantification 31

3.4.2 Polyacrylamide gel electrophoresis 32

3.4.3 Electroblotting.....32

3.5 IMMUNOCYTOCHEMISTRY 33

3.5.1 Immunofluorescence analysis 33

CHAPTER 4 – RESULTS AND DISCUSSION.....35

4.1 EXPERIMENTAL DESIGN.....35

4.1.1 Preliminary tests for threshold choice.....35

4.1.2 Optimization 42

4.1.2.1 Goodness of fit 46

4.1.2.2 Residual plots 47

4.1.2.3 Predicted versus Actual plots 48

4.1.2.4 Main effects 48

4.1.2.5 Surface and contour plots 50

4.1.2.6 Analysis of variance 53

4.1.3 Model validation 56

4.2 – TRANSFECTION EFFICIENCY 58

4.2.1 Western blot 59

4.2.2 – Immunocytochemistry 60

4.2.2.1 E6 immunofluorescence quantification 62

4.2.2.2 E7 immunofluorescence quantification 64

CHAPTER 5 – CONCLUSIONS AND FUTURE PERSPECTIVES 69

REFERENCES 73

List of Figures

CHAPTER 1- Introduction

Figure 1. Schematic representation of APC pathways 3

Figure 2. Vectors used in DNA-based therapy trials. 4

Figure 3. HPV genome structure. 10

Figure 4. Evolution of infection by HPV. 11

Figure 5. Degradation of p53 in the presence of HPV 16 E6. 12

Figure 6. Schematic representation of HPV-16 E6/E7 pDNA 15

Figure 7. Production, recovery and purification of pDNA in order to achieve high pharmaceutical grade for application 16

Figure 8. Schematic representation of arginine modified monolithic support 20

Figure 9. Schematic representation of CCF design 21

CHAPTER 3 – Materials and methods

Figure 10. Calibration curve with plasmid DNA standards (1-25 µg/mL) 30

Figure 11. Calibration curve with protein standards (0.2-10 µg/µL) 32

CHAPTER 4 – Results and discussion

Figure 12. Chromatographic profile of the *E. coli* clarified lysate sample injected in the arginine monolith and agarose electrophoresis of the different peaks obtained 36

Figure 13. Evaluation of the flow rate effect in sc pDNA purification 37

Figure 14. pH influence on the sc pDNA purification with the arginine monolith 38

Figure 15. Example of the chromatographic profiles obtained in the analysis with the CIMac™ analytical column. 41

Figure 16. Graphical representation of the statistical coefficients of the model 46

Figure 17. Residual plot or recovery and purity 47

Figure 18. Predicted versus Actual plot for recovery and purity 48

Figure 19. Surface plot and contour plot for recovery response..... 51

Figure 20. Surface plot and contour plot for purity response 52

Figure 21. Chromatographic replicate runs for model validation of the optimal point 57

Figure 22. Western Blot for E6 and E7 proteins 59

Figure 23. CHO-1 cells immunocytochemistry images for E6 staining 61

Figure 24. CHO-1 cells immunocytochemistry images for E7 staining 62

Figure 25. E6 protein immunofluorescence comparison between the different groups (pvalue<0.001) 64

Figure 26. E7 protein Immunofluorescence comparison between the CT group and sc pDNA transfected group (pvalue<0.05)..... 66

List of Tables

CHAPTER 1- Introduction

Table 1. DNA-based therapy clinical trials 2

Table 2. Classification of several HPVs..... 9

Table 3. Specifications for pDNA to be considered as safe for administration. 17

CHAPTER 4 – Results and discussion

Table 4. Chromatographic conditions for pH range assessment..... 39

Table 5. Chosen factors versus levels of CCF design..... 42

Table 6. CCF design for three levels..... 43

Table 7. Designed experiments according to CCF design 43

Table 8. Responses of sc pDNA recovery and purity obtained for each run defined by CCF design..... 45

Table 9. Statistical coefficients of the model..... 46

Table 10. Summary of the main effects..... 49

Table 11. ANOVA table for recovery response 54

Table 12. ANOVA table for purity response..... 55

Table 13. Optimal range for recovery and purity maximization 56

Table 14. Responses obtained from the optimization/validation of the model 58

Table 15. Confidence intervals for both responses..... 58

Table 16. Immunofluorescence for CT, kit, oc and sc pDNA transfected groups for E6 protein (n=3) 63

Table 17. Immunofluorescence for CT, kit, oc and sc pDNA transfected groups for E7 protein (n=3) 65

List of Acronyms

ANOVA	Analysis of variance
µg	Microgram
µL	Microliter
µm	Micrometer
APCs	Antigen-presenting cells
BCA	Bicinchoninic acid assay
BSA	Bovine Serum Albumin
CCF	Central Composite Face
°C	Centigrade
CI	Confidence intervals
CIN	Cervical intraepithelial neoplasia
Cl-	Chloride ion
CO2	Carbon dioxide
CR	Conserved regions
CT	Control
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
E6AP	E6 Association Protein
ECF	Enhanced chemifluorescence
EDTA	Ethylene-diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid

F	Ratio of means of squares
FBS	Fetal bovine serum
FDA	Food and Drug Administration
gDNA	Genomic DNA
HCl	Hydrochloric acid
HPV	Human Papilloma Virus
IMAC	Immobilized metals affinity chromatography
K_2HPO_4	Dipotassium phosphate
kDa	Kilo Daltons
KH_2PO_4	Monopotassium phosphate
LAL	Limulus amebocyte lysate
Ln	Linear
M	Molar
mA	Miliampere
mAU	Miliabsorbance units
MHC	Major histocompatibility complex
min	Minute
mL	Mililliter
mM	Milimolar
NaCl	Sodium chloride
NaOH	Sodium hydroxide
nm	Nanometer
oc	Open circular

OD_{600}	Optical density at 600 nm
OH-	Hydroxide ion
ORFs	Open reading frames
PBS	Phosphate buffered saline
PBS-T 0.1%	PBS with 0.1% Tween
PCR	Polymerase chain reaction
pDNA	Plasmid DNA
PEI	Polyethylenimine
pKa	Acid dissociation constant
PLL	Poly-L-Lysine
PMSF	Phenylmethylsulfonyl fluoride
pRB	Gene product of retinoblastoma tumour suppressor
p-value	Significance probability
RNA	Ribonucleic acid
rpm	Rotations per minute
sc	Supercoiled
SCC	Squamous skin cancer
SDS	Sodium dodecylsulfate
TAE	Tris-acetate-EDTA
TBS-T 0.1%	Tris-buffered saline solution with 0.1% Tween 20
THAC	Triple-helix affinity chromatography
Tris-EDTA	10 mM Tris-HCl and 10 mM EDTA
URR	Upstream regulatory region

V	Volts
v:v	Volume:volume
VLPs	Viral-like proteins

xxx

CHAPTER 1 - Introduction

1.1 DNA-based therapy

Science has evolved in the past century at an astonishing velocity. Great discoveries such as the DNA structure by Watson and Crick [1], in 1953, allowed the Science progress at a faster rhythm, changing thoughts and breaking dogmas along the way. DNA is responsible for the coding of all the information within the cell machinery, allowing the synthesis of proteins and other cell components. From the Stanley Cohen's successful attempt in 1973 to join DNA molecules with different origins into a plasmid construct [2], the DNA recombinant technology was born and promptly grew throughout the years. The possibility for the arrangement of different genes with different functions permitted to establish the conditions for the hatching of a new area in Genetic Engineering, the DNA-based therapy. The possibility of administrating genes in order to achieve a therapeutic effect started to draw attention, rising as a promising pathway for therapy of several pathologies. Unravelling the human genome sequence in 2001 [3] allowed to understand all the information contained in the human genome, opening new horizons in the area of DNA-based therapy. Nowadays, DNA-based therapy can be subdivided into two widely studied areas, the gene therapy and DNA vaccines. Both DNA-based approaches are further discussed below.

1.1.1 Gene therapy

In the last decades, gene therapy has been targeted as a promising treatment for acquired diseases and genetic disorders. Ever since the evolution boost in molecular biology, other related fields gained some benefits, for example the production of considered amounts of nucleic acids through the use of polymerase chain reaction or the progression of the biotechnology field towards the improvement of transfection technology, which were permitted by the study and access to several different methods and techniques [4]. Gene therapy is based on the transfection of foreign nucleic acids encoding therapeutic information into the host cells, leading to a signal capable of correcting the target malfunction [5]. Therefore, this therapeutic design can be used for several purposes such as to add, eliminate or modify a particular function, sequence or expression of a given gene [6]. The ideal gene therapy vector should be able to carry the DNA information necessary to correct the target impairment, efficiently transfect the target cells or tissues and maintain an adequate expression or inhibition, without raising any biosafety concerns [7-9]. Nowadays, the areas where gene therapy may be applied are widening, although cancer diseases remain the most studied and, currently, represent the area with more ongoing gene therapy clinical trials worldwide, as portrayed in table 1.

Table 1. DNA-based therapy clinical trials (adapted from [10]).

Indications	Number	%
Cancer diseases	1274	63.8
Cardiovascular diseases	162	8.1
Gene marking	50	2.5
Healthy volunteers	52	2.6
Infectious diseases	164	8.2
Inflammatory diseases	13	0.7
Monogenic diseases	178	8.9
Neurological diseases	37	1.9
Ocular diseases	31	1.6
Others	35	1.8
Total	1996	

1.1.2 DNA vaccines

Conventional vaccines were based on administration of weakened or modified infectious agents to healthy individual only to prevent a specific disease, promoting antibody immunity [11]. Owing to the substantial morbidity and mortality associated with particular diseases worldwide, such as tuberculosis, malaria, leishmaniasis, human papillomavirus (HPV) infection or human immunodeficiency virus infection, an understanding of the mechanisms involved in generating long-lived cellular immune responses arises like a critical point [12].Therefore, a new form of vaccination, by using DNA vaccines, can stimulate both humoral and cellular immune responses [12, 13].

DNA vaccines consist in a novel approach in the immunology field, taking advantage of genetic information that is delivered to the system able to induce an immune response against a given antigen. Upon inoculation, the individual shall produce a strong and enduring immune response against the encoded protein antigen, associated to the pathology [14]. The pathway by which DNA vaccines are able to trigger an immune response begins with the transfection of antigen-presenting cells (APCs) and non-APCs, leading to the expression of the

antigen codified in the DNA-based vector. As depicted in figure 1.(a), transfected cells may present the antigen, through Major Histocompatibility Complex (MHC) class I, to CD8+ lymphocytes [15, 16]. The target antigen production by non-APCs results in its exogenous take up by APCs, which will present the antigen to CD4+ lymphocytes via MHC class II, as can be seen in figure 1.(b) [15-17]. Such antigen presentation leads to the rise of a solid immune response [18].

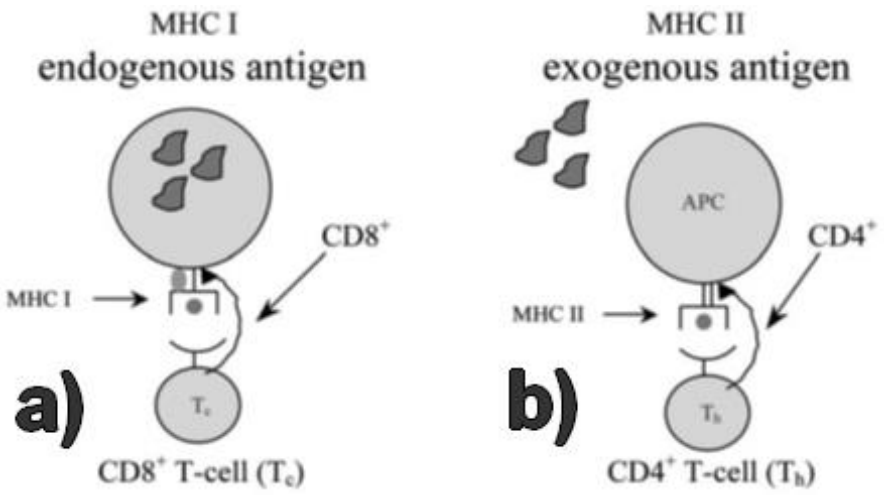

Figure 1. Schematic representation of APC pathways. A) Antigen presentation by MHC I. B) Antigen presentation by MHC II (adapted from [16]).

Beyond the capacity of DNA vaccines to generate all types of desired immunity, other advantages are related with DNA manufacturing, which can be easily obtained in large quantities with great purity, minimizing the risk of vaccine contamination with potential pathogens, being safety for a broad population administration [12]. In addition, the multivalent vaccination by administration of several antigenic genes reduces the total number of vaccinations that must be administered [19]. Furthermore, DNA vaccines also provide the potential for longer-lived antigen exposure *in vivo*, which could lead to increased immunogenicity, and could be easily transported and stored, because DNA provides more stability than other biological molecules [20].

Vaccination progress revealed that DNA vaccines can be used not only for preventive immunization but also as a versatile method to control and treat cancer [21-23]. It is estimated that, in 2012, 8.2 million people worldwide died of cancer, according to World Health Organization, and part of this problematic is caused by bacteria and viruses infection. Such numbers indicate that it is urgent to focus in alternative treatment strategies and recent studies suggest that DNA vaccines might be a partial answer for this problem.

1.1.3 Biological vs non-biological DNA therapy approaches

Nucleic acids, viruses or genetically engineered microorganisms, according to Food and Drug Administration (FDA), are an example of different vectors used for DNA therapy. Each vector present different advantages and drawbacks and can be divided in two separate categories, the biological and non-biological vectors [24].

1.1.3.1 Biological approaches

Viruses and bacteria are two different types of vectors that fit in this group. While viruses have been extensively studied and described in DNA therapy, bacteria have not yet been able to draw much attention in this area. Viruses represent 67% of the currently vectors in use in DNA therapy trials, as represented in figure 2. The strong application of this sort of vector might be related to the profound study of viruses in gene therapy, being referred to as the first-generation therapy vectors [9].

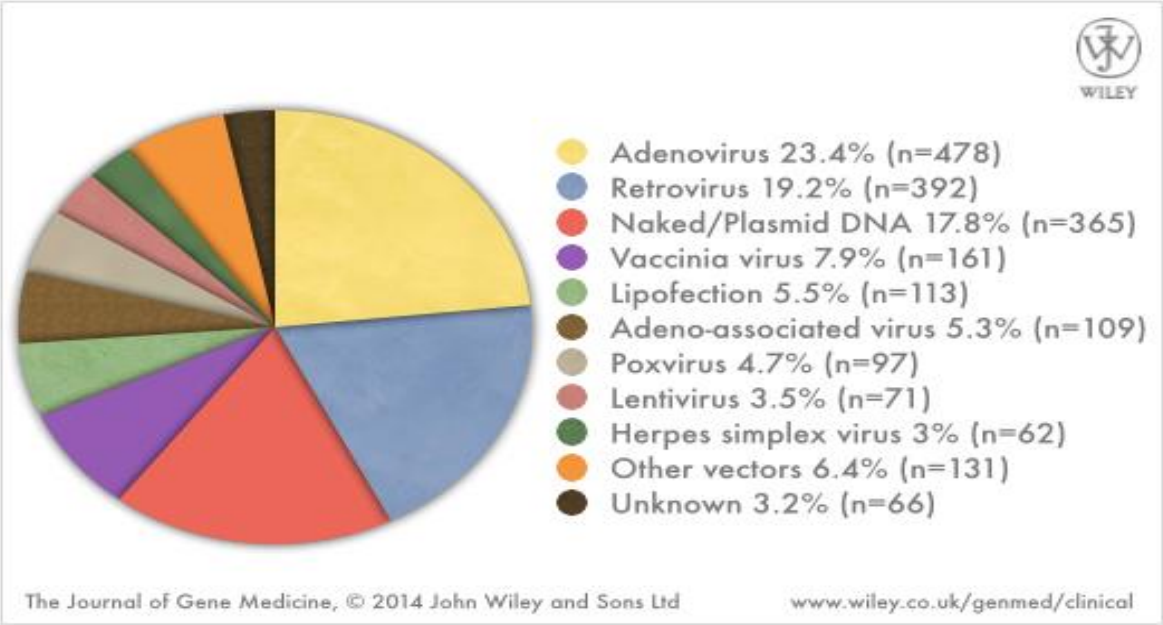

Figure 2. Vectors used in DNA-based therapy trials (adapted from [10]).

In general, the ability to easily infect the host cells is the main benefit of viruses in DNA therapy, contributing to increased gene expression levels and representing the key factor for viruses popularity. Different viruses are known to be used as vectors, however, adenoviruses and retroviruses stand out for being the most used. Adenoviruses are characterised by high efficiency transduction rates, capacity to transduce non-dividing and dividing cells and oncolytic power [9, 25, 26]. These appealing characteristics are probably the main reason behind the extensive study of this virus. However there are some drawbacks that cannot be overlooked.

In 1999, gene therapy faced the first death of a patient undergoing gene therapy treatment for nitrogen metabolism disorder ornithine transcarbamylase deficiency with an adenovirus which carried the gene for ornithine transcarbamylase [27]. Adenoviruses, being human pathogens, are highly immunogenic and its administration may be seen by the host defences as an infectious agent, triggering an immune response and, thereby, limiting its effectiveness and following administration, as well as compromising the patients health [9, 27, 28].

On the other hand, retroviruses present very low immunogenicity and are easily manipulated [25]. Nevertheless, their ability to randomly integrate the host cells in a permanent way raises several biosafety concerns [25]. The insertion of foreign nucleic acids into the host genome may disrupt important genetic information or be placed near a strong promoter, leading to its overexpression, perhaps contributing to the development of major health issues [26]. Another drawback in the use of viruses as DNA therapy vectors is the very small amount of DNA that they can carry, limiting the number and the size of genes to be inserted, and the high cost in its production [4].

Bacteria are microorganisms which present several properties useful for their application as DNA therapy vectors. Since 1980, when it was first reported the transfer of DNA into mammalian cells by bacteria [29], the study of bacteria in DNA therapy has come a long way. Nowadays, bacteria might be used in DNA therapy through two different approaches, namely, bactofection and tumour-specific bacterial replication [24]. Bactofection consists in the transfer of DNA into mammalian cells through the use of bacteria [24, 30, 31]. This procedure usually relies on using invasive type bacteria, which are capable of fully entering the target cells, and so deliver the genetic material [24, 30]. The tumour overgrowth can lead to the development of hypoxia, a condition which consists in low levels of oxygen due to insufficient blood supply [24, 30]. Under these circumstances, the use of anaerobic bacteria is useful in targeting tumour tissues, providing a specific delivery system of genetic material. Therefore, different non-invasive bacterial strains may be genetically engineered in order to secrete therapeutic proteins within the tumour cells [24, 30]. Invasive bacterial strains may also be used with the purpose of improving bactofection targeting [24]. However, safety issues associated with immunogenicity, bacterial sepsis and reversion of pathogenicity, suggest that further improvements are necessary [31].

1.1.3.2 Non-biological approaches

In the past years, non-biological vectors and techniques have been widely used, and their popularity is expected to keep rising. In 2004, 14% of DNA-based therapy clinical trials were developed with non-viral techniques, increasing to 18.3% in 2012 [32]. Although the general transfection efficiency of non-biological approaches is usually lower than biological vectors, non-biological techniques present certain characteristics that definitely surpass the apparent drawbacks. The main advantage of using such methodology relies on the biosafety of the

transfection procedure, besides allowing to transfect larger amounts of DNA at a reduced cost [33]. Plasmid DNA (pDNA) is currently very popular as a non-biological vector and has been widely studied in the biotechnology field. This double-stranded biomolecule has a simple low cost manufacture and is able to maintain acceptable expression levels, without presenting signs of pathogenicity [8, 32]. This vector is minutely described below, in section 1.3.

In general, the cellular entry of naked DNA is quite difficult, resulting in poor transfection and gene expression [5, 9, 34]. In order to overcome this limitation, naked DNA has been coupled with several physical and chemical methods which are able to increase the naked DNA cellular uptake.

1.1.3.2.1 Physical methods

Needle injection is a physical method with clinical interest for direct injection of naked DNA into tissues, organs or blood streams. The use of this methodology has been specially directed for DNA vaccines, due to the triggering of the immune system through the rupture of the skin by the needle [35]. Although needle injection has its advantages, like its simplicity and lack of toxicity, the poor level of gene expression has lead researchers to search for other alternatives [5].

Gene gun transfer is also a physical method often used when it is necessary to transfect skin, mucosa, surgically exposed tissues or tumour cells [5, 8]. This methodology consists in using gold particles coated with the target DNA [8]. Afterwards, the particles shall be accelerated through pressurized gas and expelled onto the target cells or tissue, delivering the DNA directly to the cytoplasm [8, 36]. Gene gun transfer is an appealing technique, especially for its simplicity and relative safety [8]. However, there are some drawbacks concerning the transient gene expression, which may imply several applications in order to achieve the therapeutic effect [5, 37].

On the other hand, electroporation is another physical method capable of successfully increase transgene expression in several types of tissue [5, 8]. Through the use of electrodes, a high-voltage electrical current is applied to the target cells, leading to the appearing of nanometric pores by which the naked DNA can enter into the cell [5, 38]. This technique allows delivering large DNA molecules, besides sustaining transgene expression for over a year [5]. Nevertheless, several disadvantages arise when using electroporation, namely the limited transfection area, tissue damage due to high voltage applied, and obligatory surgical procedure for treatment of internal organs [5, 8, 39].

