Journal of Chromatography B 1105 (2019) 184–192

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/jchromb

DoE to improve supercoiled p53-pDNA purification by *O*-phospho-L-tyrosine chromatography

J.F.A. Valente, A. Sousa, J.A. Queiroz, F. Sousa*

CICS-UBI - Health Sciences Research Centre, Universidade da Beira Interior, Avenida Infante D. Henrique, 6200-506 Covilhã, Portugal

ARTICLE INFO

Keywords:
Composite Central Face design
Design of Experiments
O-Phospho-L-tyrosine chromatography
Supercoiled p53-encoding plasmid

ABSTRACT

P53 is implicated in various cellular functions and several studies have shown that transfection of cancer cells with wild-type p53-expressing plasmids could directly drive cells into growth arrest and/or apoptosis. In the present work, the 6.07 kbp pcDNA3-FLAG-p53 plasmid, which encodes the p53 tumor suppressor, was produced and recovered from a recombinant cell culture of *Escherichia coli* DH5α. Following plasmid biosynthesis, the *O*-phospho-L-tyrosine chromatographic matrix was explored to purify the supercoiled p53-encoding plasmid. In order to quickly determine the optimal chromatographic performance and to obtain the required purity degree, maximizing the recovery yield of the supercoiled plasmid DNA, the Composite Central Face design was applied. The model revealed to be statistically significant (p-value < 0.05), with coefficient of determination of 0.9434 for the recovery yield and 0.9581 for purity and the central point was successfully validated. After the chromatographic process optimization by using the design of experiments tool, 49.7% of the supercoiled p53-encoding plasmid was recovered with 98.2% of purity, when a decreasing ammonium sulphate gradient was applied. The dynamic binding capacity of the *O*-phospho-L-tyrosine agarose column was 0.35 ± 0.02 mg pDNA/mL matrix at 50% of the breakthrough. Finally, the purified sample was analysed to assess the content of endotoxins, proteins and genomic DNA, showing that all these impurity levels were below the recommendations of the regulatory agencies.

1. Introduction

The TP53 gene, which encodes p53, is one of the most frequently mutated genes in human cancers. It is reported that approximately half of all cancers have the p53 inactivated. This tumor suppressor is implicated in a number of known cellular functions, including the DNA damage repair, the cell cycle arrest in G1/S and apoptosis [1]. Several studies have shown that cells transfection with wild-type p53-expressing plasmids could directly drive cells into apoptosis and/or growth arrest when p53 is overexpressed, suggesting that the gene therapy approach for cancer treatment can re-establish the normal p53 function [2].

Recent advances have led to many different strategies to p53-targeted cancer therapy, including TP53 gene therapy, p53 vaccines, and rescue of mutant p53 function by small-molecule inhibitors. The cases of TP53 gene therapy and p53 vaccines have been studied extensively in patients with solid cancers [3–5]. In fact, several preclinical studies have shown that adenoviral vectors engineered with a complementary DNA copy of the p53 gene trigger a dramatic tumor regression response in cancers, such as head and neck cancer [6], lung cancer [7], colorectal cancer [8], among others [9].

For gene therapy purposes, the use of plasmids to replace or correct a defective gene has been evaluated. Plasmid DNA (pDNA) is the non-biologic vector mostly used in DNA-based therapies, due to its simplicity, low-cost manufacture, acceptable expression levels and lack of immunogenicity [10,11]. Among the different conformations that pDNA can acquire, it has been proved that the supercoiled isoform (sc) is more advantageous as it is more effective for gene transfection [12,13]. Also, in accordance with regulatory agencies, such as Food and Drug Administration (FDA) or the European Medicines Agency (EMA), pDNA should accomplish several specifications regarding its purity and global quality, in order to be suitable for therapeutic applications [14,15].

To isolate and purify the sc pDNA vector, several chromatographic methodologies have been used, namely size exclusion, anion exchange, hydrophobic interaction, reversed phase, thiophilic adsorption and affinity chromatography [14]. In particular, the affinity chromatography has already been explored to successfully isolate the sc p53 encoding plasmid by using matrices modified with arginine and methionine ligands. In these studies, the arginine support allowed the recovery of sc pDNA with a purity higher than 95%, while 97% was obtained with the

* Corresponding author at: Av. Infante D. Henrique, 6200-506 Covilhã, Portugal.
E-mail address: fani.sousa@fcsaude.ubi.pt (F. Sousa).

https://doi.org/10.1016/j.jchromb.2018.12.002
Received 4 May 2018; Received in revised form 5 November 2018; Accepted 2 December 2018
Available online 15 December 2018
1570-0232/

methionine matrix [12,16]. However, the recovery yield is usually low-to-moderate and the methodology commonly used to optimize the ideal elution conditions is based on trial and error approaches, which implies a huge number of experiments until the suitable conditions for the sc pDNA isolation are reached. Regarding this, it is still essential to develop and study new matrices and methods to accomplish the sc pDNA purification with higher specificity, recovery yield, robustness, in a more cost-effective manner to be considered for biopharmaceutical industry.