1.1.3.2.2 Chemical methods

Cationic lipids represent one of the most commonly used complex agents for non-biological gene delivery, due to its lack of pathogenicity, low price and relatively simple production [5, 39, 40]. Cationic lipids include liposomes, closed spherical constructs which contain one or more structures of concentric phospholipid bilayers surrounded by an aqueous phase [40]. Plasmid DNA, when combined with liposomes, condenses into small particles named lipoplexes, which will protect the nucleic acids from enzymatic degradation and improve transfection efficiency [8]. Even though cationic lipids have been the subject of profound study in the past years, the successful mechanism of gene delivery and expression is still not fully understood. It is certain that factors such as chemical structure, size, surface charge, colloidal stability and others, contribute of transfection efficiency improvement [5, 8, 39]. After binding to cell surface due to opposite charges between lipoplexes and the cell-surface, the lipoplexes internalization occurs mainly by endocytosis [5, 34, 39]. In order to evade the lysosomal degradation, the endosomal release is necessary. Lipoplexes are known to destabilize the endosomal membrane, particularly with the help of co-lipids, which are able to relocate DNA contained in lipoplexes into cytoplasm, allowing the nuclear entry of DNA, since lipoplexes are too large to pass the nuclear pores [5, 8, 34, 39]. While this DNA therapy approach has been target of major improvements throughout the years, there are still some concerns regarding lipoplexes toxicity and the rapid plasma clearance [7].

Cationic polymers popularity has been growing in the past years due to its ability to condense nucleic acids through mild electrostatic interactions [41]. Nonetheless, different polymers result in different transfection activity and underlying toxicity. For instance, chitosan is one of the most studied non-biological gene carrier, due to its biosafety, even in high concentrations, and ability to effectively bind and compact DNA. However, it has low delivery efficiency in most cell lines, a drawback that has been trying to be surpassed by production of chitosan derivatives in order to improve its transfection ability [4, 41]. On the other hand, Poly-L-Lysine (PLL) was one of the first polymers to be utilized in gene delivery. Its charged amine groups interact with negative charged DNA, resulting in a good packaging of the DNA into nanosized particles. Although the uptake of PLL-DNA complexes is effective, the endosomal escape is still an obstacle when using such cationic polymer, leading to poor transfection activity. Polyethylenimine (PEI) is another polymer often used as a gene delivery carrier. Although its transfection rates *in vivo* are quite satisfactory, this polymer is nonbiodegradable, which may lead to an increase of toxicity in the cellular environment [8].

1.1.4 Ongoing clinical trials

Nowadays there are 1996 ongoing clinical trials focused on DNA therapy. The majority, representing 63.8% as depicted in table 1, are directed towards cancer diseases. Most of the 1274 cancer diseases targeting DNA-based therapies rely on the use of non-viral techniques.

As mentioned before, DNA vaccines have great potential in the treatment of such pathologies, due to the creation of an immune response towards to a target antigen.

For instance, Koen Oosterhuis and colleagues, have an ongoing trial with a DNA vaccine, consisting in the administration of naked pDNA which encodes a fusion protein domain1 of tetanus toxin fragment C and the shuffled version of the HPV E7 oncoprotein in order to produce E7 specific T cell immunity for the treatment of stage IV squamous cell cancer patients [42, 43]. Another example is the work developed by Cornelia Trimble and colleagues, which has an ongoing trial for the treatment of patients with stage III or IV HPV 16-positive head and neck squamous cell carcinoma, based on the development of a DNA vaccine, which also targets an immune response towards E7 oncoprotein through the administration of pDNA pNGVL4a-Sig/E7(detox)/HSP70 [44]. The E7 oncoprotein is produced when an infection by Human Papillomavirus (HPV) occurs and is responsible for transforming the host's cellular cycle, coupled with E6 protein, as it shall be further described. This problematic issue has drawn attention from several researchers, whom have been working in the development of different DNA-based therapies toward such oncoproteins [23, 45, 46]. With this in mind, our research group has focused in the development of a DNA vaccine able to encode both oncoproteins E6 and E7 with the purpose of potentiating an immune response towards both antigens, preventing the emergence of a future infection and helping in the treatment of an ongoing infection.

1.2 Human papillomavirus

The link between the development of several types of cancer with the infection by HPV has been extensively described in the past years. Cervical cancer is the second most common cancer amongst women, according to the World Health Organization, and a recent study made in Brazil found HPV DNA in 99% of 172 cases of invasive cervical cancer [47].

Besides cervical cancer, HPV has been also associated with anal, vaginal, vulval, penile, head and neck cancers [49]. There are several types of HPV, which are identified by DNA sequence similarity and reflect on the tissue tropism of the virus. HPV can be divided into two different phylogenetic groups, the alpha genus and the beta genus. While alpha HPVs infect mucosal tissues, beta HPVs infect cutaneous tissue, leading to different clinical outcomes [50]. HPV can also be categorized by risk of inducing cancer, being thus classified as high, intermediate and low risks. Table 2 serves as an example of classification for different HPVs.

Table 2. Classification of several HPVs (adapted from [48]).

Group	Prototypes	Site of infection	Acute consequences	Chronic consequences	Other features
Cutaneous (alpha)	HPV1, HPV2	Skin	Warts	None	Synchronous regression, lasting immunity
Mucosal (beta)	HPV6, HPV11,	Genital mucosa	Warts	None	Slow resolution in immunosuppressed individuals
Mucosal High risk (beta)	HPV16, HPV18, HPV31, HPV33, HPV45	Anogenital mucosa (other mucosa surfaces)	Flat lesion (CIN 1)	~2% persist, ~1% progress to invasive cancer	Slow resolution in immunosuppressed individuals, variable malignant potential
Cutaneous High risk (alpha)	HPV5, HPV8	Skin	Flat lesion or none warts	Promotes SCC	SCC more common in immunosuppressed individuals

CIN, cervical intraepithelial neoplasia; SCC, squamous skin cancer

HPV 16 and HPV18 are high risk HPVs, which are usually associated with development of cervical cancers. Therefore, the development of an efficient therapeutic against these HPV types gains a significant importance, in order to prevent or eliminate an infection by HPV.

1.2.1 Human papillomavirus molecular biology

Human Papillomavirus is one of the most common viruses responsible for sexually transmitted infections, despite not being only sexually transmitted [51]. Around 40 different types of HPVs can be found in the genital tract, from over 100 HPV types molecularly characterized [52]. This virus is strongly associated with the transformation of infected cells, what might lead to solid replication of abnormal cells and, consequently, to the tumour outgrowth [53].

The HPV genome, represented in figure 3, apart from being responsible for encoding 8 proteins, 6 of them at an early stage (E1, E2, E4-E7) and 2 of them at a late stage (L1 and L2), has a non-protein coding upstream regulatory region (URR) which contains cis-elements and is able of regulating the gene expression, genome replication and viral packaging [52]. While the early proteins are associated with the infection and transformation of the infected cells, the late proteins are accountable for the spread of the infection in the host system.

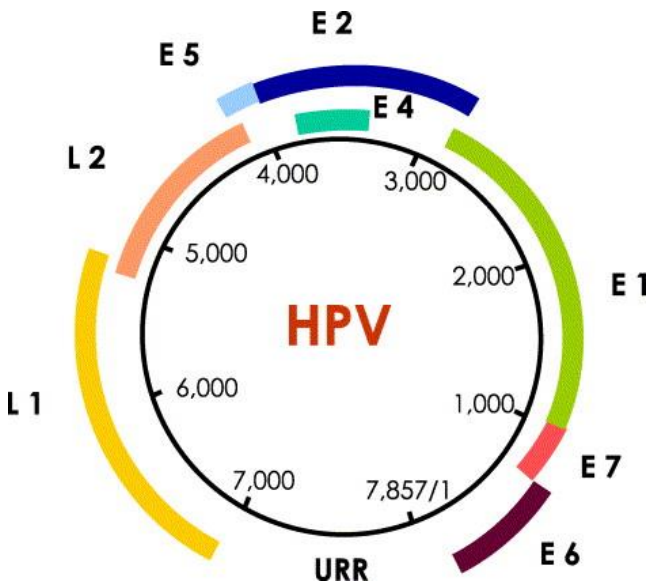

Figure 3. HPV genome structure. HPV genome is represented as a 7857 base pair circular DNA molecule. The open reading frames (ORFs) for the 8 proteins encoded are represented in different colours. The blank space between L1 and E6 ORFs represents the upstream regulatory region [52].

In figure 4, it is represented the evolution of HPV infection. Through microabrasions, the virus is able to penetrate the epithelium [48]. The expression of proteins E1 and E2, upon infection of undifferentiated basal cells, is responsible for the regulation of viral replication and expression of the other early stage proteins [54]. Hereafter, the expression of oncoproteins E5, E6 and E7 leads to the uncontrolled cell proliferation, cell survival and keratinocyte cell

differentiation. In cancers associated to HPV, E6 and E7 are responsible for the cancer phenotype [53, 55]. L1 and L2 are components of the viral capsid, allowing the assembling of the virus particles, which will be released within the superficial epithelial cells [48].

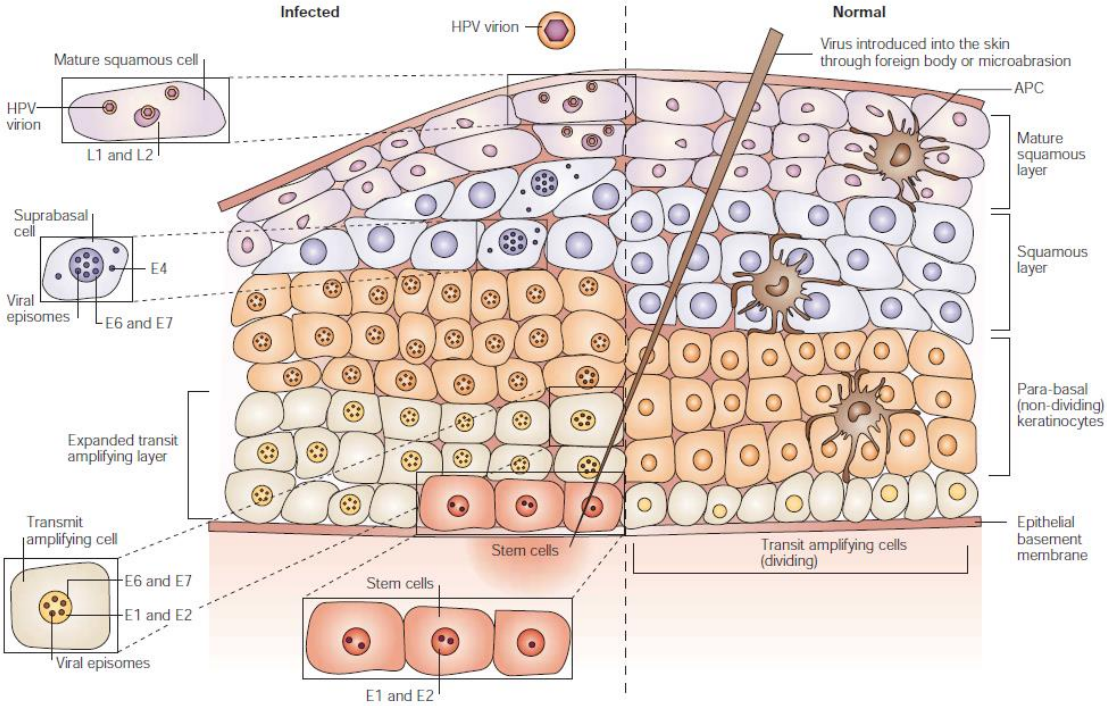

Figure 4. Evolution of infection by HPV [48].

1.2.2 Papillomavirus E6 oncoprotein

Papillomavirus E6 oncoprotein is a 151 amino acid protein characterized by two zinc finger domains, with a molecular weight of 18 kDa [56, 57]. The main action of this oncoprotein in the cellular cycle of the infected cells involves the ability of E6 protein to interfere with tumour suppressor p53, hindering the p53 protein capacity to trigger signalization pathways for cell repair and apoptosis [58]. The disabling of p53 contributes to continuous replication of damaged DNA and survival of abnormal cells, which would usually be repaired when p53 is expressed in normal levels. There are several mechanisms by which E6 protein is able to incapacitate p53 protein role [55, 59]. E6 is known to bind to E3 ubiquitin ligase, commonly known as E6 Association Protein (E6AP), through LXXLL motifs. E6AP is unable to bind to p53 without E6 oncoprotein. The docking of such proteins leads to the recognition of p53, preventing the binding of p53 to specific sequences of DNA [56, 60].

In addition, the association between E6AP-E6 complex to p53, ubiquitinate p53 leading to its degradation by the proteasome, as shown on figure 5 [61]. On the other hand, E6 oncoprotein is able to affect p53 functions through hindering of transactivation of p53 responsive genes. Such mechanism may be seen, for instance, through the p300 interaction with E6. The

promotor p300 is capable of acetylation of p53 when cellular damage exists, leading to higher recognition of specific DNA sequences by p53. The binding of E6 to p300 obstructs p53 acetylation, assisting in the expression decrease of responsive p53 genes [59, 60]. High risk HPV E6 proteins show increased affinity to p53, resulting in greater levels of inhibition, therefore contributing to a tumorous phenotype unlike low risk HPVs [59].

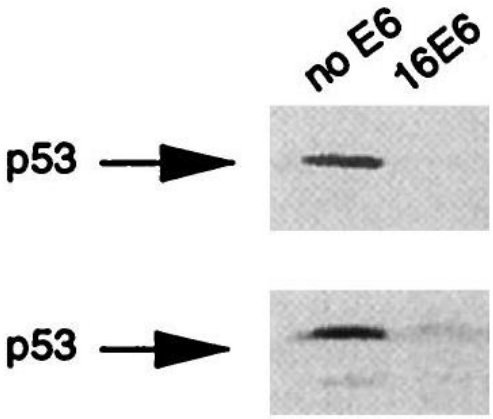

Figure 5. Degradation of p53 in the presence of HPV 16 E6 protein (adapted from [61]).

1.2.3 Papillomavirus E7 oncoprotein

Papillomavirus E7 oncoprotein is a 98 amino acid protein with a molecular weight of 16-17 kDa, which comprises three functional conserved regions (CR) [57, 62]. CR1 is associated with transformation of cell cycle independent of the gene product of retinoblastoma tumour suppressor (pRB), while CR2 is linked to the transformation of cell cycle by binding to pRB through its LXCXE motif [57]. The main transforming pathway of E7 oncoprotein relies on the binding to pRB, disrupting its binding to transcription factors such as E2F. Usually, pRB has the ability to phosphorylate E2F, forming a pRB-E2F complex which is capable of interfering with the progression of the cellular cycle. When the binding of E7 to pRB occurs, E2F disengagement results in stimulation of E2F responsive genes, which are responsible of cellular cycle transition to S phase [54, 57]. CR1 and CR3 domains of E7 protein are involved with destabilization of p107 and p130, two proteins which are also able of forming complexes with E2F. The association of E7 protein to pRB, p107 and p130 leads to uncontrolled cell proliferation and progression to malignant transformation [57, 62]. Similarly to E6 and p53 binding, high risk HPV E7 proteins show increased affinity for pRB binding [54].

1.2.4 Preventive and therapeutic vaccination

The rising of awareness for prevention of HPV infection has been increasing in the past years. Screening tests, such as pap smears, have been implemented as a routine test, contributing to an early diagnose and a decrease in the cancer mortality [51]. Prophylactic vaccines have been developed for prevention of infection. Its functionality resides in the administration of viral-like proteins (VLPs) that shall mimic a natural HPV infection. These VLPs consist in HPV late protein L1, a constituent of the viral capsid [63]. So far, there are only two available vaccines commercialized for the prevention of HPV infection: Gardasil and Cervarix [51]. Up to now, efficacy of prophylactic vaccines has been point out to be high [63]. However, it is estimated that the sexual behaviour of patients undergoing vaccination has a major influence in the outcome of the vaccines efficacy [51, 54]. The risk of infection increases with the number of lifetime sexual partners of the patient. However, the main failure of the VLPs therapy is related with the vaccination of a person already infected with HPV, which shall not produce any kind of effect within the individual [54].

Therefore the investment in a therapeutic alternative is extremely necessary, which might be able to prevent and, at the same time, treat infection by HPV. Nowadays there are several ongoing investigation and clinical trials regarding therapeutic vaccination treatments for HPV infection, taking mostly advantage of the previously described DNA vaccine technology. Each research has focused on different antigens and different antigen modifications in order to produce the ideal therapeutic HPV vaccine. Nonetheless, it is undeniable the preference for choosing E6 and E7 as target antigens for immunization and treatment responses [18, 45, 46, 64-66]

1.3 Plasmid DNA technology

Plasmid DNA has been widely studied throughout the years for its use as a successful cloning tool and for its ability in the eukaryotic cell transfection and subsequent induction of proteins expression [67]. Its capacity for carrying DNA information and lack of toxicity led this biomolecule to be strongly exploited as a DNA-based therapy vector, currently resulting in the most popular non-viral technique used in DNA-based therapy clinical trials, as previously depicted in figure 2 [68]. This circular double-stranded DNA molecule has bacterial origin and its size varies until 120 kilo base pairs, allowing this vector to carry larger DNA amounts than viral systems [28, 67, 69, 70].

The pDNA can be presented in different isoforms, the supercoiled (sc), open circular (oc), linear (ln), oligomeric or denatured isoforms which can occur depending on DNA sequence and different environment conditions. However, the sc conformation is the most produced isoform by *Escherichia coli* (*E. Coli*) host. The increasing of temperature, for instance, may lead to the uncoiling of the double DNA helix, promoting the conversion of the sc pDNA isoform to other pDNA isoforms. Also, undamaged sc pDNA can be converted to other isoforms through random enzymatic cleavage [71]. The isoform with target biological function is the sc pDNA, which has been proven to provide greater transfection rates than other isoforms [72, 73]. Thus, the sc pDNA isolation from the non-effective pDNA topologies and from the host components, that may be unveiled as toxic, is extremely important.

1.3.1 Construction, production and primary isolation of pDNA

The construction of a recombinant plasmid was firstly reported by Stanley Cohen and colleagues in 1973, making use of restriction enzymes, which are able to recognize specific DNA sequences and consequently cutting it, and DNA ligase in order to re-attach such DNA sequences [2]. A pDNA vector shall comprise an origin of replication for efficient cloning of the molecule in the target bacterial host, a strong promoter for expression in eukaryotic cells, the therapeutic target gene, an antibiotic resistance gene and finally, a polyadenylation termination sequence. The origin of replication is necessary for the amplification of the pDNA molecule within the bacterial host. The introduction of a gene capable of inducing a specific-antibiotic resistance is necessary to guarantee the solely growth of transformed bacteria with the target pDNA. The polyadenylation termination sequence is important for the translated RNA protection from degradation[74]. In figure 6 it is schematically represented the plasmid which was chosen to be manipulated throughout the present work.

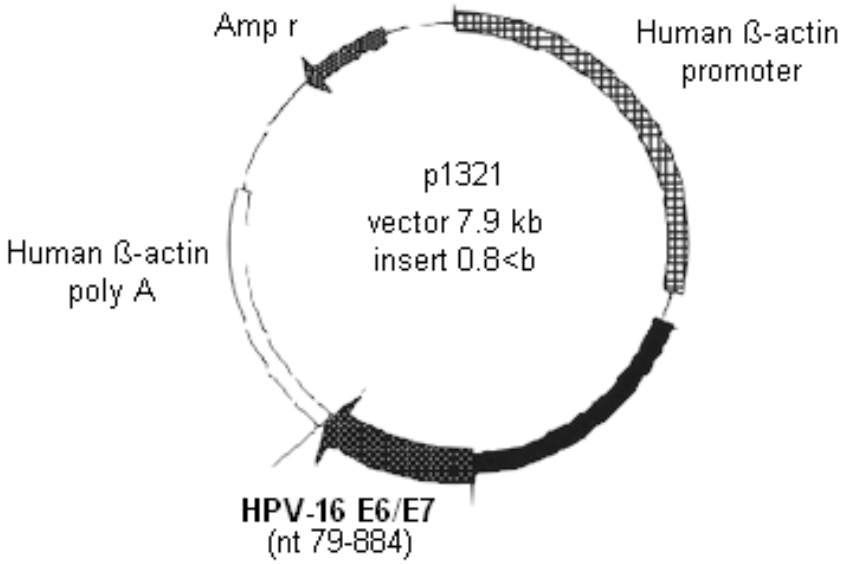

Figure 6. Schematic representation of HPV-16 E6/E7 pDNA. (plasmid 8641, adapted from [75]).

Fermentation allows the production of pDNA in transformed bacteria, by adjusting the growth conditions according to the chosen bacterial host, which is usually *E. coli* [67, 76]. Once the target optical density of transformed bacteria is achieved, it is necessary to proceed to the downstream processing of the cell lysate in order to obtain the biopharmaceutical pDNA.

Cell lysis should be performed through mechanical or chemical methods. The most used method is the alkaline lysis, which relies on the use of basic solutions capable of disrupting the cells membrane, releasing the pDNA and all cell components. Concentration and washing steps of the lysate sample can be performed to recover a clarified lysate to be further purified [77]. In order to achieve the suitable purity of the target biomolecule, the chromatography field has very much evolved in the past years, providing us a variety of techniques for such purpose. In figure 7 it is represented the main steps by which a target pDNA shall pass in order to accomplish a high pharmaceutical grade.