In the present work, it was explored the *O*-phospho-L-tyrosine chromatographic matrix to purify the sc p53-encoding pDNA. This affinity matrix was already studied in our research group for RNA isolation, showing an interesting potential for the purification of nucleic acids [17]. To accomplish this aim, the Composite Central Face design (CCF) was applied to optimize the chromatographic process regarding the maximization of pDNA purity and recovery yield. The design of experiments (DoE) tool enables to systematically and simultaneously vary several parameters in order to obtain more information about the purification process with few experiments. Thus, we are able to faster reach the best purification conditions, to accomplish the final result, which allows a reduction of the costs associated to the random experiment approach in the search for the suitable isolation conditions.

2. Experimental section

2.1. Materials

O-phospho-L-tyrosine agarose matrix was obtained from Sigma Aldrich (St. Louis, M.O., USA) and the NZY MaxiPrep Kit was purchased from NZYTech (Lisbon, Portugal). All the water used to prepare solutions was ultra-pure grade, purified with a Milli-Q system from Millipore. Ammonium sulphate ($(NH_4)_2SO_4$) was purchased from VWR and tris(hydroxymethyl) aminomethane (Tris) was from Merck (Darmstadt, Germany). Binding and elution buffers were filtered through a 0.20 µm pore size membrane (Schleicher Schuell, Dassel Germany) and degassed ultrasonically. The 6.07 kbp pcDNA3-FLAG-p53 (ref. 10838) plasmid [18] was purchased from Addgene, and all the reagents used in the bacterial growth were obtained from Sigma-Aldrich. The DNA ladder was obtained from Bioline (London, UK).

2.2. Methods

2.2.1. Plasmid and bacterial growth conditions

The pcDNA3-FLAG-p53 plasmid was amplified in a cell culture of *Escherichia coli* (*E. coli* DH5α). The growth was carried out at 37 °C, 250 rpm, in an Erlenmeyer with 500 mL of Terrific Broth medium (20 g/L of tryptone, 24 g/L of yeast extract, 4 mL/L of glycerol, 0.017 M KH_2HPO_4, 0.072 M K_2HPO_4) supplemented with 30 µg L^{-1} of ampicillin. The cells were grown until the log phase (O.D. 600 nm ± 9) and then were collected by centrifugation and stored at −20 °C.

2.2.2. Lysis and plasmid isolation

In order to study a complex lysate sample, after cells recovery, they were lysed through a modified alkaline method previously described. Completed the lysis procedure, and expecting the elimination of some impurities, the pDNA was first precipitated with isopropanol and then, an ammonium sulphate precipitation step was performed to eliminate proteins and RNA [15,19].

2.2.3. Preparative chromatography

Chromatographic assays were performed in an ÄKTA pure system with UNICORN 6.3 software (GE Healthcare, Uppsala, Sweden). To perform these experiments, a 16 × 40 mm (approximately 8 mL) column was packed with a commercial *O*-phospho-L-tyrosine agarose gel (Sigma Aldrich). This matrix is characterized by the manufacturer as a crosslinked 4% beaded agarose support with a one-atom spacer and an extent of labelling between 5 and 15 µmol/mL. The experiments were performed by using a circulating water bath to control the temperature. After the matrix equilibration, the lysate was loaded onto the column using a 100 µL loop, at 1 mL/min. The absorbance of the eluate was continuously monitored at 260 nm. The column was equilibrated with different concentrations of ammonium sulphate depending on the experiment. The elution of bound species was carried out by using a decreasing ammonium sulphate stepwise gradient.

2.2.4. Design of experiments

To optimize the sc pDNA purity and also maximize its recovery yield, a CCF design was applied. Concerning that, three factors (inputs) were taken into account, namely i) the sc binding step: to efficiently bind the sc pDNA to the column and at the same time eliminate some of the impurities; ii) the sc elution step: to elute most of sc pDNA with the required purity degree; iii) the Temperature. The inputs were studied at three levels (−1; 0; +1) and the range was defined according to preliminary results obtained from the random experiments approach. The responses (outputs) evaluated were the % of purity, that should be higher than 97% of sc pDNA and the % of recovery yield, that should be maximized maintaining the required purity.