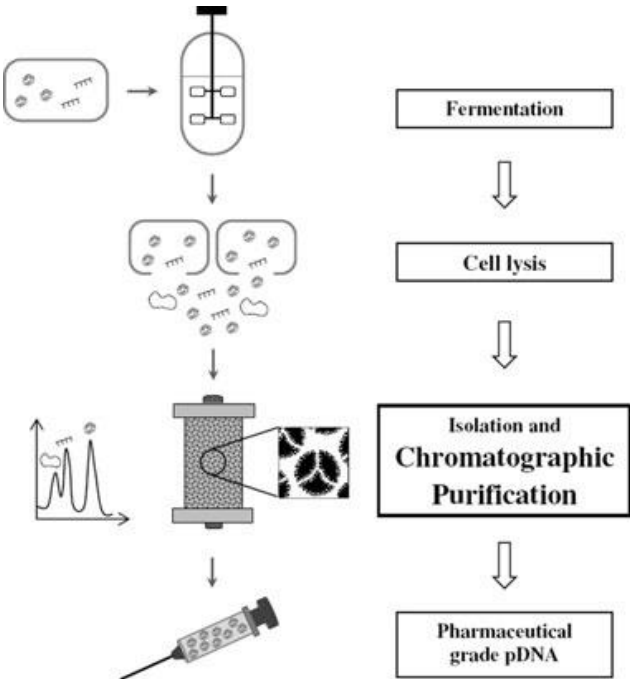

Figure 7. Production, recovery and purification of pDNA in order to achieve high pharmaceutical grade for application [69].

1.3.2 Plasmid DNA purification

Chromatography is an interesting area in biotechnology, providing methodologies and processes able to eliminate contaminants from a target molecule and purifying it. Impurities such as genomic DNA (gDNA), oc plasmid and RNA represent a hurdle to the sc pDNA purification that shall be surpassed through the aid of chromatography techniques. This methodology is very important to guarantee the compound safety for application. In matter of fact, purification of DNA vaccines has major significance and contributes to most of the expenses associated with vaccine development. In this way, the constant development of suitable chromatographic strategies to isolate the sc pDNA from a lysate sample in a single step is required, in order to reduce the costs associated to the DNA vaccine manufacturing. Table 3 represents the several specifications which pDNA must fulfil in order to be safely administrated. To achieve such specifications, several chromatographic techniques have been explored, taking advantage of biomolecule characteristics, namely size, charge, hydrophobicity or affinity between the different molecules present in the extract to be purified and the chromatographic ligands used [78].

Table 3. Specifications for pDNA to be considered as safe for administration (adapted from [78]).

Characteristics	Specifications
Appearance	Clear Colourless solution
Plasmid Homogeneity	>97% sc
Proteins	Not detectable, via BCA
RNA	Not detectable, via 0.8% agarose gel
gDNA	<2 µg/mg plasmid, via PCR
Endotoxins	<10EU/mg plasmid, via LAL essay

Size exclusion chromatography relies in the separation of different biomolecules according to their molecular size [79]. The greater the biomolecules are (as for instance the gDNA and pDNA) the less time it will take to pass through the support, since they do not enter to the pore particles, while smaller molecules like RNA will need more time [80, 81]. Nonetheless, such technique is unable to separate different isoforms of pDNA, showing poor selectivity and capacity for this molecule [76, 78].

On the other hand, anion exchange chromatography is known to fractionate the sample elements by charge differences. The principle behind this chromatographic process is based on the attraction of the target molecule to its opposite charge, immobilized into the stationary phase [81, 82]. Polymeric amines are usually the ligands used in anion exchanger chromatographic supports for pDNA purification [69]. After the sample binding, the salt concentration should gradually vary in order to elute the different molecules bound to the support. Nonetheless, some hurdles arise with the use of such technique due to its inability to distinguish different molecules with similar chemical and physical properties [81].

Hydrophobic interaction chromatography, also a very popular chromatography procedure, relies on the hydrophobic interaction between the target molecule and the chromatographic support. High concentrations of salt shall be applied for the sample binding, followed by a gradual decrease of ionic strength in order to separate the molecules with an increasing

hydrophobicity degree [76, 81]. Even so, the use of high salt concentrations represents a downside due to the high costs and the environmental impact associated [81].

Affinity interactions between target molecules and certain ligands have been object of intense study in the past years, leading to the development of affinity chromatography. Such technique is based on the specific interactions, as for instance molecular recognition, in order to selectively purify the target biomolecule [76]. The type of ligand used in this procedure should be chosen according to the biomolecule function or chemical structure [81]. The main drawback pointed to this chromatographic principle is the strong binding of the target molecule to the support, sometimes making it difficult to elute [83]. Within affinity chromatography, different strategies may be distinguished for the purification of pDNA, namely, immobilized metals affinity chromatography, triple-helix affinity chromatography, protein-DNA and amino acid-DNA affinity chromatography.

Immobilized metals affinity chromatography (IMAC) is based on the fractionation of molecules according to their affinity with metallic ligands of the stationary phase [76, 81]. RNA and single-stranded nucleotides show high affinity for the IMAC matrices, while gDNA, pDNA and other double-stranded nucleic acids have very low binding affinity [76, 81, 84]. This technique in particular can be used for polishing purposes, eliminating RNA and endotoxin impurities, but it has no significant importance in pDNA isoforms separation [76, 81].

On the other hand, triple-helix affinity chromatography (THAC) is based on the recognition of specific sequences of DNA by immobilized oligonucleotides, forming a triple-helix at acidic pH [76, 81, 85]. In comparison to IMAC, THAC has the advantage of being able to successfully fractionate different DNA isoforms, isolating the sc pDNA in one single chromatographic step. However, long chromatographic runs and low yields, in part due to the stronger retention of sc pDNA, suggest this might not be the ideal procedure for sc pDNA purification [81].

On the other side, amino acid-DNA affinity chromatography allows multiple interaction between the target DNA and the immobilized amino acid ligands, natural compounds whose chromatographic use provides an efficient single step chromatographic purification of sc pDNA and successful elimination of RNA, gDNA and other possible contaminants [76, 81]. Our research group has focused the purification of pharmaceutical-grade pDNA to be therapeutically applied, through the use of histidine, lysine and arginine amino acids as affinity ligands.

For instance, the sc pDNA purification with histidine and arginine agarose matrices suggested the involvement of different interactions between each amino acid and the different nucleic acids, revealing a selective retention according to the oligonucleotide bases sequence in the target biomolecule [86, 87]. Moreover, the isolation of sc pVAX1-*LacZ* from clarified cell lysate was successfully achieved, through the use of a histidine-agarose matrix, with an overall yield of 40% [88]. Likewise, through the use of lysine-agarose matrix, the sc pVAX1-*LacZ* purification was achieved with an overall yield of 45% [89]. Although the high-grade

purity attained with this two amino acids, the recovery yields were relatively low. Meanwhile, the use of arginine-agarose to purify the sc pVAX1-*LacZ* revealed a recovery yield of 79%, with the recommended purity degree [90].

Overall, the arginine agarose chromatography seems to be a good strategy to purify sc pDNA, owing to the selectivity achieved with this ligand under mild elution conditions and the satisfactory recovery yield. However, conventional stationary phases present some limitations related with work at high flow rates and low binding capacity for large biomolecules, such as pDNA, therefore leading to the exploration of alternative supports.

1.3.3.1 Monoliths: a new chromatographic support technology

Monoliths are chromatographic supports that have been gaining attention in the past years due to attractive characteristics, which are responsible for the outstanding behaviour of this new technology in comparison with conventional stationary phases. Such innovative stationary phases consist in continuous beds, resembling several stacked membrane sheets [69]. These supports are highly porous, whose sizes are dependent of the polymerization temperature, and comprise an interconnected three-dimensional network, allowing that all channel surface to be convectively accessible [91]. Such structural characteristics result in faster mass transfer, higher access of target molecule to the total monolithic structure and higher working flow rates [69]. The greatest advantage of monolith utilization relies on the combination of the convective flow with its high interconnectivity, allowing a swift separation in short beds [92]. The use of monoliths in sc pDNA purification has great value, considering conventional supports limitations [69].

Taking into account the high specificity and selectivity of the arginine ligand by the sc pDNA, as well as the high versatility and capacity of the monolithic supports, the coupling of both strategies, by the immobilization of arginine ligand into the monolith, emerged as a promising solution for the pDNA vaccine purification [93]. Thus, the arginine monolith was able to purify sc HPV-16 E6/E7 pDNA through amino acid-DNA affinity chromatography (as it is schematized in figure 8). Although 100% of purity was achieved, only 39% of sc HPV-16 E6/E7 pDNA was recovered [94]. To improve such recovery yield, maintaining the high purity degree, the experimental design tool could be used.

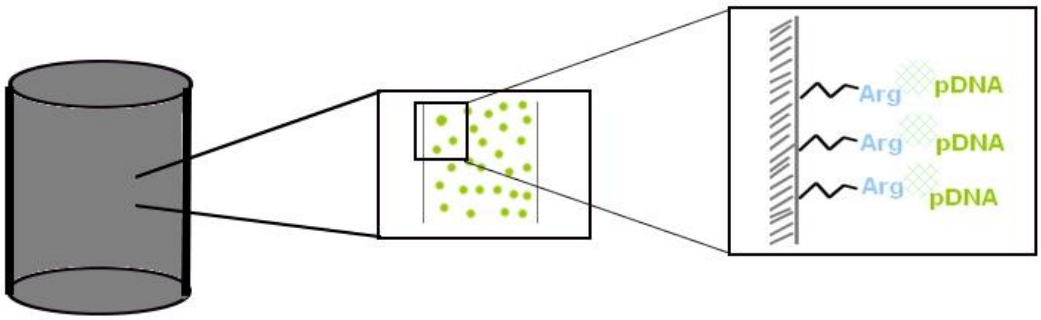

Figure 8. Schematic representation of arginine modified monolithic support.

1.3.4 Design of experiments

To attain a high pDNA yield and purity, some tests of screening, optimization and robustness should be performed. Such achievement might be difficult by the usual 'one by one' method, in which factors are separately tested, revealing to be time-consuming and exhaustive. Design of Experiments, a statistical technique which allows to plan, conduct, analyse and interpret data from experiments, arises as a solution to such demanded procedure, permitting to evaluate different factors through the use of a small amount of experiments [95, 96]. The user usually chooses the different factors to take in consideration, as well as the range in which values should vary. This technique has different designs available that can be chosen accordingly to the final purpose of the experimental design, which will suggest different experiments that must be performed. This technique emerges as a useful tool to help investigators to achieve better results with few experiments and can be used in miscellaneous areas.

In the chromatography field, the application of experimental design is usually associated with analytical chromatography, particularly in high performance liquid chromatography and gas chromatography [95]. Nonetheless, our research group recently published the use of DoE in preparative chromatography, accomplishing the purification of pre-miR-29 with high purity and considerable yield through an O-phospho-l-tyrosine affinity chromatographic support [97].

For optimization purposes, there are several designs which can be chosen in order to achieve the target result. For instance, Central Composite Face (CCF), Box–Behnken or Doehlert are experimental design tools used with the intent of optimizing a given experiment. Central composite designs are usually the preferred strategy when choosing an optimization design, even though it requires more experiences than other optimization designs such as Box-Behnken or Doehlert designs [95]. CCF design can be a good choice to optimize a purification strategy, because this method provides a wider perspective of the optimum gap, when

compared for example to Box-Behnken, which excludes experiments that meet every higher or every lower level for all factors [96].

In figure 9 it is schematically represented the CCF design. Each dot represents a combination of different values of the different factors for one experiment. The black dots in the cube vertices represent the factorial design points in evaluation, while the centre black dots represent the star points, which in this design lie within the faces of the factorial design points [95]. In red is represented the centre point, which has to be replicated in order to provide a measure of pure error and stabilize the variance of the predicted response [96].

.

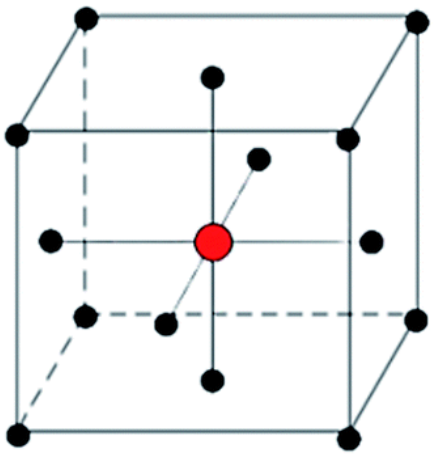

Figure 9. Schematic representation of CCF design (adapted from [98]).

CCF design is preferably used for optimization of experiments, which uses a full quadratic model that can be generally described as (n) = 2k–p + 2k + cp, where 'k' is the number of factors to be considered, 'p' the fractionalization number and 'cp' the number of centre points required for curvature estimation [99].

1.3.5 Plasmid application

Once the pDNA purity is guaranteed, further testing is necessary in order to assess the DNA vaccines success in transfecting the host cells, expressing the target gene and developing an immune response. With this in mind, several *in vitro* and *in vivo* procedures might be useful in assessing such information. Western blot and immunocytochemistry, for instance, are *in vitro* procedures which allow the evaluation of the transfection efficiency and target gene expression. Western blot technique allows to stain and quantify target proteins from a whole protein extract, while immunocytochemistry can be used to guarantee the pDNA entrance in the host cells by tracking the expression of the target gene, as well as quantifying it [100,

101]. Immunocytochemistry can also be used to monitor the nuclear structures, in order to verify if the transfection materials used are damaging the cells. On the other hand, *in vivo* studies are presented as the following step for pDNA application studies, permitting to verify the DNA vaccine ability to induce an immune response directed to the target antigens. Procedures such as flow cytometry for targeting of specific cytokines, serve as an example of studies that might be used to assess such application in in vivo experiments [102].

In this work, western blot and immunocytochemistry procedures were performed in order to assess the sc pDNA HPV-16 E6/E7 ability to transfect and produce the E6 and E7 target genes, also comparing its biological activity with other plasmid isoforms or plasmids prepared by using different purification methodologies.

CHAPTER 2 – Global Aims

The main objective of this work is to obtain a successful DNA vaccine for prevention and treatment of HPV infections. Such goal can only be reached if purification of the therapeutic molecule is achieved, considering the safety concerns regarding its administration. In order to accomplish this aim, the optimization of the implemented purification strategy for sc pDNA HPV16 E6/E7 has extreme importance, maximizing the performance of the purification process. In addition, the assessment of the purified target molecule ability for transfection is important, to evaluate if the DNA vaccine is able to perform its expected purpose. Hence, the goals of the present thesis are the optimization of sc pDNA HPV-16 E6/E7 purification by arginine monolith and evaluation of the transfection efficiency and gene expression ability of the purified molecule.

Therefore, the present thesis can be divided in two different steps:

- Establishment of an optimized purification strategy for sc HPV-16 E6/E7 by arginine monolith, in order to increase the recovery percentage of the target molecule, making use of design of experiments statistical tool.

- Transfection studies with purified sc HPV-16 E6/E7 pDNA molecule, as well as with the commercially obtained pDNA and oc pDNA, with the intent of assessing the transfection efficiency and gene expression ability of the different molecules.

CHAPTER 3 - Materials and methods

3.1 Production

Addgene plasmid 8641 [53], corresponding to HPV -16 E6/E7 DNA vector with 8.702 kbp, was obtained from Addgene (Cambridge, MA, USA). This vector is able to encode human E6 and E7 proteins of HPV-16 and contains the human beta-actin mammalian expression promoter and the ampicillin resistance gene.

HPV-16 E6/E7 plasmid was amplified through *E. coli* DH5α fermentation. Bacterial growth was performed at 37 ºC and 250 rpm shaking and the medium used was Terrific Broth (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. Growth was suspended at late log phase (OD_{600}~8), cells were recovered by centrifugation, at 4500 g for 10 min at 4 ºC, and the cell pellets were stored at -20 ºC.

3.2 Purification

3.2.1 Alkaline lysis with NZYTech kit

Plasmid DNA was recovered by alkaline lysis and purified through the use of NZYMaxiprep kit (NZYTech, Lda. - Genes and Enzymes, Lisbon, Portugal), according to manufacturer's instructions. After cells lysis, the purification starts with the cellular debris elimination, followed by pDNA binding to anion-exchange columns under the appropriate low-salt and pH conditions. Removal of RNA, proteins and low-molecular-weight impurities was carried out by medium-salt washing, while pDNA elution occurs by the use of high-salt buffer. After elution, pDNA was concentrated with isopropanol precipitation. The final pDNA (oc + sc) sample was used as a control of the purification procedure in transfection experiments.

3.2.2 Modified alkaline lysis

In order to evaluate purification by chromatography, it is necessary to obtain a complex sample by performing cell lysis by the use of the modified alkaline method [103], as described by Diogo *et al.*, 2000 [104]. Therefore, 250 mL bacterial pellets were resuspended in 20 mL of solution I (50 mM glucose, 25 mM tris-hydroxymethil aminomethane (Tris) and 10 mM ethylene-diamine tetraacetic acid (EDTA) at pH 8.0). Alkaline lysis was then performed by adding 20 mL of solution II (200 mM NaOH and 1% (w/v) sodium dodecylsulfate (SDS)), followed by 5 minutes of incubation at room temperature. In order to neutralize solution II, 20 mL of solution III (3 M potassium acetate at pH 5.0) were added, followed by 20 minutes of

incubation on ice. Elimination of cellular debris, gDNA and proteins was carried out through centrifuging twice at 20 000g (30 minutes, 4 ºC) with an Allegra™ 25R centrifuge (Beckman Coulter, Miami, FL, USA). The concentration and reduction of the impurity content was performed according to a previously published method [104]. Summing up, the nucleic acids present in supernatant were precipitated through the addition of 0.7 volumes of isopropanol, followed by 30 minutes of incubation on ice. Then, the precipitate was recovered by centrifugation at 16 000 g (30 minutes, 4 ºC). Afterwards, the pellet was resuspended in 4 mL of 10 mM Tris-HCl and 10 mM EDTA (Tris-EDTA) buffer, pH 8.0. Hereafter, proteins and RNA were precipitated by adding ammonium sulphate to pDNA solution up to a final concentration of 2.5 M, followed by 15 minutes of incubation on ice. Impurities were then removed by centrifugation at 16 000 g (20 minutes, 4 ºC). Finally, the supernatant was desalted by passing through PD-10 desalting columns (GE Healthcare, Buckinghamshire, UK), according to manufacturer's instructions, using mM Tris-EDTA (pH 8.0) as elution buffer. Afterwards, the resulting lysate sample was used in affinity chromatography.

3.2.3 Affinity chromatography

The equipment used to perform the chromatographic experiments was an AKTA Avant system (GE Healthcare, Buckinghamshire, UK), which consists in a compact separation unit and a personal computer with UNICORN™ 6.1 software. Arginine monolith, with a column bed volume of 0.34 mL and average pore sizes of 1500 nm prepared and provided by BIA Separations (Ajdovscina, Slovenia), was used according to the purification strategy previously optimized by our research group to isolate the sc HPV-16 E6/E7 pDNA [105]. Sodium chloride (NaCl) was purchased from Panreac (Barcelona, Spain) while Tris was purchased from Merck (Darmstadt, Germany). In each chromatographic experiment, all solutions were prepared through the use of freshly deionized water ultra-pure grade, purified with Milli-Q system (Millipore, Billerica, MA, USA), and analytical grade reagents. Filtration of elution buffers was performed with 0.20 µm pore size membrane (Schleicher & Schuell Bioscience GmbH, Dassel, Germany), followed by ultrasonic degassing. Briefly, the monolithic disk was firstly equilibrated with 600 mM NaCl in Tris-EDTA buffer (pH 8.0) at a flow rate of 1 mL/min. Lysate sample resulting from modified alkaline lysis was injected into the column using a 500 µL loop in the same flow conditions. Unbound species were immediately eluted and the ionic strength was then increased to 740 mM NaCl in Tris-EDTA buffer (pH 8.0) and then 1 M NaCl in Tris-EDTA buffer (pH 8.0) using a stepwise gradient, so that the different bound species could gradually elute. All the experiments were carried out at room temperature and the absorbance was constantly monitored at 260 nm. Fractions were recovered, according to obtained chromatograms, concentrated and desalted with Vivaspin® 6 Centrifugal Concentrator (Vivaproducts, Littleton, MA, USA), at 800 g in the SIGMA 3-18K centrifuge (SIGMA Laborzentrifugen GmbH, Germany), for further electrophoretic analysis. In the end of

the chromatographic experiments, the column and AKTA Purifier system were cleaned with MiliQ water.

3.2.3.1 Agarose gel electrophoresis

Horizontal electrophoresis was carried out with pooled fractions of each chromatographic experiment, using 15 cm long 0.8% agarose gel (Hoefer, Holliston, MA, USA) stained with 0.016 µL/mL Greensafe premium (NZYTech, Lda. - Genes and Enzymes, Lisbon, Portugal). Electrophoresis was performed at 110 V for 30 minutes in TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA, pH 8.0). The gel was later visualized under UV light in a FireReader (Uvitec Cambridge, Cambridge, UK).

3.2.4 Design of experiments

In order to improve the purity and yield of sc HPV-16 E6/E7 pDNA purification, experimental design was performed. In this design, three factors (inputs) were evaluated, namely, binding concentration, washing concentration and solutions pH. Recovery yield and purity (outputs) were the evaluated responses. The column was equilibrated within the range of 500 mM and 600 mM NaCl in Tris-EDTA buffer and washing step was performed within the range of 650 mM and 750 mM NaCl in Tris-EDTA buffer. The solutions were prepared with a pH variation between 7.5 and 8.0. Elution of sc HPV-16 E6/E7 pDNA would be performed through the increasing of ionic strength to 1M Tris-EDTA buffer. Each peak was pooled, desalted and concentrated to a final volume of 300 µL with Vivaspin® 6 Centrifugal Concentrator (Vivaproducts, Littleton, MA, USA), at 800 g in SIGMA 3-18K centrifuge (SIGMA Laborzentrifugen GmbH, Germany). The resulting purified sc HPV-16 E6/E7 plasmid was later used for *in vitro* transfection studies. Statistical analysis was performed through the use of UNICORN™ 6.1 software and Design Expert version 9 trial software.

3.2.4.1 Supercoiled pDNA quantification

The evaluation of sc pDNA purity and recovery was accomplished through the application of a modified quantification method by using the CIMac™ pDNA analytical column [106, 107]. Briefly, a calibration curve was performed with pDNA standards of 1 to 25 µg/mL, as it is presented in Figure 10. The pDNA standards were prepared through dilution of the highest concentration of pure pDNA with 200 mM Tris-HCl (pH 8.0) and the concentration was confirmed by Ultrospec 3000 UV/Visible Spectrophotometer (Pharmacia Biotech, Cambridge, England). CIMac™ pDNA analytical column was equilibrated with 600 mM NaCl in 200 mM Tris-HCl (pH 8.0). After injection of 20 µL of plasmid sample at 1 mL/min a linear gradient from 600mM to 700 mM NaCl in 200 mM Tris was applied for 10 minutes, leading to elution of all

DNA species. The areas from the different peaks of different DNA species allowed the calculation of purity and recovery.

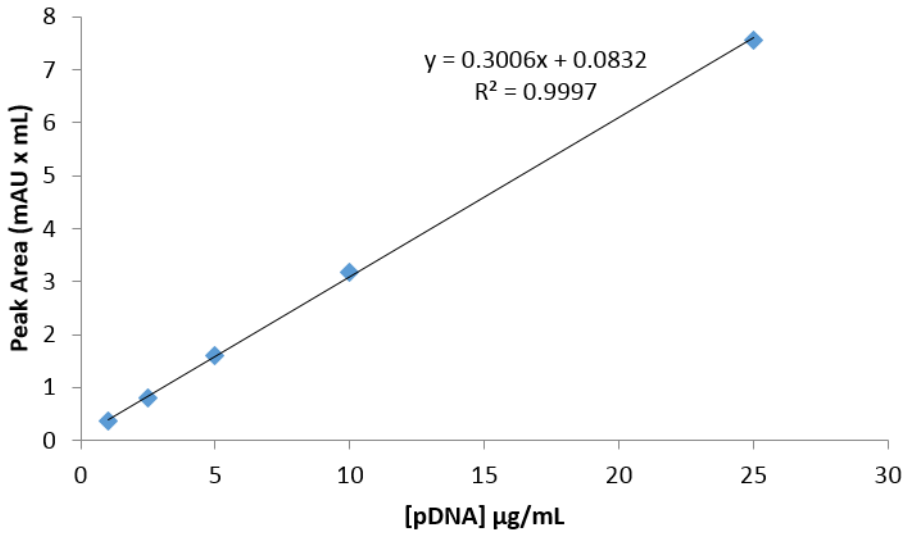

Figure 10. Calibration curve with pDNA standards (1-25 µg/mL).

3.3 Cell culture

CHO-1 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM)-F12 High Glucose Medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% (v:v) of fetal bovine serum (FBS) and a mixture of penicillin (100 µg/mL) and streptomycin (100 µg/mL). All cells were grown at 37 ºC in a humidified atmosphere with 5% of CO_2 in air.

3.3.1 Transfection

The purpose of the transfection protocol was to verify the expression of E6 and E7 proteins with different pDNA isoforms and preparation methods of HPV-16 E6/E7 plasmid, in order to identify the most effective procedure. Thus, the transfection protocol was performed with K2® Transfection System (Biontex, Martinsried/Planegg, Germany) according to manufacturer's instructions, in order to transfect CHO-1 cells with sc HPV-16 E6/E7 plasmid obtained by affinity chromatography and native pDNA (sc+ and oc) HPV-16 E6/E7 plasmid obtained by NZYTech kit. Summing up, CHO-1 cell line was cultured until a confluence of 80% to 90% is obtained. Then, complete medium was substituted by medium without antibiotic and FBS and cells were incubated for 24 hours. Thereafter, K2 transfection reagent and different plasmid samples were incubated at room temperature for 20 minutes and added to the 24-wells, followed by 24 hours of incubation. Hereafter, the cell line medium without

antibiotic and FBS was replaced with the respective medium with antibiotic, followed by 72 hours of incubation. Protein extraction and immunocytochemistry were performed with transfected cells. Cells intended to be used for protein extraction were cultured in 6-wells, while cells planned to be used for immunocytochemistry were cultured in 24-wells with 10 mm coverslips.

3.3.2 Protein extraction

Transfected and control cells were also submitted to protein extraction. Briefly, cells were scrapped with cell scrapper and phosphate buffered saline (PBS) and divided in microtubes. Then, the microtubes were centrifuged at 11500 rpm (7 minutes, 4ºC) in a Hettich Mikro 200R centrifuge (Andreas Hettich GmbH & Co. KG, Tuttlingen, Germany). The supernatant was discarded and 500 µL of PBS were added, followed by centrifugation in the same conditions. Afterwards, supernatant was discarded again and complete lysis medium was added according to pellets volume, followed by 10 minutes of incubation on ice. Complete lysis medium consists in 25 mM Tris (pH 7.4), 2.5 mM EDTA, 2% Triton X-100, 2.5 mM EGTA, 1 mM PMSF, 10 µL/mL complete EDTA free protease inhibitor cocktail (Roche, Indianapolis, USA). The resulting proteins were used to perform western blot technique.

3.4 Western blot

3.4.1 Protein quantification

The Bradford micro-assay from BioRad (Hemel Hempstead, UK) was used to measure the protein concentration. In summary, triplets of mixing 1 µL of sample, 159 µL of MiliQ water and 40 µL of Biorad reagent were prepared in a microplate for each sample. Upon 15 minutes of incubation at room temperature, absorbance was recorded at 595 nm in an Anthos 2020 microplate reader (Biochrom, Cambridge, United Kingdom). Afterwards, a calibration curve was constructed, using Bovine Serum Albumin (BSA) from Sigma-Aldrich (St. Louis, MO, United States of America) as a standard protein (0.2-10 µg/µL), as depicted in figure 11.

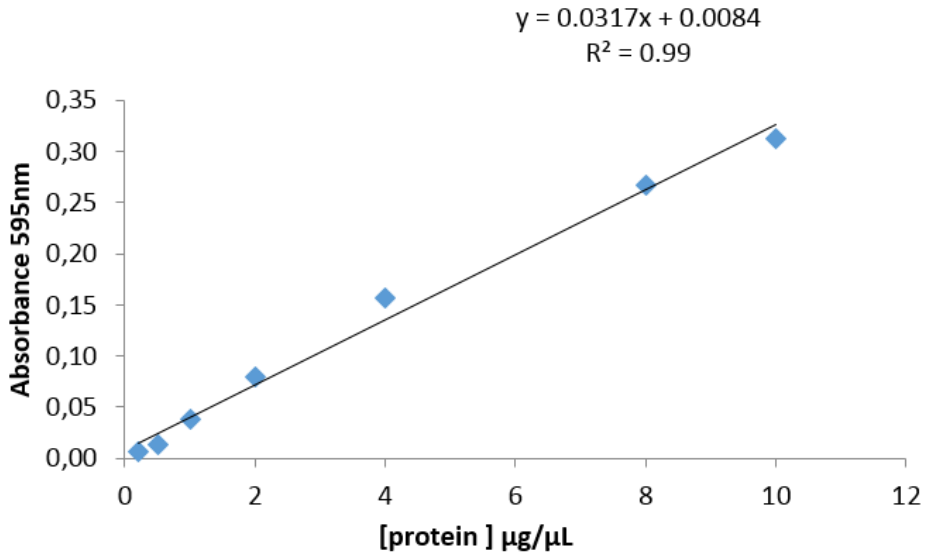

Figure 11. Calibration curve with protein standards (0.2-10 µg/µL).

3.4.2 Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis was performed with pooled fractions of proteins of transfected and control cells, using 12.5% polyacrylamide resolving gel and 4.7% polyacrylamide stacking gel. The samples were prepared with 50 µg of each protein sample, whose volume was calculated through the calibration curve constructed by protein quantification, and 5 µL of loading buffer, followed by 15 minutes of denaturing at 100 ºC. The molecular weight marker used was NZYColour Protein Marker II (NZYTech, Lda. - Genes and Enzymes, Lisbon, Portugal). Afterwards the electrophoresis was carried out at 120 V, for 90 to 120 minutes. The resulting electrophoresis gel was visualized or used for electroblotting.

3.4.3 Electroblotting

Protein gel electrotransfer is based on the transference of total proteins of an electrophoresis gel to a polyvinylidenedifluoride membrane. Electroblotting was carried out through a Cleaver system with protein gel resulting from previous electrophoresis. Membrane activation was performed with sheer methanol, followed by MiliQ water equilibration. Electrotransfer was performed at 0.75 mA and 250 V for 45 minutes. Afterwards, the membrane was washed 3 times with Tris-buffered saline solution with 0.1% Tween 20 (TBS-T 0.1%) for 5 minutes, followed by 2 hours of incubation at room temperature with 5% BSA, under constant agitation.

Afterwards, the membrane was washed 3 times for 15 minutes, cut-out and incubated at 4 ºC, overnight, with mouse anti-ß-actin monoclonal antibody (Sigma-Aldrich, MO, United States of America), mouse anti-E6 IgG monoclonal antibody (Santa Cruz Biotechonology, Heidelberg, Germany) and mouse anti-E7 IgG monoclonal antibody (Santa Cruz Biotechnology, Heidelberg, Germany). Hereafter, the membranes were washed 3 times with TBS-T 0.1% for 15 minutes and incubated at room temperature with goat anti-mouse IgG polyclonal antibody (Santa Cruz Biotechnology Heidelberg, Germany) for 1 hour, followed by washing 3 times for 15 minutes with TBS-T 0.1%. Then, the membranes were incubated for 5 minutes with ECF (GE Healthcare, Buckinghamshire, United Kingdom) and were visualized in BioRad FX-Pro-plus (Bio-Rad, Hemel Hempstead, UK).

3.5 Immunocytochemistry

In order to perform immunocytochemistry it was necessary to place the coverslips previously to the seeding of cells in 24-wells. After 72 hours of transfection, the coverslips were gently washed 3 times with PBS, fixed with 4% paraformaldehyde for 15 minutes, washed again 3 times with PBS, and permeabilized with 1% Triton X-100 in PBS for 5 minutes. In order to eliminate non-specific binding, a blocking step must be performed with 20% FBS in PBS with 0.1% Tween at room temperature, followed by 3 times washing with PBS Tween-20 0.1% (PBS-T O.1%). Mouse anti-E6 IgG monoclonal antibody and mouse anti-E7 IgG monoclonal antibody with a dilution of 1:25, were added to each coverslip and incubated overnight at 4 ºC. After overnight incubation, the coverslips were washed 6 times with PBS-T 0.1%. Incubation of 1 hour at room temperature with secondary antibodies Alexa 546 goat anti-mouse (Life Technologies, CA, United States of America) to the cells marked with the E6 primary antibody and Alexa goat anti-mouse 488 (Life Technologies, CA, United States of America) to the cells marked with the E7 primary antibody, was executed. The coverslips were washed 3 times with PBS-T 0.1%, followed by 10 minutes incubation with Hoescht 33342 (Invitrogen, CA, United States of America) for nucleus staining. The coverslips were mounted with a drop of mounting medium and stored at 4 ºC or -20 ºC.

3.5.1 Immunofluorescence analysis

Coverslips were analysed using a Zeiss LSM 710 laser scanning confocal microscope (Carl Zeiss SMT, Inc., Oberkochen, Germany) and image achievement was performed with Zeiss Zen software (2010). Immunofluorescence analysis and image manipulation was performed with ImageJ 1.48u software. Statistical analysis was performed with one-way ANOVA and unpaired rat-student statistical tools by GraphPad Prism version 6.04 trial software.

CHAPTER 4 – Results and discussion

The increase of cancer diseases having bacterial and viral origins has led research studies to obtain adequate treatments for such matter. DNA vaccines arise as a potential solution for such cancer diseases, due to its ability to induce the development of targeted immune responses for bacterial and viral antigens. Purification strategies are essential to ensure the lack of toxicity of these DNA vaccines, eliminating impurities derived from the DNA vaccines bacterial production process.

HPV is an example of virus capable of inducing the formation of tumours upon infection through the production of oncogenic E6 and E7 proteins. Our research group has been able to produce and successfully purify the sc pDNA HPV-16 E6/E7 DNA vaccine. However, the recovery of the target molecule obtained from the arginine monolith was relatively low. Thus, the experimental design arises as a useful statistical tool which provides several designs for experiments optimization, helping investigators to achieve optimum conditions to maximize a target response surface.

In the present work, the optimization of the sc pDNA HPV-16 E6/E7 purification by arginine monolith was performed through the use of design of experiments tool, in order to attain a higher recovery yield while maintaining 100% purity. Afterwards, *in vitro* transfection studies were performed in order to assess the ability of the purified pDNA vaccine for transfection and expression of target proteins, comparing such ability with other pDNA isoforms obtained by different purification strategies.

4.1 Experimental design

4.1.1 Preliminary tests for threshold choice

For the experimental design to be performed, there are a few issues that have to be delineated beforehand, namely, the choice of factors to be considered in the design of experiments. Preliminary studies, represented in figure 12, showed that purification of the sample was able to be achieved with 600 mM of NaCl for binding step, 740 mM of NaCl for washing step and 1 M for elution step. The sample chromatographic behaviour is different from the one previously published [94], eluting at 1 M of NaCl, instead of 740 mM of NaCl, without the presence of RNA species. This behaviour is probably due to the absence of different RNA species resulting from pDNA primary isolation.

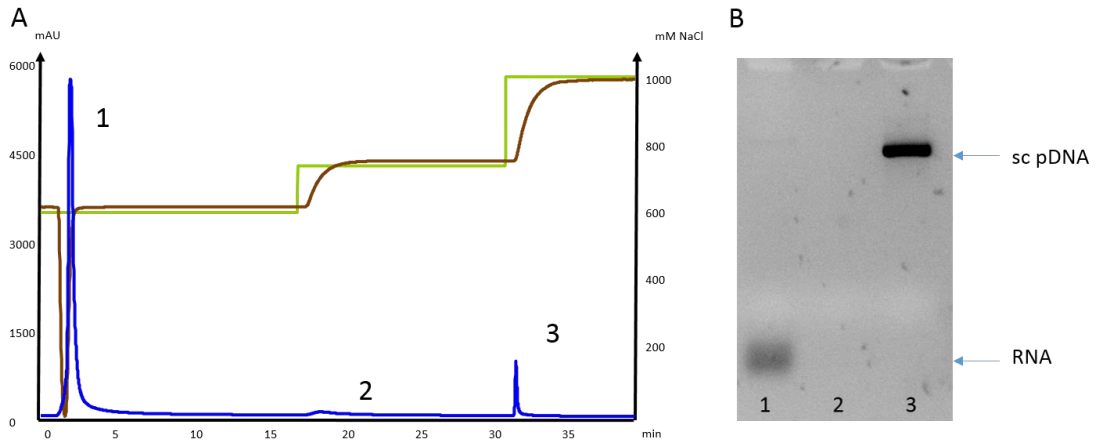

Figure 12. A) Chromatographic profile of the *E. coli* clarified lysate sample injected in the arginine monolith. Chromatographic run performed with 600 mM NaCl in Tris-EDTA (pH = 8.0) for binding step, 740 mM NaCl in Tris-EDTA (pH = 8.0) for washing step and 1 M NaCl in Tris-EDTA (pH = 8.0) for elution step. B) Agarose gel electrophoresis of the different peaks obtained. Lane 1 – Binding peak; lane 2 – Clarification peak; lane 3 – Elution peak.

Therefore the chromatographic runs were intended to be performed in 3 different steps: binding, washing and elution steps. The baseline for binding step comprises the ionic strength necessary to guarantee the efficient binding of the target molecule to the chromatographic matrix, while eliminating great part of impurities present in the clarified lysate sample, mainly RNA, as visible in lane 1 of figure 12 (B). The baseline for washing step consists as an intermediate chromatographic step, which uses a higher ionic strength than the previous step, being responsible for the elimination of small impurities bound in the first step (such as RNA and oc pDNA) but maintaining the sc pDNA retained. This baseline is at the same time a threshold for the recovery of the target molecule, because a slight increase of ionic strength can be enough for the partial or total elution of the target sample. The baseline for the elution step relies in a high ionic strength buffer in order to totally elute and recover the target molecule.

The factors firstly chosen to be evaluated by design of experiments were the binding step, the washing step, flow rate, and the pH.

Taking into account the preliminary studies previously described (figure 12), the ranges for binding and washing were chosen in order to comprise the concentrations of NaCl used to purify the sc pDNA. Thus, the range for binding step optimization was delineated as 500-600 mM, while the range for optimization of washing step was defined as 650-750 mM.

In order to guarantee the relevance in varying the flow rate, several preliminary chromatographic runs were also performed. In figure 13 it is represented the effect of flow rate variation in the sc pDNA purification with the arginine monolith. Both experiments were executed using the same pH, binding, washing and elution conditions (pH 8.0, 550 mM of NaCl in Tris-EDTA for binding step, 750 mM of NaCl in Tris-EDTA for washing step and 1M of NaCl in

Tris-EDTA for elution step). The flow rate used in the first run was 1.5 mL/min (figure 13 (A)) and in the second run was 2 mL/min (figure 13 (B)). However, as shown in figure 13, there were no significant differences between both chromatographic runs, suggesting that the flow rate has no relevant influence in the sc pDNA purification. These results are in agreement with previous publications about the monolith functionality, revealing purification strategies without influence of flow rate [69, 91, 108]. Therefore, the obtained data led to the disposal of this factor from experimental design.

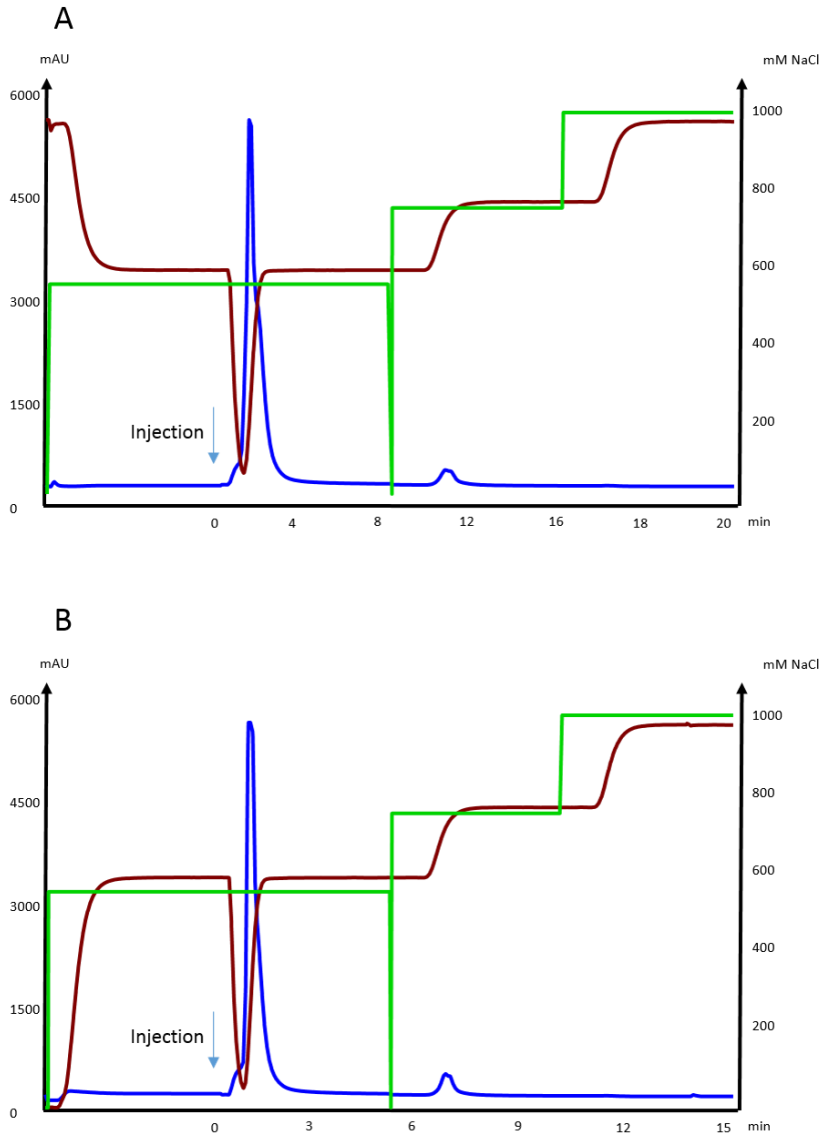

Figure 13. Evaluation of the flow rate effect in sc pDNA purification. Injected sample: *E. coli* clarified lysate sample. Both chromatographic runs were performed with 550 mM NaCl in Tris-EDTA (pH = 8.0) for binding step, 750 mM NaCl in Tris-EDTA (pH = 8.0) for washing step and 1 M NaCl in Tris-EDTA (pH = 8.0) for elution step. Different flow rates were used A) 1.5 mL/min; B) 2mL/min.

Finally, the selection of the adequate range for the pH factor was also performed by several preliminary experiments with different pH values. Given that the double-stranded DNA is known to be more stable within the range of pH 4.0 to pH 9.5 [109], the selected range for pH testing was pH 6.0 to pH 9.0. Hence, five chromatographic runs with five different pH values were executed, as depicted in figure 14. The flow rate and NaCl concentration used in the experiments is portrayed in table 4. As discussed before, different flow rates had no effect on the chromatographic profile, or in the recovery and purity of the target molecule.

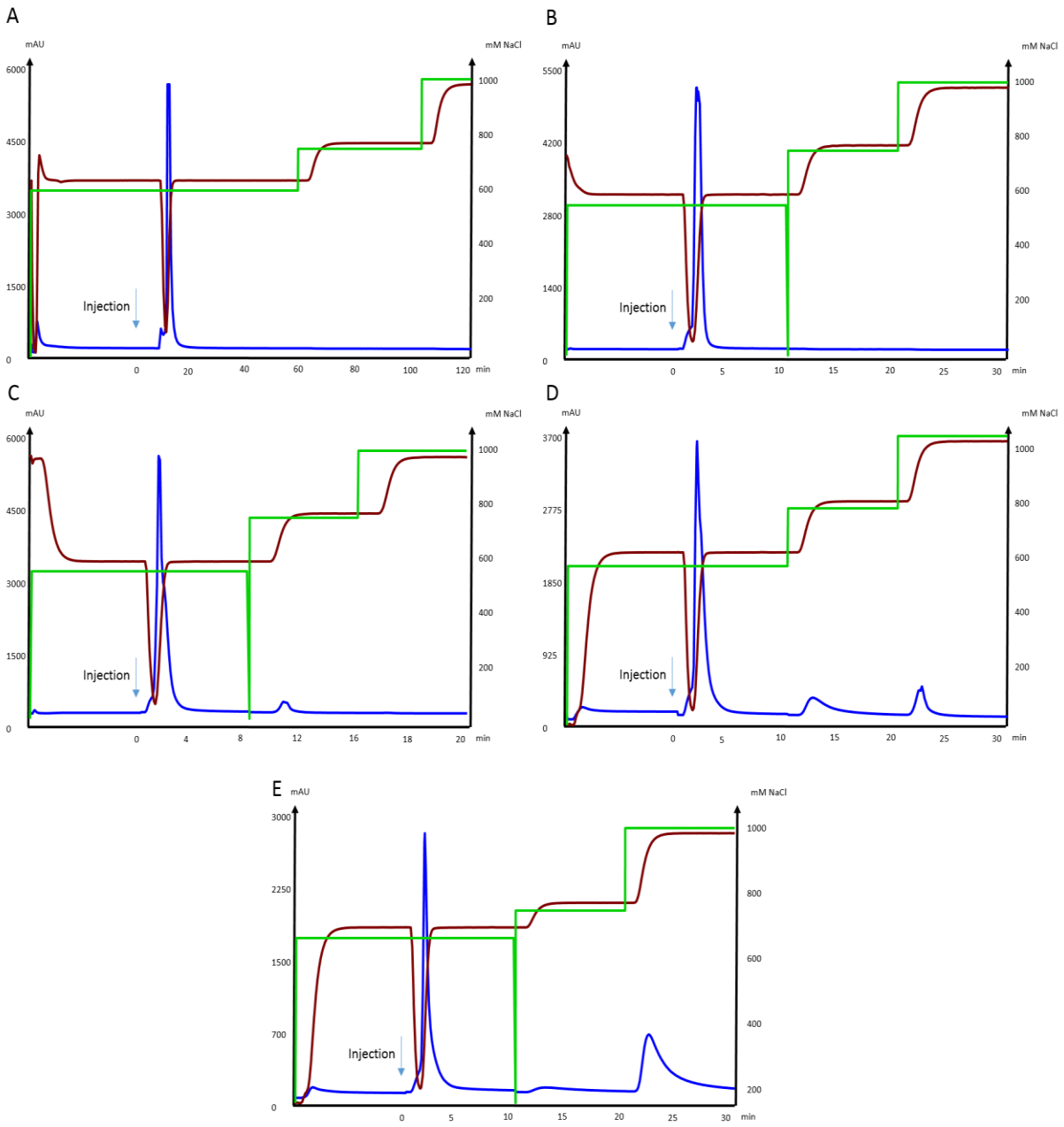

Figure 14. pH influence on the sc pDNA purification with the arginine monolith. Chromatographic conditions are presented in table 4.

At pH 9.0 and pH 8.5 (figure 14 (A) and (B)) it was observed the total elution of the injected sample at the binding step. At pH 6 (figure 14 (E)) it was observed a stronger binding of pDNA to arginine monolith, being more retained and some species only eluted at higher salt concentration. Taking into consideration that chromatographic runs at pH 6.0 and pH 9.0 were performed with the same binding condition (table 4), it is possible to confirm that pH variations have some effects in the retention and elution of the target molecule. Given that the pKa of arginine is 12 [110], the lower working pH will increase the effectiveness of the arginine positive charges, leading to a stronger interaction with pDNA. Thus, the strong retention of pDNA at pH 6 can be mainly related with the high positive charge density of the arginine ligand, not disregarding the involvement of other non-covalent multiple interactions.

Table 4. Chromatographic conditions for pH range assessment.

Chromatogram	pH	Flow rate (mL/min)	NaCl Concentration (mM in Tris-EDTA buffer)		
			Binding step	Washing step	Elution step
A	9	0.2	600	750	1000
B	8.5	1	550	750	1000
C	8	1.5	550	750	1000
D	7.5	1	550	750	1000
E	6	1	600	700	1000

These data also suggest that beyond the effect of the ligand charge, the additional influence of the ions negatively charged can promote nucleic acid elution by competition [111]. When the pH is higher than 7, there are more OH- ions present in the solution, which can compete with pDNA negative charge and bind to arginine, allowing the pDNA elution with lower NaCl concentration due to the cumulative competition effect from the Cl^- ions. Otherwise, when the pH is lower than pH 7, there are less OH- ions present in the solution, contributing to a stronger binding of the pDNA to the arginine ligand, being necessary higher NaCl concentrations to promote the pDNA elution.

Therefore, according to previous results of the pH effect (pH 9, 8.5 and 6) on pDNA retention to the arginine monolith, the range for NaCl concentration in the binding and elution steps would have to vary considerably when compared to the established purifying strategy

(present in figure 12) to promote the retention or elution of pDNA at pH 8. For instance, a decrease of NaCl concentration for binding purposes is necessary in order to prevent the total elution of the pDNA species at pH higher than 8.5. The reverse behaviour should be used at pH lower than 7.

For chromatographic runs with the same binding NaCl concentration (550 mM) but pH 8 (figure 14 (C)) and pH 7.5 (figure 14 (D)), the pDNA was recovered in the washing and elution peaks, suggesting that this pH range can be easily manipulated in order to obtain the pDNA in the elution peak. Thus, the chosen pH range for design of experiments was between 7.5 and 8.

All samples were analysed with the CIMac™ pDNA analytical column in order to assess the purity degree and recovery yield, in special of the sc pDNA eluted in the third peak. In figure 15 it is represented three examples of different analysis obtained through the analytical column. Elution of the injected DNA species was performed at 1 mL/min with a 10 minutes application of a linear gradient from 600 mM to 700 mM NaCl in 200 mM Tris-HCl buffer (pH 8.0) [107].

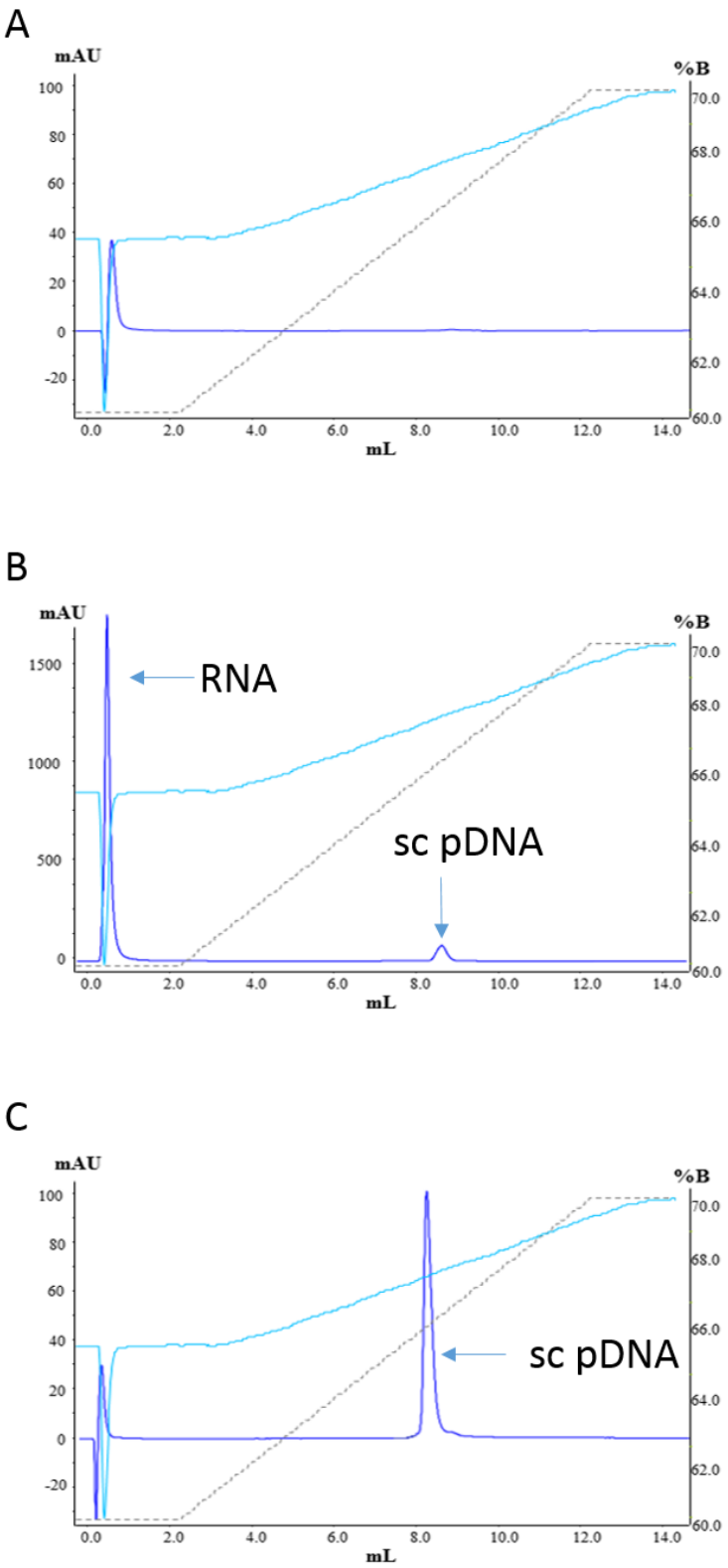

Figure 15. Example of the chromatographic profiles obtained in the analysis with the CIMacTM analytical column. All experiments were performed at 1 mL/min, during 10 min of linear gradient from 600 mM to 700 mM NaCl in 200 mM Tris-HCl pH 8.0. A) Blank sample; B) Lysate sample; C) Pure pDNA sample.

In figure 15 (A) is shown a blank sample, where only 200 mM Tris-HCl buffer (pH 8.0) was injected in order to simulate the effect of the ionic strength of the analysed samples in the elution profile. The peak depicted in the chromatogram occurred after the injection and corresponds to the difference of ionic strength between the loading buffer and the injected sample. Thus, the area from this peak shall be discounted when analysing pDNA samples. In figure 15 (B) is portrayed an example of lysate sample injected, where the first peak corresponds to RNA and the second peak corresponds to pDNA. In figure 15 (C) is visible an example of pure sc pDNA sample injection. Through the comparison of the areas of the injected sample and the calibration curve (1-25 µg/mL) it was possible to access the recovery and purity of the analysed peaks.

4.1.2 Optimization

Experimental design can provide several designs suitable for optimization purposes. In this work, CCF design was chosen over other optimization designs, such as Box-Behnken, due to providing a wider perspective of the optimum gap, even requiring more experiments [96].

Once the factors and their ranges were adequately chosen, CCF design was used in order to reach an optimization for this purification strategy. CCF design may work with three levels factors, in which the lower level -1 represents the minimum value for the defined range, while the higher level +1 represents the maximum value for the defined range. Intermediate level 0 represents the mean of the given range. In table 5 are represented the chosen factors, their ranges and their correspondent levels.

Table 5. Chosen factors versus levels of CCF design.

Levels		Factors		
		Binding (mM Na Cl in Tris-EDTA)	Washing (mM Na Cl in Tris-EDTA)	pH
Low	-1	500	650	7.50
Medium	0	550	700	7.75
High	1	600	750	8

After setting the range of each factor, table 6 represents the general combinations of the three levels factors for CCF design. Through the use of UNICORN™ 6.1 software, a matrix was designed with the different experiments to be performed, represented in table 7. In grey is represented the central point, which has be replicated twice in order to access the model's reproducibility.

Table 6. CCF design for three levels.

Experiment	Factors		
	A	B	C
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1
9	-1	0	0
10	+1	0	0
11	0	-1	0
12	0	+1	0
13	0	0	-1
14	0	0	+1
15	0	0	0

Table 7. Designed experiments according to CCF design. Binding and washing step represent NaCl concentration in Tris-EDTA buffer.

Experiment	Factors		
	Binding (mM)	Washing (mM)	pH
1	500	650	7.50
2	600	650	7.50
3	500	750	7.50
4	600	750	7.50
5	500	650	8
6	600	650	8
7	500	750	8
8	600	750	8
9	500	700	7.75
10	600	700	7.75
11	550	650	7.75
12	550	750	7.75
13	550	700	7.5
14	550	700	8
15	550	700	7.75

Therefore, considering the replicates, 17 experiments were randomly performed with three different gradient steps. First, binding of the target molecule and elimination of most of the impurities was performed, followed by washing, through the elimination of small impurities, and finally elution of the target molecule with 1 M of NaCl in Tris-EDTA buffer. Different pH buffers were used, correspondingly to the proposed value by experimental design.

As it was previously refereed, each eluted peak was pooled, concentrated, desalinized and analysed through CIMac™ analytical column for recovery and purity assessment, the chosen outputs for experimental design. In table 8 is represented the relative percentage of recovery and purity of the sc pDNA (responses) eluted in the third peak of each chromatographic experiment, resultant from the CCF design combinations.

Table 8. Responses of sc pDNA recovery and purity obtained for each run defined by CCF design. Binding and Clarification step represent NaCl concentration in Tris-EDTA buffer.

Experiment	Binding (mM)	Clarification (mM)	pH	% Recovery	% Purity
1	500	650	7.50	56.1	18.5
2	600	650	7.50	59.1	29.9
3	500	750	7.50	39.8	43.4
4	600	750	7.50	68.1	62.1
5	500	650	8	80	100
6	600	650	8	70.2	100
7	500	750	8	0	0
8	600	750	8	0	0
9	500	700	7.75	69.4	44.4
10	600	700	7.75	62.7	53.5
11	550	650	7.75	77.6	70.5
12	550	750	7.75	56.7	82.2
13	550	700	7.5	54.5	29.8
14	550	700	8	0	0
15	550	700	7.75	53.7	80.7
16	550	700	7.75	53	87.6
17	550	700	7.75	68	62.1

The resulting responses were introduced in UNICORN™ 6.1 software and Design Expert version 9 trial software for statistical analysis.

4.1.2.1 Goodness of fit

Goodness of fit data, depicted in table 9 and figure 16, allows understanding if the model used for optimization is good. Coefficient of determination, represented as R^2, explains if the model fits the data. R^2 should vary between 0 and 1, being ideal a percentage close to 1, indicating the data points fit the statistical model [112].

Table 9. Statistical coefficients of the model.

Response Surface	R^2	Adjusted R^2	Q^2	Model validity	Reproducibility
% Recovery	0.92	0.81	0.39	0.74	0.90
% Purity	0.83	0.62	0.03	0.67	0.85

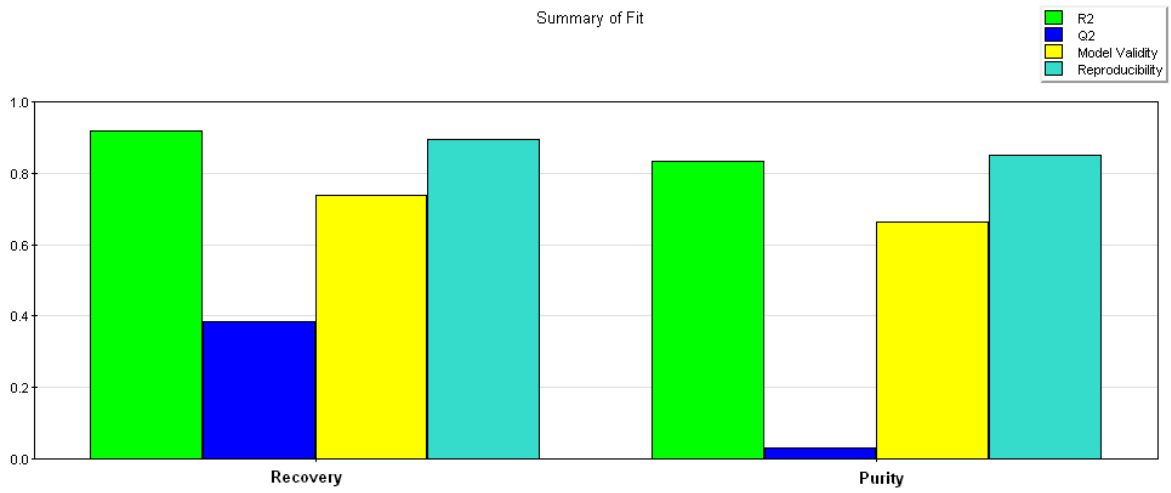

Figure 16. Graphical representation of the statistical coefficients of the model.

Recovery response has a better R^2 than purity response, 0.92 and 0.83 respectively. However both R^2 are high, leading to the assumption that both models fit the data. On the other hand, the adjusted R^2 allows to assess if the theoretical values adjust satisfactorily the experimental data by this model [113]. If the adjusted R^2 is significantly lower than R^2 perhaps the sample size is not adequate to the model [114]. Adjusted R^2 for recovery has a difference of 0.1 when compared to the respective R^2, while adjusted R* for purity has a difference of 0.3 when compared to the respective R^2. This suggests that some experiments could be eliminated in order to improve the statistical model.

The predicted values variation, represented by Q^2, allows to assess if the model is good in predicting new data. The higher the Q^2 is, the useful the model is. Negative values for Q^2 suggest a model devoid of predictive power [115]. Both recovery and purity Q^2 values are positive, suggesting that the model can predict new data. However, Q^2 purity has a low value (0.03), suggesting a worse predictive power when compared to Q^2 recovery (0.38). This suggests that some experiments could be excluded in order to increase the predictive power of the purity model.

On the other side, the model validity of both responses was relatively high, 0.74 for recovery and 0.66 for purity, suggesting that the quadratic model was a good choice. In general, a model is considered valid from values higher than 0.25 [115].

The reproducibility of the model assesses the variation of the response upon the same conditions, through the use of the data obtained for the centre points. Both responses achieved a high reproducibility (0.90 for recovery and 0.85 for purity), suggesting that this is a reproducible model [115].

4.1.2.2 Residual plots

The residual plots serve to identify the outlined experiments, providing a useful insight of experiments that might be interfering with the goodness of the model. Figure 17 consists in the representation of the residual points for the recovery and purity responses. A fitting model is expected to have residual points randomly distributed, falling in a straight line. The points deviated from the straight line are considered as outliners [114].

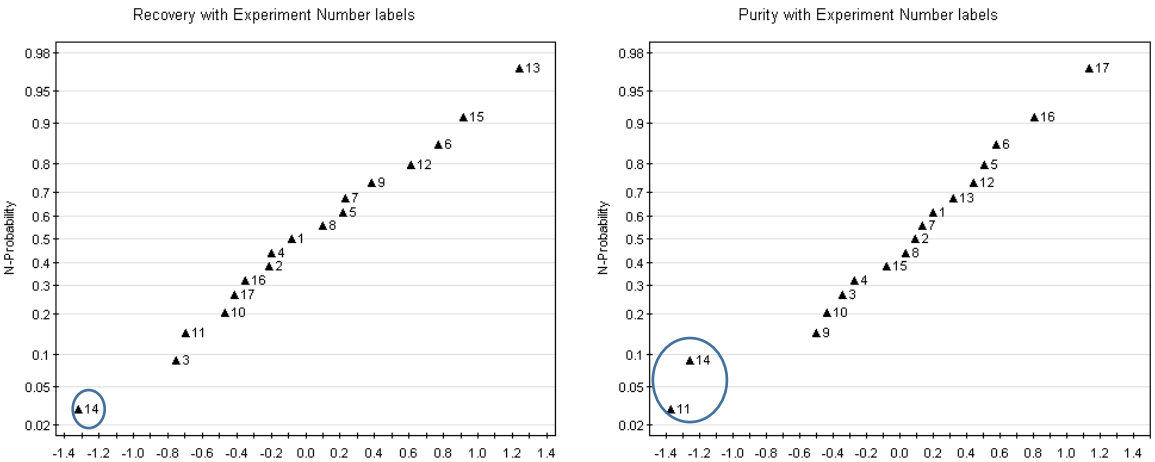

Figure 17. Residual plot or recovery and purity. The experiment numbers are labelled on the graphic.

Thus, figure 17, obtained by UNICORN™ 6.1 software, identifies the existence of two outliers (circled in blue), which might be responsible for the reduced Q^2 value for purity. No

experiment was excluded, since its exclusion would have a negative side effect in the recovery model. Taking into account that both models had acceptable values, outlined experiments were ignored.

4.1.2.3 Predicted versus Actual plots

In figure 18 are represented the predicted versus actual values for recovery (A) and purity (B), provided by Design Expert version 9 trial software. In a good model the points are expected to fall in a 45 degree line.

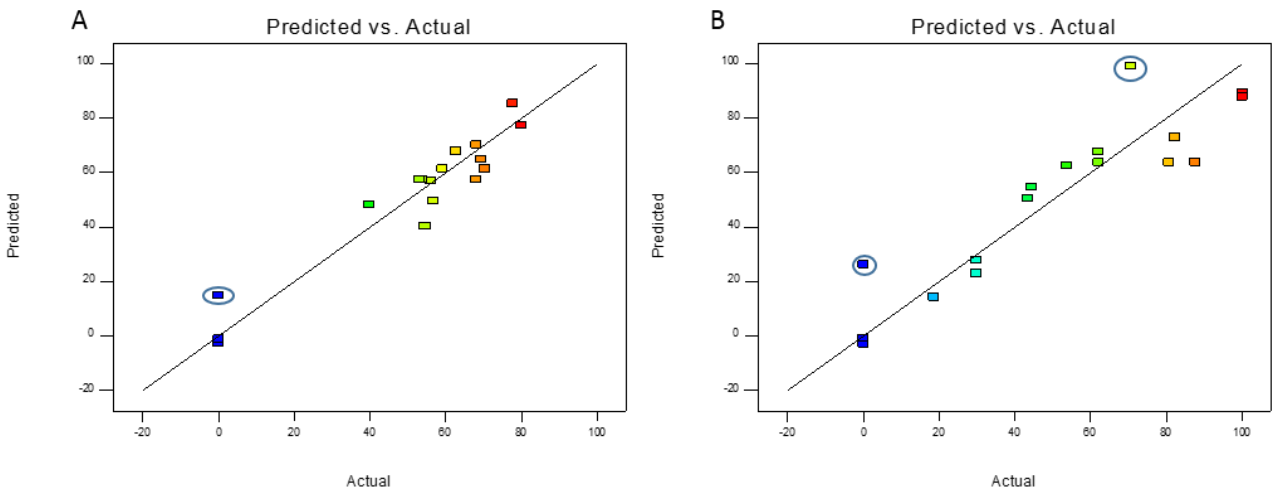

Figure 18. A) Predicted versus Actual plot for recovery. B) Predicted versus Actual plot for purity.

As depicted in figure 18, most of the points fall closer to the 45 degree line, suggesting a good model fitting. The blue and green points circled, which deviate from the 45 degree line, correspond to the outliners points previously identified in the residual plots.

4.1.2.4 Main effects

Through the Design Expert version 9 trial software, it was possible to withdraw the coded multiple regression equation, which permits to evaluate the main effects of the factors. Below are the obtained equations, where A stands for binding, B for washing and C for pH factors:

Recovery = 57.63 + 1.48 * A - 17.84 * B - 12.74 * C + 4.39 * AB - 5.14 * AC - 17.86 * BC + 8.87 * A^2 + 9.97 * B^2 - 29.93 * C^2

Purity = 63.81 + 3.92 * A - 13.12 * B + 1.63 * C + 0.9125 * AB - 3.76 * AC - 32.14 * BC -5.12 * A^2 + 22.28 * B^2 - 39.17 * C^2

Table 10 presents the information withdrawn of the multiple regression equation, corresponding to the main effects of the factors.

Table 10. Summary of the main effects.

Factor	Recovery Effect	Purity Effect
Binding	Positive	Positive
Washing	Negative	Positive
pH	Negative	Positive
Binding * Washing	Positive	Positive
Binding * pH	Negative	Negative
Washing * pH	Negative	Negative
$Binding^2$	Positive	Negative
$Washing^2$	Positive	Positive
pH^2	Negative	Negative

As depicted in table 10, binding factor has a positive effect in both responses, indicating that at a higher binding level the recovery and purity increase. However, washing and pH factors have different behaviours for each response. Both factors have negative effects for recovery and positive effects for purity. This might be explained by the fact that with higher concentrations for the washing step, the target molecule might elute in washing peak and not in the elution peak, due to higher ionic strength, leading to less recovery. On the other hand, other impurities elute as well, leading to a higher purity of the recovered molecule. As described in section 4.1.1 (Preliminary tests for threshold choosing), in the screening for the adequate range of pH, it was found that higher pH would result in unbound of the target molecule, while lower pH would result in higher binding of the molecule to the chromatographic matrix. Therefore, the negative effect in recovery can be related with the early elution of the target molecule at higher pH, resulting in lower recovery. The positive effect on purity can be explained in the same way as the washing factor. At higher pH there is less binding of impurities, leading to purity increase of the recovered sc pDNA.

4.1.2.5 Surface and contour plots

Surface and contour plots are a representation of the multiple regression equation, allowing to assess interaction effects of the factors, while finding the optimal range for maximizing the response values [115]. Figure 19 represents the 3D surface and contour plots for recovery, while figure 20 represents the 3D surface and contour plots for purity, both obtained by Design Expert version 9 trial software. The colours represent a range for the optimal points, in which blue represents the range for the lowest response, followed by green, yellow and finally red, which represents the range for the highest response. In addition, the significance of the interactions occurring between the different factors can be assessed through the contour plots. Significant interactions result in elliptical line shapes in the contour plots, while insignificant interactions result in circular shaped contour lines. Surface plots consist in the 3D representation of the contour plots, allowing to easily visualizing the overall obtained data. Nonetheless, contour plots provide a more precise examination of the obtained data [116].

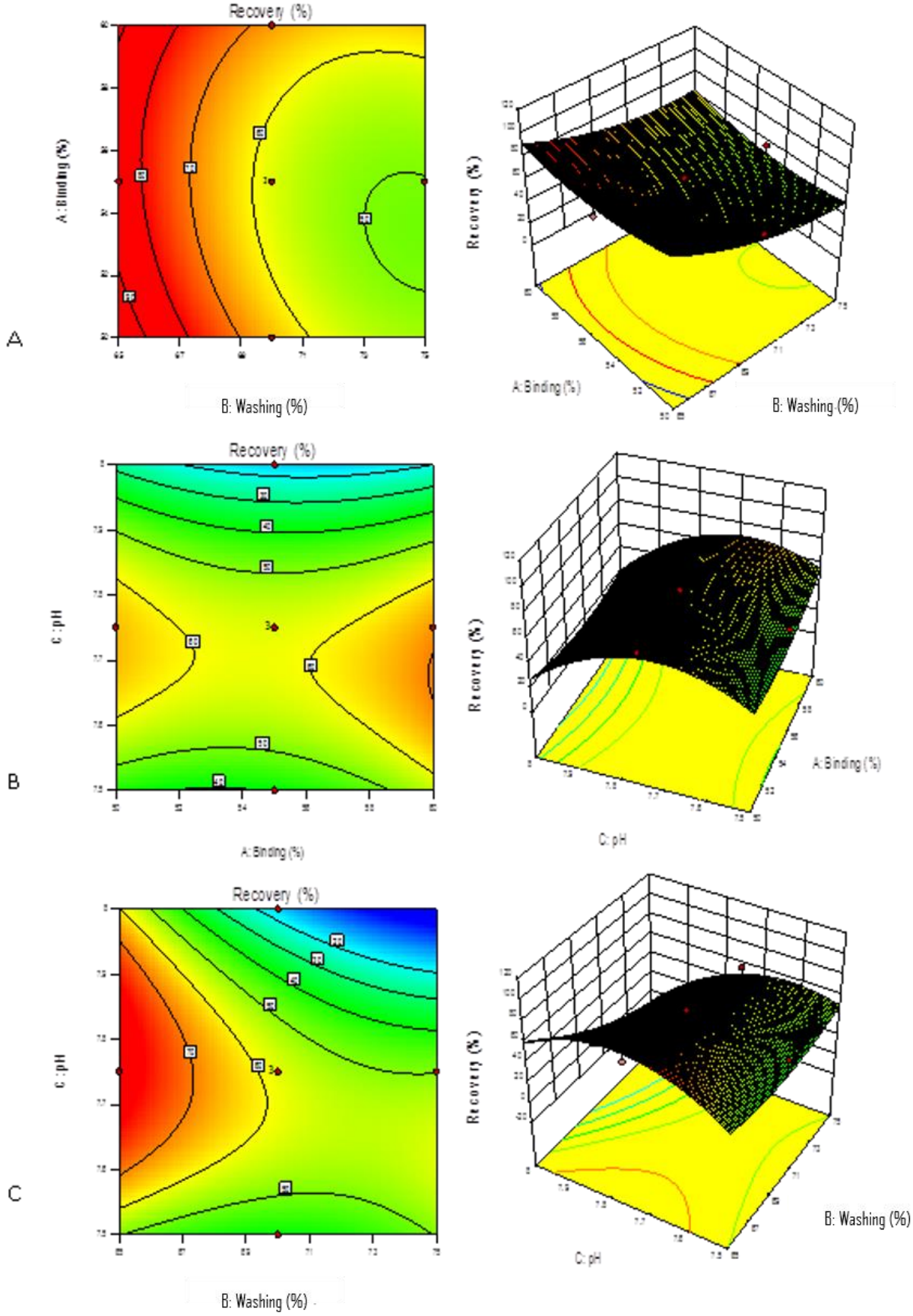

Figure 19. Surface plot and contour plot for recovery response. A) Binding versus washing; B) pH versus binding; C) pH versus washing.

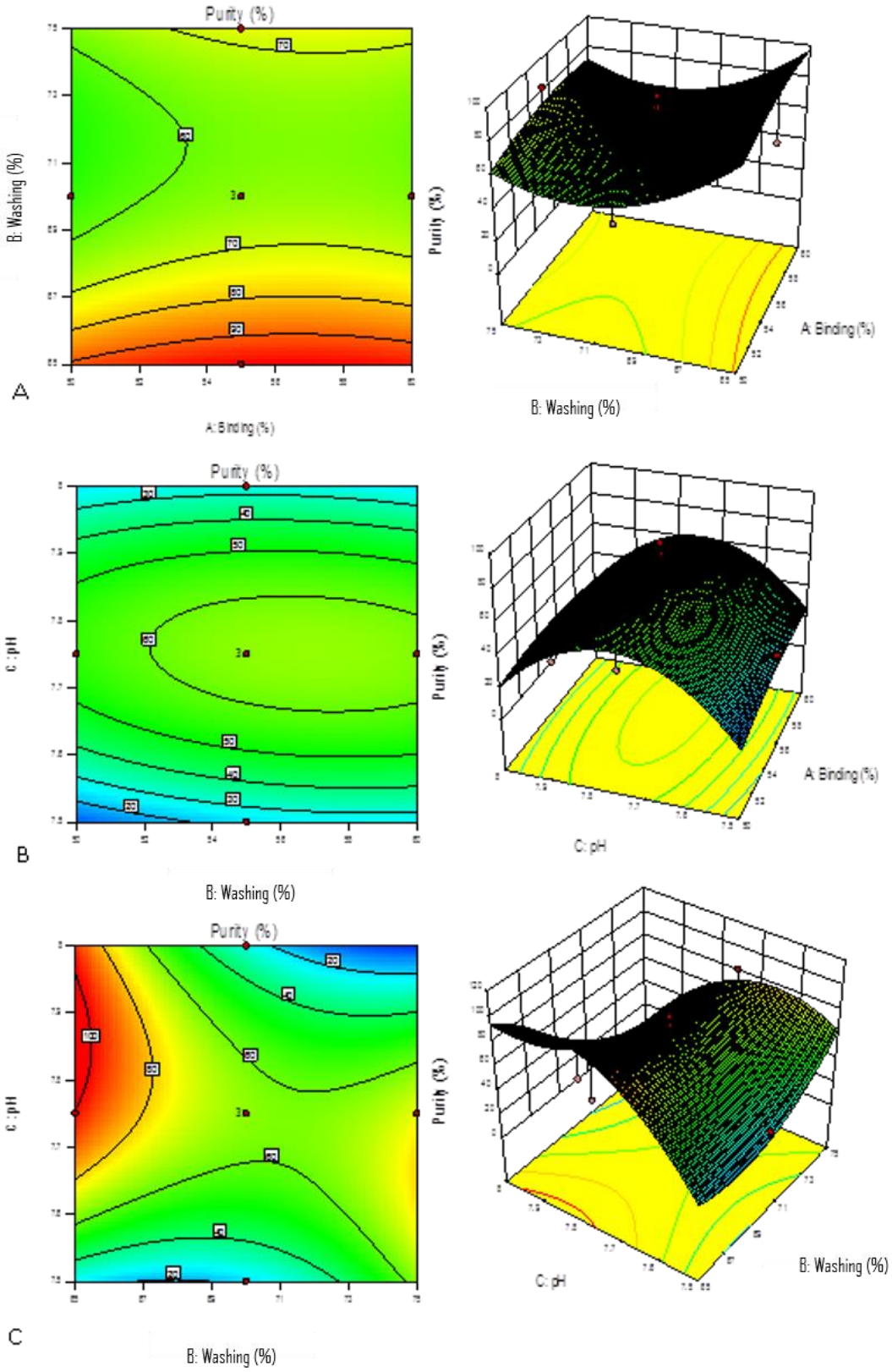

Figure 20. Surface plot and contour plot for purity response. A) Binding versus washing; B) pH versus binding; C) pH versus washing.

Figure 20.(B) suggests there is a strong interaction between washing and pH factors for purity. The elliptical shape of the pH versus binding contour plot demonstrates that the interaction occurs between both factors. In the provided overall contour plots there are different interactions between the different factors for the two responses, except for figure 19.(A), where an almost perfect circular shaped contour plot is presented for binding versus washing factors in recovery response, suggesting no interaction between both factors.

4.1.2.6 Analysis of variance

Analysis of variance (ANOVA) is a statistical tool important in linear regression analysis, which is very used amongst researchers [117]. The concept of ANOVA is based on partitioning the variability of all scores, also described as sum of squares, into two different components, the model and residual sum of squares. Mean of squares is obtained from these two components, providing an estimation of population variances. The ratio of means of squares (F) allows to test if all the treatments have equal effects [112].

Table 11 and 12 represent the ANOVA treatment for recovery and purity responses, respectively. Taking into consideration that the significance probability (p-value) for both responses is lower than 0.05, it is safe to assume that the model data is statistically significant.

Table 11. ANOVA table for recovery response.

Source	Sum of Squares	Degrees of freedom	Mean Square	F Value	p-value	
Model	10152.95	9	1128.11	8.80	0.0045	Significant
A-Binding	21.90	1	21.90	0.17	0.6917	
B-Clarification	3182.66	1	3182.66	24.84	0.0016	
C-pH	1623.08	1	1623.08	12.67	0.0092	
AB	154.00	1	154.00	1.20	0.3092	
AC	211.15	1	211.15	1.65	0.2401	
BC	2552.55	1	2552.55	19.92	0.0029	
A^2	210.95	1	210.95	1.65	0.2403	
B^2	266.49	1	266.49	2.08	0.1925	
C^2	2399.56	1	2399.56	18.73	0.0034	
Residual	896.99	7	128.14	2.10		
Lack of Fit	753.66	5	150.73	8.80	0.3529	not significant
Pure Error	143.33	2	71.66			
Cor Total	11049.94	16				

Table 12. ANOVA table for purity response.

Source	Sum of Squares	Degrees of freedom	Mean Square	F Value	p-value	
Model	15423.61	9	1713.73	3.92	0.0427	Significant
A-Binding	153.66	1	153.66	0.35	0.5720	
B-Clarification	1721.34	1	1721.34	3.94	0.0877	
C-pH	26.57	1	26.57	0.061	0.8124	
AB	6.66	1	6.66	0.015	0.9052	
AC	113.25	1	113.25	0.26	0.6265	
BC	8262.55	1	8262.55	18.89	0.0034	
A^2	70.23	1	70.23	0.16	0.7006	
B^2	1330.01	1	1330.01	3.04	0.1247	
C^2	4110.68	1	4110.68	9.40	0.0182	
Residual	3061.31	7	437.33			
Lack of Fit	2713.37	5	542.67	3.12	0.2604	not significant
Pure Error	347.94	2	173.97			
Cor Total	18484.92	16				

Lack of fit consists in the model data variation and, if p-value is significant, suggests the model does not fill well the data. As presented in tables 11 and 12, the lack of fit shows no significance, suggesting the model fits the data, as previously analysed in section 4.1.2.1 (Goodness of fit).

4.1.3 Model validation

After statistical analysis, model validation was accomplished through execution of three replicates of the optimal point for maximized recovery and purification of sc HPV-16 E6/E7 pDNA. The data recovered from the experimental design was introduced in UNICORN 6.1 software, allowing the prediction of the optimal point, whose conditions are represented in table 13.

Table 13. Optimal range for recovery and purity maximization.

Factors	Value
Binding (NaCl concentration in Tris-EDTA)	501 mM
Washing (NaCl concentration in Tris-EDTA)	650 mM
pH	7.9
Response surface	**Value**
Recovery	91.39
Purity	98.82

Therefore, three chromatographic runs were performed with 501 mM NaCl concentration for binding step, 650 mM NaCl concentration for washing a 1 M NaCl concentration for elution step, with all buffers with pH = 7.9. The execution of three assays allows to assess the reproducibility of the statistical method. In figure 21 are represented the chromatograms of the three experiments, which revealed a similar chromatographic profile, suggesting similar interactions and reproducibility within the three experiments. The elution peaks were pooled and analysed by CIMac™ analytical column in order to determine the recovery and purity of the samples.

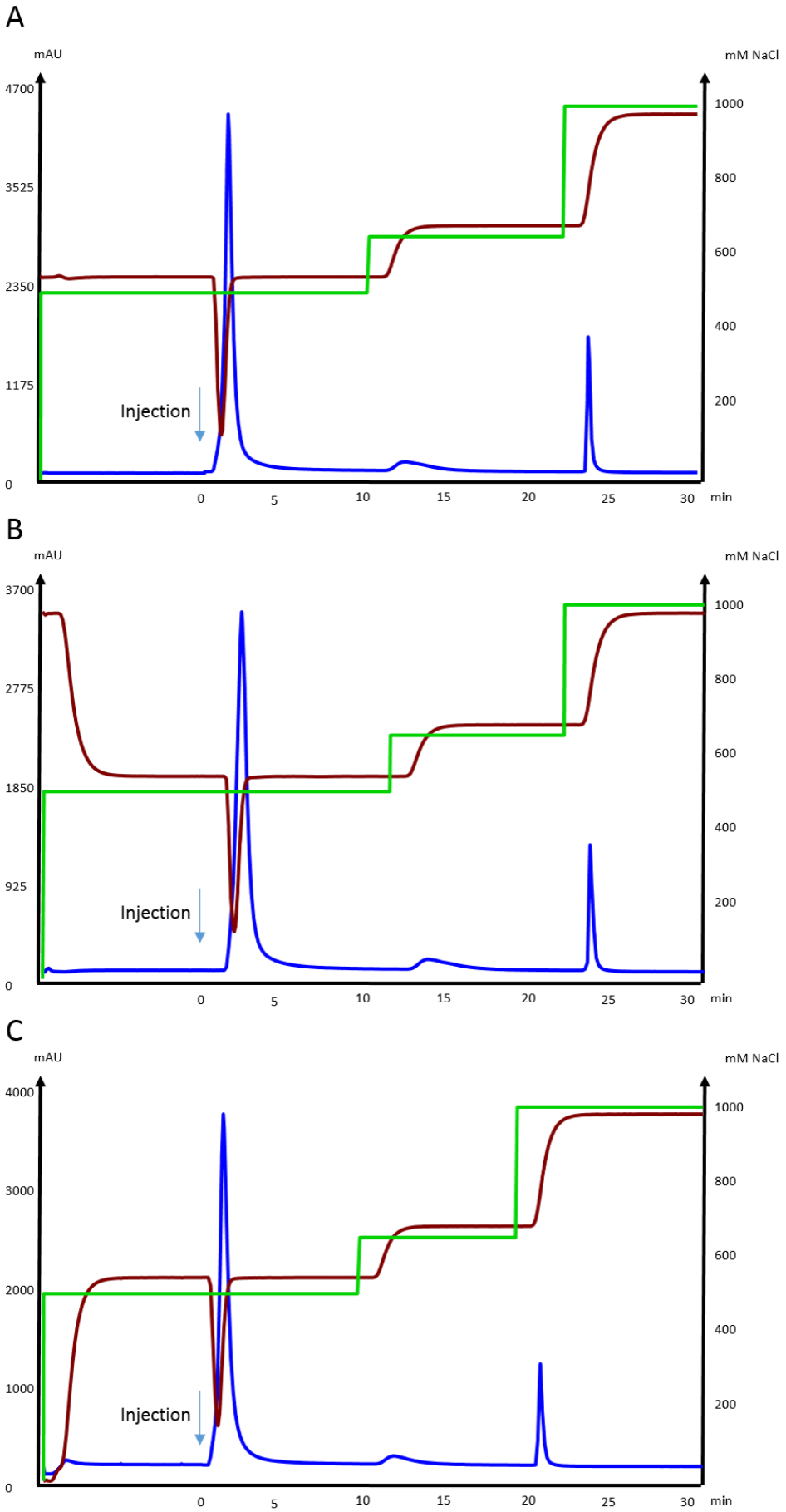

Figure 21. Chromatographic replicate runs for model validation of the optimal point.

In table 14 are represented the responses for the three chromatographic runs as well as the mean of both responses, while in table 15 is represented the confidence interval in which the responses are valid. Although the optimal recovery response was not achieved, the value of the three experiments was close to the predicted value and within the confidence interval, besides outdoing the recovery percentage obtained in the established strategy. Although the optimal point suggested 98.8% of purity, 100% was achieved as expected.

Table 14. Responses obtained from the optimization/validation of the model.

Recovery	Purity
88.76	100
75.83	100
85.82	100
Mean	
Recovery	83.47
Purity	100

Table 15. Confidence intervals for the two responses.

Response	95% CI low	95% CI high
Recovery	63.27	119.51
Purity	46.87	150.77

These data underlie the versatility of the arginine monolith for high grade sc pDNA purification.

4.2 – Transfection efficiency

The principle behind DNA vaccines relies on successful transfection of nucleic acids, which contain genetic information for coding of a given antigen, into host cells. Therefore, transfection efficiency studies are extremely important in order to preliminary assess the ability for transfection of a DNA vaccine. For that, after transfection of CHO-1 cells with the purified sc pDNA obtained by the optimized purification strategy using the arginine monolith and with pDNA obtained with a commercial kit and oc pDNA HPV-16 E6/E7, western blot and immunocytochemistry techniques were performed with the intent of assessing the transfection efficiency and gene expression differences.

4.2.1 Western blot

At a preliminary stage, in order to verify if transfection and gene expression was occurring, Western Blot technique was performed. Cultured cells were transfected with different pDNA isoforms, namely oc and sc pDNA, and with pDNA purified with a commercial technique, as opposed to the sc pDNA purified by the previously described optimized purification strategy. In addition, a control (CT) group, where no transfection was performed, was used as comparison for the obtained data. This procedure relies on the detection of the target proteins, E6 and E7 in this case, from the total proteins extracted from different cell cultures. A polyacrylamide gel electrophoresis was performed with protein samples extracted from cultured cells, by applying 50 µg of total proteins in each well. Afterwards, electroblotting was performed with antibodies for E6 and E7, as it was described in section 2. As depicted in figure 22, it is noticeable an increase staining for kit and sc pDNA when compared with CT for E6 protein. For E7 protein, the detection for sc pDNA is detached when compared to the remaining samples.

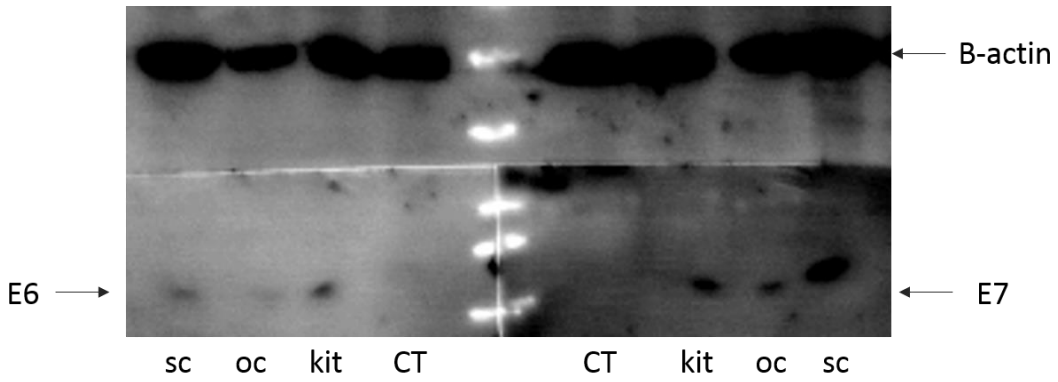

Figure 22. Western Blot for E6 and E7 proteins. At left, E6 protein staining is represented, while E7 protein staining is represented at right. CT – control group, Kit- commercialy purified pDNA transfected group, oc – oc pDNA transfected group, sc – sc pDNA transfected group. B-Actin staining was used for normalization purposes.

The transfection pattern was higher for sc pDNA, followed by kit and oc pDNA. This result is in accordance with the expected, taking into consideration that several studies have proven the sc pDNA greater ability for transfection than other isoforms [72, 73]. The main objective of this preliminary study was to assess if the transfection occurred. Staining of a single band of E6 and E7 proteins was observed for samples transfected with kit, oc and sc pDNA, suggesting that transfection was successful.

In order to compare the transfection pattern of different pDNA samples, quantification of E6 and E7 protein is required. Within western blot technique, quantification of target proteins needs to be normalized through a housekeeping gene, as for instance ß-Actin. However, the

protein quantification by western blot procedure can be affected by several extern factors, and for this reason, such quantification was performed with a more reliable technique such as immunocytochemistry.

4.2.2 – Immunocytochemistry

With the intention of directly assess the structural cell conditions upon transfection, while locating and quantifying the protein expression between the different transfection circumstances, immunocytochemistry was performed. Figure 23 presents different images taken to cultured cells, which were submitted to different transfection conditions and stained with E6 antibody. In figure 23 (A), (D), (G) and (J) it is possible to visualize cell nuclei staining with Hoescht. There are no differences between the nuclei structure of the different transfected groups (D, G and J) to the CT group (A), suggesting that the transfection agent and the vector used for transfection are not inducing cell damage. In figure 22 (B), (E), (H) and (K) it is depicted the staining for the E6 protein with the secondary antibody Alexa 488 anti-mouse (with green fluorescence), while in (C), (F), (I) and (L) both stainings are presented.

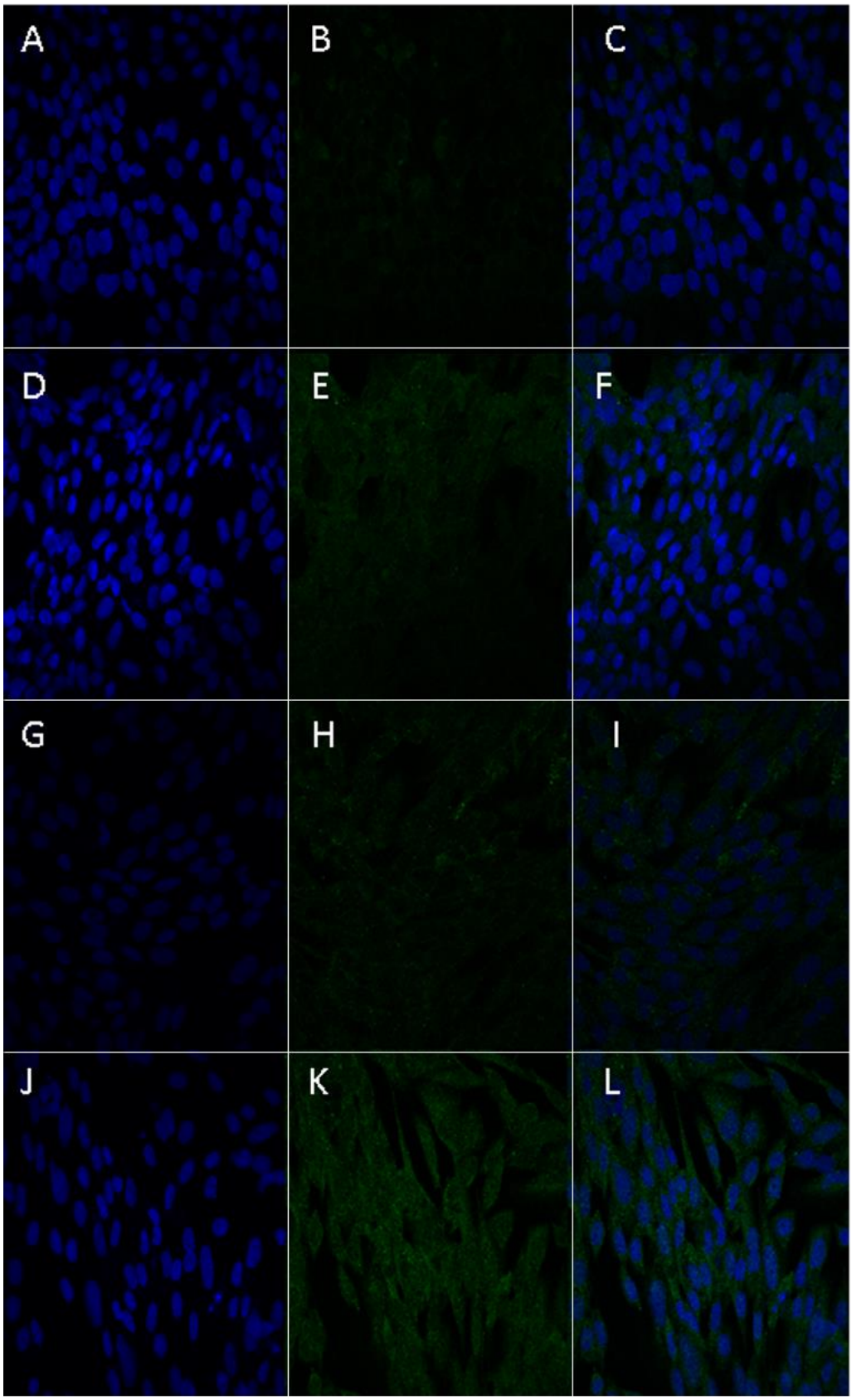

Figure 23. CHO-1 cells immunocytochemistry images for E6 staining. A-C) Control group; D-F) pDNA commercial kit transfected cells; G-H) oc pDNA transfected cells; J-L) sc pDNA transfected cells.

On the other side, different images taken to cultured cells, which were submitted to transfection with sc pDNA HPV-16 E6/E7 and stained with E7 antibody, are depicted in figure 24. In figure 24 (A), is shown cell nuclei staining with Hoescht. There are no structural differences between the nuclei of the transfected group (D) and the CT group (A), suggesting that the transfection agent and the vector used for transfection are not inducing cell damage, as previously described for E6 staining. In figure 24 (B) and (E), it is depicted the staining for the E7 protein with the secondary antibody Alexa 546 anti-mouse (with red fluorescence), while in (C) and (F) both stainings are represented.

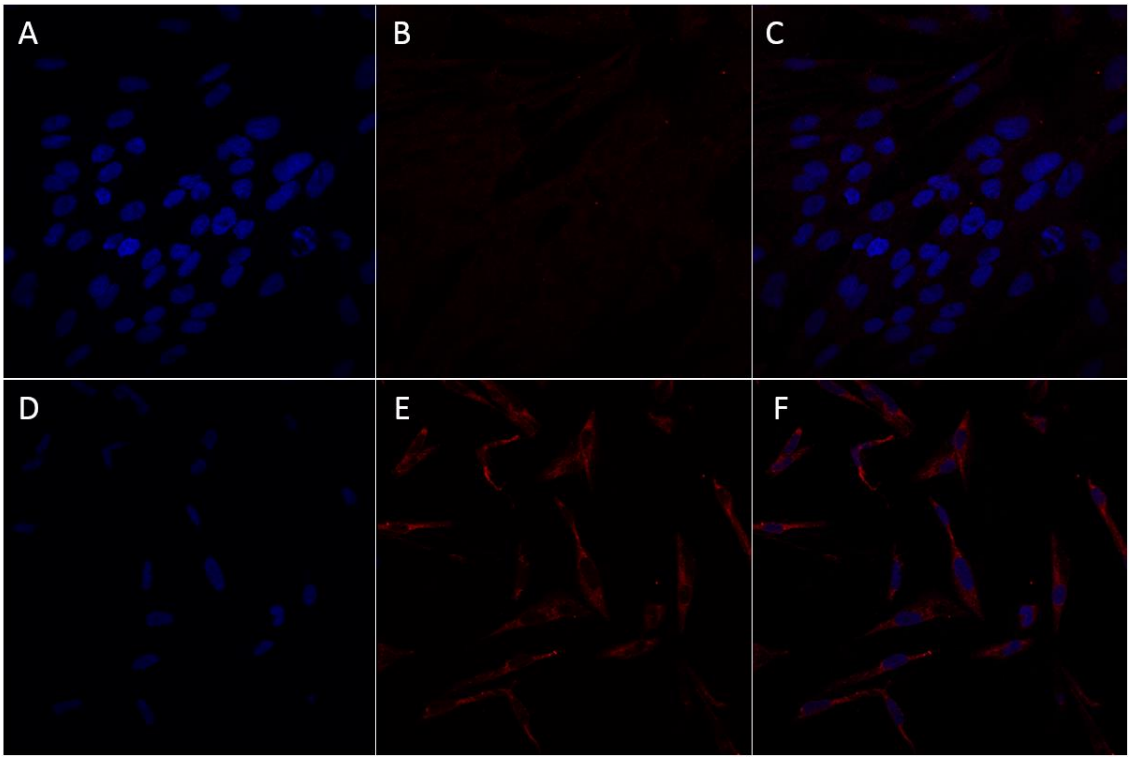

Figure 24. CHO-1 cells immunocytochemistry images for E7 staining. A-C) Control group; D-F) sc pDNA transfected cells.

4.2.2.1 E6 immunofluorescence quantification

Immunofluorescence was quantified through ImageJ 1.48u software in order to further compare protein expression between the different analysed groups with GraphPad Prism version 6.04 trial software statistical tool. Table 16 represents the different immunofluorescence values obtained for 15 images taken from the different sample groups. Each set of 5 pictures represents a different experiment replicate (n=1).

Table 16. Immunofluorescence for CT, kit, oc and sc pDNA transfected groups for E6 protein (n=3).

Picture	CT	Kit	Oc	sc
1	7.676	12.848	9.253	19.549
2	8.055	11.830	9.975	20.315
3	8.647	10.511	9.629	21.535
4	8.087	9.908	9.952	22.950
5	9.012	9.806	9.566	17.807
6	8.236	11.414	9.765	15.804
7	8.503	9.531	9.854	15.047
8	9.543	9.163	9.24	17.370
9	8.379	8.796	7.665	16.820
10	9.739	8.954	9.463	17.915
11	7.571	10.787	11.081	19.014
12	7.556	9.471	13.275	15.714
13	7.837	10.852	10.645	15.594
14	7.623	10.450	10.328	15.598
15	7.818	10.170	10.784	15.966

In figure 25 is graphically represented the statistical analysis of the different immunofluorescence group values by one-way ANOVA. It was verified that sc pDNA-induced immunofluorescence was significantly higher than the other analysed groups, suggesting higher transfection efficiency and target gene expression. Also, there was no significant immunofluorescence difference between kit, oc pDNA and CT groups

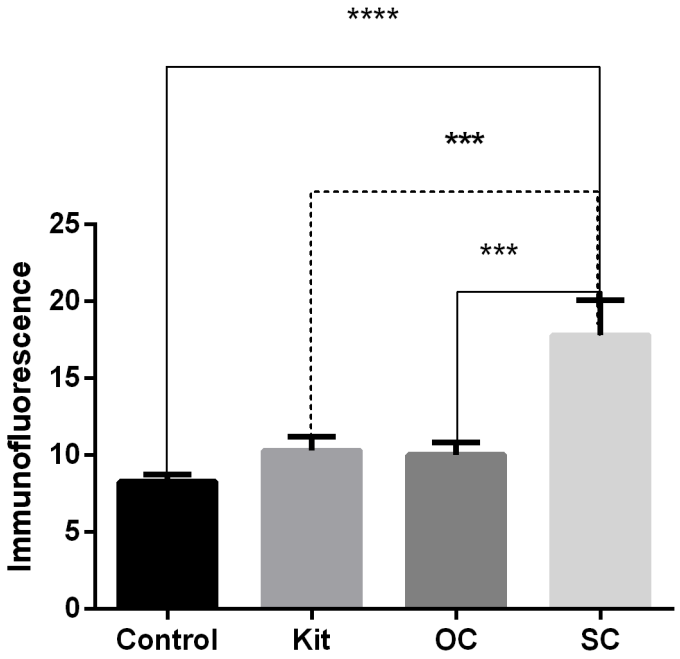

Figure 25. E6 protein immunofluorescence comparison between the different groups (pvalue<0.001).

4.2.2.2 E7 immunofluorescence quantification

Similarly to the previously presented E6 protein expression, in order to perform immunofluorescence quantification of E7 protein, the ImageJ 1.48u software was used. The immunofluorescence values were later introduced in GraphPad Prism version 6.04 trial software, with the intent of executing statistical analysis of the obtained data by one-way ANOVA. Table 17 represents the different values of the 15 pictures acquired for immunofluorescence of the different analysed groups, namely CT and sc pDNA treated samples. Each set of 5 pictures represents a different experiment replicate (n=1).

Table 17. Immunofluorescence for CT, kit, oc and sc pDNA transfected groups for E7 protein (n=3).

Pictures	CT	sc
1	10.218	35.829
2	9.302	48.648
3	7.240	40.113
4	7.934	44.171
5	5.395	32.281
6	7.847	27.099
7	7.094	21.938
8	5.116	22.448
9	8.418	25.221
10	6.453	17.616
11	9.156	19.194
12	6.060	21.14
13	8.709	22.07
14	10.222	17.562
15	7.110	20.101

As visible in figure 26, a statistical representation by unpaired t-student test of the different immunofluorescence group values, sc pDNA immunofluorescence was significantly higher when compared to CT group. These data are in agreement with the information obtained from E6 immunofluorescence quantification, suggesting that the sc pDNA obtained by the optimized purification strategy previously described is fitted for transfection.

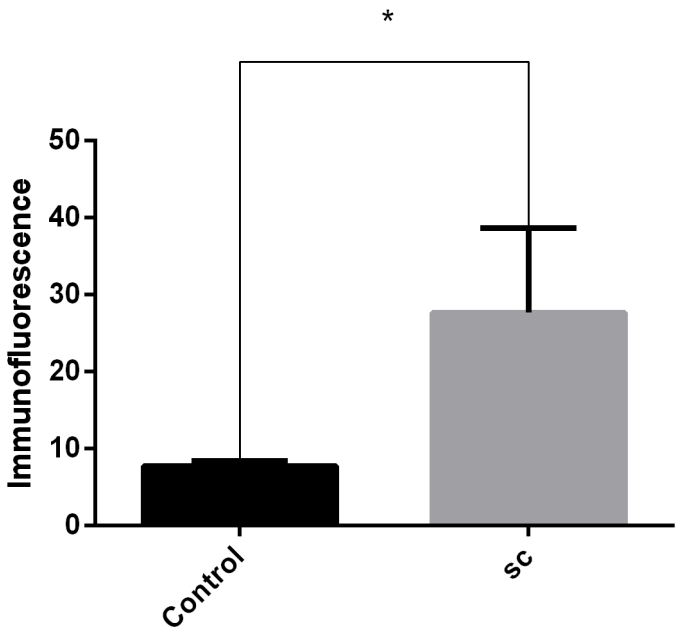

Figure 26. E7 protein Immunofluorescence comparison between the CT group and sc pDNA transfected group (pvalue<0.05).

CHAPTER 5 – Conclusions and Future Perspectives

In the current days, it has been observed an increase in patients number diagnosed with cancer, great part developed by viral or bacterial infections. This problematic might be attenuated with DNA vaccine technology, which permits the prevention and treatment of such infections. Purification strategies are extremely important to guarantee the administration of qualified vaccines, without biosafety concerns. The development of chromatographic techniques should focus on the maximum recovery and purity of the target molecule, in order to take fully advantage of the purification process. Our research group recently published a purification strategy to efficiently purify sc pDNA HPV-16 E6/E7 DNA vaccine with an arginine monolith. This strategy was able to reach 100% purity, but only 39% of recovery was achieved. Thus, design of experiments is a statistical tool able to aid researchers to screen, optimize and test the robustness of their experiments. For optimization purposes, this tool allows the design of different experiments that have to be performed in order to achieve the optimal range for the response chosen to be improved, evaluating different factors chosen a priori by the investigator. Thus, the present thesis was designed in order to achieve the optimization of the purification strategy of sc pDNA HPV-16 E6/E7 vaccine in order to perform transfection studies for gene expression assessment.

Firstly, several screening experiments were performed in order to identify the factors which should be evaluated and the range between their values should vary. Once the factors and their range were chosen, Composite Central Face experimental design was applied in order to assess the optimal point for maximum recovery and purity of sc pDNA HPV-16 E6/E7. The factors chosen were the binding step, the washing step and the pH. The chromatographic runs were performed with three steps, in which the target molecule was expected to elute within the last step. Pooled fractions of the third peak were analysed through CIMac™ analytical column with the intent of determining the recovery and purity percentage of the experiment. Statistical analysis provided an adequate insight of the designed optimization model, suggesting that the model was significant and valid. The prediction of the optimal range allowed the execution of three experiment replicates, in which 83.5% of recovery and 100% of purity were achieved. Therefore, the optimization of the purification strategy of sc pDNA HPV-16 E6/E7 by arginine monolith was successfully achieved.

Then, transfection studies with the sc pDNA HPV-16 E6/E7, obtained by the optimized strategy, were performed in order to evaluate transfection efficiency and gene expression of the purified molecule in comparison to oc pDNA and pDNA purified with a commercial kit. CHO-1 cells were cultured and transfected with the different pDNA samples. Western blot and immunocytochemistry techniques were performed and gene expression quantification for E6

and E7 proteins was accomplished through adequate software. For E6 gene expression assessment, commercially purified pDNA, oc and sc pDNA efficiency was evaluated in comparison to control group. It was found that sc pDNA gene expression was significantly higher than the remaining groups. On the other hand, for E7 expression assessment, only sc pDNA was evaluated in comparison to control group. Gene expression for E7 protein was significantly higher than control group. These data suggest the sc pDNA HPV-16 E6/E7 purified by our optimized purification strategy has higher transfection efficiency and gene expression when compared to other isoforms and other purification methodologies.

However further testing is necessary in order to fully support the obtained data. For instance, evaluation of E7 gene expression in comparison to commercially purified pDNA and oc pDNA should be performed, as well as protein quantification by western blot. Toxicity tests have also major importance in order to guarantee DNA vaccine biosafety. In addition, reverse transcriptase polymerase chain reaction would be an interesting technique to perform in order to assess the messenger RNA for both proteins, keeping track of pDNA genetic expression throughout the whole transfection procedure. Also, transfection studies with a dendritic cell line could provide an insightful perspective of antigen presentation. Afterwards, *in vivo* studies should be perform with the intent of determining the sc pDNA HPV-16 E6/E7 ability in providing an adequate immune response, such as detection and determination of cytokines through flow cytometry.

Overall, the optimization of sc pDNA HPV-16 E6/E7 purification strategy through the use of experimental design tools allowed improving the recovery of the target molecule, while maintaining a 100% of purity. In addition, the transfection efficiency and gene expression ability of this DNA vaccine was proven, supporting the suitability of sc isoform for transfection studies as well as the importance in developing adequate purification strategies.

References

1. Watson, J.D. and F.H. Crick, *The structure of DNA*. Cold Spring Harb Symp Quant Biol, 1953. **18**: p. 123-31.
2. Cohen, S.N., et al., *Construction of biologically functional bacterial plasmids in vitro*. Proc Natl Acad Sci U S A, 1973. **70**(11): p. 3240-4.
3. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
4. Ragusa, A., I. Garcia, and S. Penades, *Nanoparticles as nonviral gene delivery vectors*. IEEE Trans Nanobioscience, 2007. **6**(4): p. 319-30.
5. Wang, W., et al., *Non-viral gene delivery methods*. Curr Pharm Biotechnol, 2013. **14**(1): p. 46-60.
6. Wirth, T., N. Parker, and S. Yla-Herttuala, *History of gene therapy*. Gene, 2013. **525**(2): p. 162-9.
7. Glover, D.J., H.J. Lipps, and D.A. Jans, *Towards safe, non-viral therapeutic gene expression in humans*. Nat Rev Genet, 2005. **6**(4): p. 299-310.
8. Gao, X., K.S. Kim, and D. Liu, *Nonviral gene delivery: what we know and what is next*. AAPS J, 2007. **9**(1): p. E92-104.
9. Rubanyi, G.M., *The future of human gene therapy*. Mol Aspects Med, 2001. **22**(3): p. 113-42.
10. Edelstein, M.L. *Gene Therapy Clinical Trials Worldwide*. 2014 January-2014 [cited 2014 21 January]; Available from: http://www.abedia.com/wiley/indications.php.
11. Abdelnoor, A.M., *Plasmid DNA vaccines*. Curr Drug Targets Immune Endocr Metabol Disord, 2001. **1**(1): p. 79-92.
12. Liu, M.A., *DNA vaccines: a review*. J Intern Med, 2003. **253**(4): p. 402-10.
13. Gurunathan, S., D.M. Klinman, and R.A. Seder, *DNA vaccines: immunology, application, and optimization**. Annu Rev Immunol, 2000. **18**: p. 927-74.
14. Josefsberg, J.O. and B. Buckland, *Vaccine process technology*. Biotechnol Bioeng, 2012. **109**(6): p. 1443-60.
15. Kim, J.V., et al., *The ABCs of artificial antigen presentation*. Nat Biotechnol, 2004. **22**(4): p. 403-10.
16. Lichtor, T. and R.P. Glick, *Immunogene therapy*. Adv Exp Med Biol, 2012. **746**: p. 151-65.
17. Cho, J.H., J.W. Youn, and Y.C. Sung, *Cross-priming as a predominant mechanism for inducing CD8(+) T cell responses in gene gun DNA immunization*. J Immunol, 2001. **167**(10): p. 5549-57.
18. Klencke, B., et al., *Encapsulated plasmid DNA treatment for human papillomavirus 16-associated anal dysplasia: a Phase I study of ZYC101*. Clin Cancer Res, 2002. **8**(5): p. 1028-37.
19. Manthorpe, M., et al., *Plasmid vaccines and therapeutics: from design to applications*. Adv Biochem Eng Biotechnol, 2005. **99**: p. 41-92.
20. Rottinghaus, S.T., et al., *Hepatitis B DNA vaccine induces protective antibody responses in human non-responders to conventional vaccination*. Vaccine, 2003. **21**(31): p. 4604-8.

21. Liao, J.C., et al., *Vaccination with human tyrosinase DNA induces antibody responses in dogs with advanced melanoma*. Cancer Immun, 2006. **6**: p. 8.

22. Weide, B., et al., *Plasmid DNA- and messenger RNA-based anti-cancer vaccination*. Immunol Lett, 2008. **115**(1): p. 33-42.

23. Hung, C.F., et al., *DNA vaccines for cervical cancer: from bench to bedside*. Exp Mol Med, 2007. **39**(6): p. 679-89.

24. Baban, C.K., et al., *Bacteria as vectors for gene therapy of cancer*. Bioeng Bugs, 2010. **1**(6): p. 385-94.

25. Vannucci, L., et al., *Viral vectors: a look back and ahead on gene transfer technology*. New Microbiol, 2013. **36**(1): p. 1-22.

26. Kay, M.A., J.C. Glorioso, and L. Naldini, *Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics*. Nat Med, 2001. **7**(1): p. 33-40.

27. Sheridan, C., *Gene therapy finds its niche*. Nat Biotechnol, 2011. **29**(2): p. 121-8.

28. Sousa, F., L. Passarinha, and J.A. Queiroz, *Biomedical application of plasmid DNA in gene therapy: a new challenge for chromatography*. Biotechnol Genet Eng Rev, 2010. **26**: p. 83-116.

29. Schaffner, W., *Direct transfer of cloned genes from bacteria to mammalian cells*. Proc Natl Acad Sci U S A, 1980. **77**(4): p. 2163-7.

30. Tangney, M., *Gene therapy for cancer: dairy bacteria as delivery vectors*. Discov Med, 2010. **10**(52): p. 195-200.

31. Jones, C.H., et al., *Polymyxin B treatment improves bactofection efficacy and reduces cytotoxicity*. Mol Pharm, 2013. **10**(11): p. 4301-8.

32. Ginn, S.L., et al., *Gene therapy clinical trials worldwide to 2012 - an update*. J Gene Med, 2013. **15**(2): p. 65-77.

33. Lavigne, M.D. and D.C. Gorecki, *Emerging vectors and targeting methods for nonviral gene therapy*. Expert Opin Emerg Drugs, 2006. **11**(3): p. 541-57.

34. Al-Dosari, M.S. and X. Gao, *Nonviral gene delivery: principle, limitations, and recent progress*. AAPS J, 2009. **11**(4): p. 671-81.

35. Elnekave, M., K. Furmanov, and A.H. Hovav, *Intradermal naked plasmid DNA immunization: mechanisms of action*. Expert Rev Vaccines, 2011. **10**(8): p. 1169-82.

36. Saade, F. and N. Petrovsky, *Technologies for enhanced efficacy of DNA vaccines*. Expert Rev Vaccines, 2012. **11**(2): p. 189-209.

37. Li, L., F. Saade, and N. Petrovsky, *The future of human DNA vaccines*. J Biotechnol, 2012. **162**(2-3): p. 171-82.

38. Aihara, H. and J. Miyazaki, *Gene transfer into muscle by electroporation in vivo*. Nat Biotechnol, 1998. **16**(9): p. 867-70.

39. Wang, T., J.R. Upponi, and V.P. Torchilin, *Design of multifunctional non-viral gene vectors to overcome physiological barriers: dilemmas and strategies*. Int J Pharm, 2012. **427**(1): p. 3-20.

40. Torchilin, V.P., *Recent advances with liposomes as pharmaceutical carriers*. Nat Rev Drug Discov, 2005. **4**(2): p. 145-60.

41. Gaspar, V.M., et al., *Formulation of chitosan-TPP-pDNA nanocapsules for gene therapy applications*. Nanotechnology, 2011. **22**(1): p. 015101.

42. Oosterhuis, K., et al., *Preclinical development of highly effective and safe DNA vaccines directed against HPV 16 E6 and E7*. Int J Cancer, 2011. **129**(2): p. 397-406.

43. Henken, F.E., et al., *Preclinical safety evaluation of DNA vaccines encoding modified HPV16 E6 and E7*. Vaccine, 2012. **30**(28): p. 4259-66.

44. Trimble, C., et al., *Comparison of the CD8+ T cell responses and antitumor effects generated by DNA vaccine administered through gene gun, biojector, and syringe*. Vaccine, 2003. **21**(25-26): p. 4036-42.

45. Peng, S., et al., *A combination of DNA vaccines targeting human papillomavirus type 16 E6 and E7 generates potent antitumor effects*. Gene Ther, 2006. **13**(3): p. 257-65.

46. Lin, C.T., et al., *A DNA vaccine encoding a codon-optimized human papillomavirus type 16 E6 gene enhances CTL response and anti-tumor activity*. J Biomed Sci, 2006. **13**(4): p. 481-8.

47. de Oliveira, C.M., et al., *Human papillomavirus genotypes distribution in 175 invasive cervical cancer cases from Brazil*. BMC Cancer, 2013. **13**: p. 357.

48. Frazer, I.H., *Prevention of cervical cancer through papillomavirus vaccination*. Nat Rev Immunol, 2004. **4**(1): p. 46-54.

49. Stern, P.L., et al., *Therapy of human papillomavirus-related disease*. Vaccine, 2012. **30 Suppl 5**: p. F71-82.

50. Chan, P.K., et al., *Laboratory and clinical aspects of human papillomavirus testing*. Crit Rev Clin Lab Sci, 2012. **49**(4): p. 117-36.

51. Brankovic, I., P. Verdonk, and I. Klinge, *Applying a gender lens on human papillomavirus infection: cervical cancer screening, HPV DNA testing, and HPV vaccination*. Int J Equity Health, 2013. **12**: p. 14.

52. Munoz, N., et al., *Chapter 1: HPV in the etiology of human cancer*. Vaccine, 2006. **24 Suppl 3**: p. S3/1-10.

53. Munger, K., et al., *The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes*. J Virol, 1989. **63**(10): p. 4417-21.

54. Best, S.R., K.J. Niparko, and S.I. Pai, *Biology of human papillomavirus infection and immune therapy for HPV-related head and neck cancers*. Otolaryngol Clin North Am, 2012. **45**(4): p. 807-22.

55. Vande Pol, S.B. and A.J. Klingelhutz, *Papillomavirus E6 oncoproteins*. Virology, 2013. **445**(1-2): p. 115-37.

56. Brimer, N., C. Lyons, and S.B. Vande Pol, *Association of E6AP (UBE3A) with human papillomavirus type 11 E6 protein*. Virology, 2007. **358**(2): p. 303-10.

57. Pang, C.L. and F. Thierry, *Human papillomavirus proteins as prospective therapeutic targets*. Microb Pathog, 2013. **58**: p. 55-65.

58. Valente, J.F., et al., *Selective purification of supercoiled p53-encoding pDNA with L-methionine-agarose matrix*. Anal Biochem, 2014.

59. Howie, H.L., R.A. Katzenellenbogen, and D.A. Galloway, *Papillomavirus E6 proteins*. Virology, 2009. **384**(2): p. 324-34.

60. Zimmermann, H., et al., *Interaction with CBP/p300 enables the bovine papillomavirus type 1 E6 oncoprotein to downregulate CBP/p300-mediated transactivation by p53*. J Gen Virol, 2000. **81**(Pt 11): p. 2617-23.

61. Liu, Y., et al., *Multiple functions of human papillomavirus type 16 E6 contribute to the immortalization of mammary epithelial cells*. J Virol, 1999. **73**(9): p. 7297-307.

62. Roman, A. and K. Munger, *The papillomavirus E7 proteins*. Virology, 2013. **445**(1-2): p. 138-68.

63. Garland, S.M. and J.S. Smith, *Human papillomavirus vaccines: current status and future prospects*. Drugs, 2010. **70**(9): p. 1079-98.

64. Hong, D., et al., *Gene silencing of HPV16 E6/E7 induced by promoter-targeting siRNA in SiHa cells*. Br J Cancer, 2009. **101**(10): p. 1798-804.

65. Cun, B., et al., *Cell growth inhibition in HPV 18 positive uveal melanoma cells by E6/E7 siRNA*. Tumour Biol, 2013. **34**(3): p. 1801-6.

66. Trimble, C.L., et al., *A phase I trial of a human papillomavirus DNA vaccine for HPV16+ cervical intraepithelial neoplasia 2/3*. Clin Cancer Res, 2009. **15**(1): p. 361-7.

67. Goncalves, G.A., et al., *Rational engineering of Escherichia coli strains for plasmid biopharmaceutical manufacturing*. Biotechnol J, 2012. **7**(2): p. 251-61.

68. Walther, W., et al., *A seven-year storage report of good manufacturing practice-grade naked plasmid DNA: stability, topology, and in vitro/in vivo functional analysis*. Hum Gene Ther Clin Dev, 2013. **24**(4): p. 147-53.

69. Sousa, A., F. Sousa, and J.A. Queiroz, *Advances in chromatographic supports for pharmaceutical-grade plasmid DNA purification*. J Sep Sci, 2012. **35**(22): p. 3046-58.

70. Herweijer, H. and J.A. Wolff, *Progress and prospects: naked DNA gene transfer and therapy*. Gene Ther, 2003. **10**(6): p. 453-8.

71. Sousa, F., et al., *Separation of supercoiled and open circular plasmid DNA isoforms by chromatography with a histidine-agarose support*. Anal Biochem, 2005. **343**(1): p. 183-5.

72. Dhanoya, A., B.M. Chain, and E. Keshavarz-Moore, *The impact of DNA topology on polyplex uptake and transfection efficiency in mammalian cells*. J Biotechnol, 2011. **155**(4): p. 377-86.

73. Dhanoya, A., B.M. Chain, and E. Keshavarz-Moore, *Role of DNA topology in uptake of polyplex molecules by dendritic cells*. Vaccine, 2012. **30**(9): p. 1675-81.

74. Oliveira, P.H. and J. Mairhofer, *Marker-free plasmids for biotechnological applications - implications and perspectives*. Trends Biotechnol, 2013. **31**(9): p. 539-47.

75. Howley, P.M. *Addgene plasmid 8641*. 2014 [cited 2014 17-02]; Available from: http://www.addgene.org/8641/.

76. Ghanem, A., R. Healey, and F.G. Adly, *Current trends in separation of plasmid DNA vaccines: a review*. Anal Chim Acta, 2013. **760**: p. 1-15.

77. Ferreira, G.N., et al., *Downstream processing of plasmid DNA for gene therapy and DNA vaccine applications*. Trends Biotechnol, 2000. **18**(9): p. 380-8.

78. Sousa, A., F. Sousa, and J.A. Queiroz, *Differential interactions of plasmid DNA, RNA and genomic DNA with amino acid-based affinity matrices*. J Sep Sci, 2010. **33**(17-18): p. 2610-8.

79. Duong-Ly, K.C. and S.B. Gabelli, *Gel filtration chromatography (size exclusion chromatography) of proteins*. Methods Enzymol, 2014. **541**: p. 105-14.

80. Ahmadi-Abhari, S., et al., *Assessment of the influence of amylose-LPC complexation on the extent of wheat starch digestibility by size-exclusion chromatography*. Food Chem, 2013. **141**(4): p. 4318-23.

81. Sousa, F., D.M. Prazeres, and J.A. Queiroz, *Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA*. Trends Biotechnol, 2008. **26**(9): p. 518-25.

82. Matos, T., J.A. Queiroz, and L. Bulow, *Binding and elution behavior of small deoxyribonucleic acid fragments on a strong anion-exchanger multimodal chromatography resin*. J Chromatogr A, 2013. **1302**: p. 40-4.

83. Sousa, F., D.M. Prazeres, and J.A. Queiroz, *Binding and elution strategy for improved performance of arginine affinity chromatography in supercoiled plasmid DNA purification*. Biomed Chromatogr, 2009. **23**(2): p. 160-5.

84. Sousa, A., et al., *Successful application of monolithic innovative technology using a carbonyldiimidazole disk to purify supercoiled plasmid DNA suitable for pharmaceutical applications*. J Chromatogr A, 2011. **1218**(46): p. 8333-43.

85. Costioli, M.D., et al., *DNA purification by triple-helix affinity precipitation.* Biotechnol Bioeng, 2003. **81**(5): p. 535-45.

86. Sousa, A., et al., *Histidine affinity chromatography of homo-oligonucleotides. Role of multiple interactions on retention.* Biomed Chromatogr, 2009. **23**(7): p. 745-53.

87. Sousa, A., F. Sousa, and J.A. Queiroz, *Selectivity of arginine chromatography in promoting different interactions using synthetic oligonucleotides as model.* J Sep Sci, 2009. **32**(10): p. 1665-72.

88. Sousa, F., et al., *Selective purification of supercoiled plasmid DNA from clarified cell lysates with a single histidine-agarose chromatography step.* Biotechnol Appl Biochem, 2006. **45**(Pt 3): p. 131-40.

89. Sousa, A., F. Sousa, and J.A. Queiroz, *Impact of lysine-affinity chromatography on supercoiled plasmid DNA purification.* J Chromatogr B Analyt Technol Biomed Life Sci, 2011. **879**(30): p. 3507-15.

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

91. Nischang, I., *Porous polymer monoliths: morphology, porous properties, polymer nanoscale gel structure and their impact on chromatographic performance.* J Chromatogr A, 2013. **1287**: p. 39-58.

92. Jungbauer, A. and R. Hahn, *Monoliths for fast bioseparation and bioconversion and their applications in biotechnology.* J Sep Sci, 2004. **27**(10-11): p. 767-78.

93. Pfaunmiller, E.L., et al., *Affinity monolith chromatography: a review of principles and recent analytical applications.* Anal Bioanal Chem, 2013. **405**(7): p. 2133-45.

94. Soares, A., et al., *Purification of human papillomavirus 16 E6/E7 plasmid deoxyribonucleic acid-based vaccine using an arginine modified monolithic support.* J Chromatogr A, 2013. **1320**: p. 72-9.

95. Hibbert, D.B., *Experimental design in chromatography: a tutorial review.* J Chromatogr B Analyt Technol Biomed Life Sci, 2012. **910**: p. 2-13.

96. Ferreira, S.L., et al., *Statistical designs and response surface techniques for the optimization of chromatographic systems.* J Chromatogr A, 2007. **1158**(1-2): p. 2-14.

97. Afonso, A., et al., *Purification of pre-miR-29 by a new O-phospho-l-tyrosine affinity chromatographic strategy optimized using design of experiments.* J Chromatogr A, 2014. **1343**: p. 119-27.

98. Lebed, P.J., et al., *Optimization of solid phase extraction chromatography for the separation of Np from U and Pu using experimental design tools in complex matrices.* Analytical Methods, 2014. **6**(1): p. 139-146.

99. Martins, S.A., et al., *Application of central composite design for DNA hybridization onto magnetic microparticles.* Anal Biochem, 2009. **391**(1): p. 17-23.

100. Charette, S.J., et al., *Protein quantification by chemiluminescent Western blotting: elimination of the antibody factor by dilution series and calibration curve.* J Immunol Methods, 2010. **353**(1-2): p. 148-50.

101. Chen, H.C., et al., *The studies on the transfection of EGFP-fused porcine circovirus Type 2 genome.* J Vet Med Sci, 2011. **73**(8): p. 1097-100.

102. Arrode-Bruses, G., et al., *Immunogenicity of a lentiviral-based DNA vaccine driven by the 5'LTR of the naturally attenuated caprine arthritis encephalitis virus (CAEV) in mice and macaques.* Vaccine, 2012. **30**(19): p. 2956-62.

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

104. Diogo, M.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.

105. Soares, A., et al., *Purification of human papillomavirus 16 E6/E7 plasmid deoxyribonucleic acid-based vaccine using an arginine modified monolithic support*. Journal of Chromatography A, 2013. **1320**: p. 72-79.

106. Mota, E., et al., *Rapid quantification of supercoiled plasmid deoxyribonucleic acid using a monolithic ion exchanger*. J Chromatogr A, 2013. **1291**: p. 114-21.

107. Sousa, A., et al., *Histamine monolith versatility to purify supercoiled plasmid deoxyribonucleic acid from Escherichia coli lysate*. Journal of Chromatography A, 2014.

108. Sousa, A., et al., *Performance of a non-grafted monolithic support for purification of supercoiled plasmid DNA*. J Chromatogr A, 2011. **1218**(13): p. 1701-6.

109. Williams, M.C., et al., *Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting*. Biophys J, 2001. **80**(2): p. 874-81.

110. Lee, D., J. Lee, and C. Seok, *What stabilizes close arginine pairing in proteins?* Physical Chemistry Chemical Physics, 2013. **15**(16): p. 5844-5853.

111. Mallik, R. and D.S. Hage, *Affinity monolith chromatography*. J Sep Sci, 2006. **29**(12): p. 1686-704.

112. Myers, J.L., A. Well, and R.F. Lorch, *Research Design and Statistical Analysis*. 2010: Routledge.

113. Ferreira, S., et al., *Response surface optimization of enzymatic hydrolysis of Cistus ladanifer and Cytisus striatus for bioethanol production*. Biochemical Engineering Journal, 2009. **45**(3): p. 192-200.

114. Haaland, P.D., *Experimental Design in Biotechnology*. 1989: Taylor & Francis.

115. Eriksson, L., *Design of Experiments: Principles and Applications*. 2008: Umetrics.

116. Patil, S.A., et al., *Optimization of medium using response surface methodology for l-DOPA production by Pseudomonas sp. SSA*. Biochemical Engineering Journal, 2013. **74**(0): p. 36-45.

117. Tjur, T., *Analysis of Variance: Why it is More Important Than Ever*. 2005.