To perform the chromatographic experiments, the column was equilibrated within the binding concentration range of 2.3 M to 2.6 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl buffer pH 8, the elution step was performed with lower ionic strength within the range of 1.8 M to 2.1 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl buffer pH 8. A last step consisting on the application of 10 mM of Tris–HCl buffer pH 8 was always performed to guarantee the total elution of the remaining bound species. The Temperature of the different chromatographic experiments was changed in the range between 4 and 20 °C. Each peak was pooled and concentrated to a final volume of 500 µL. Statistical analysis was performed through the use of UNICORN™ 6.3 software and Design Expert version 7.0 trial software. The generalized second-order polynomial model equation used in the response surface analysis is presented below (Eq. (1)):

$$\mathbf{Y} = \beta_1 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + X_{23} X_2 X_3 \quad (1)$$

2.2.5. Dynamic binding capacity of plasmid DNA

A 10 × 7 mm column (GE Healthcare Biosciences, Uppsala, Sweden) was packed with the commercial *O*-phospho-L-tyrosine agarose gel giving a total bed volume of 0.5 mL. The column was connected to an ÄKTA Pure system (GE Healthcare Biosciences, Uppsala, Sweden), and used for determination of the dynamic binding capacity (DBC) of the support. The breakthrough experiments were conducted at 1 mL/min and by using a pDNA feed solution with 0.025 mg/mL. The column was equilibrated with 3 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl buffer pH 8.0 and the pDNA solution was prepared in the same buffer. The determination of DBC was carried out by recording breakthrough curves and calculating the amount of bound pDNA per mL support at 10 and 50% breakthrough. Thus, the DBC value was obtained according to the following equation (Eq. (2)):

$$\mathrm{DBC} = \frac{(\mathrm{VL} - \mathrm{V0}) \times \mathrm{Cp}}{Vc} \quad (2)$$

where DBC is the dynamic binding capacity (mg/mL), VL corresponds to the volume loaded up to the breakthrough point (mL), V0 is the void volume of the column (mL), Cp corresponds to the pDNA concentration (mg/mL) and Vc is the column volume (mL). The elution of the bound pDNA was then achieved by using 10 mM Tris–HCl buffer pH 8.0 in the mobile phase.

2.2.6. Plasmid quantification

A modified quantification method using a CIMAc™ pDNA analytical column was applied to determine the sc pDNA purity and recovery yield [20,21]. The calibration curve was constructed in the range of 1–25 μg/mL, by considering the peak area corresponding to sc pDNA standards. The recovery yield was calculated by the ratio between the obtained sc pDNA concentration and the sc pDNA concentration present in the lysate sample, while purity was calculated by the ratio between sc pDNA peak area and total peak area present in the analytic chromatogram.

2.2.7. Protein quantification

For protein quantification, the micro-BCA (Bicinchoninic acid) protein assay kit from Pierce (Rockford, USA) was used in accordance to the supplier's specifications. A calibration curve with the standard protein Bovine Serum Albumin (BSA) (0.01–0.1 mg/mL) diluted in 10 mM of Tris-HCl pH 8.0 was determined. Before analysis, samples were desalted with Tris-HCl pH 8.0 to avoid salt interference.

2.2.8. Genomic DNA quantification

Real-time polymerase chain reaction (PCR) was used to evaluate the level of genomic DNA (gDNA) in the pDNA purified samples. The analysis was performed in an iQ5 Multicolor real-time PCR Detection System (Bio-Rad), as previously described [19]. Sense (5′-ACACGGTC CAGAACTCCTACG-3′) and antisense (5′-CCGGTGCTTCTTCTGCGGGT AACGTCA-3′) primers were used to amplify a 181-bp fragment of the 16S rRNA gene. For the gDNA quantification in lysate samples, a 100-fold dilution was prepared. PCR amplicons were quantified by following the change in fluorescence of the DNA binding dye Syber Green (Bio-Rad). The concentration of gDNA was obtained through a calibration curve generated by consecutive dilutions of the purified *E. coli* DH5α gDNA in the 5 pg to 50 ng/μL range. Negative controls (no template) were run at the same time of each analysis.

2.2.9. Endotoxin quantification

The ToxiSensor™ Chromogenic Limulus Amoebocyte Lysate (LAL) endotoxin assay kit (GenScript, USA, Inc.) was used for endotoxins measurement, considering the manufacturer's instructions. The calibration curve (from 0.005 to 0.1 EU/mL) was performed using a provided stock solution of 8 EU/mL. To avoid external endotoxin interference, all samples were diluted or dissolved in non-pyrogenic water, which was also used as the blank. All the tubes and tips used to perform this analysis were endotoxins-free. All the procedure was performed inside of a laminar flow cabinet to assure aseptic conditions.

3. Results and discussion

O-phospho-L-tyrosine ligand can promote a multitude of interactions (hydrophobic, electrostatic, cation-π, van der waals forces and/or hydrogen bond) that can be established between the ligand and different biomolecules [17]. Therefore, the multiple non-covalent interactions that are established, are actually responsible for the biorecognition and specific interaction occurring with the target molecule, namely the sc pDNA. Also, DNA-amino acid affinity chromatography for the purification of sc pDNA, using amino acids as specific ligands are already several times successfully applied [11,15,16]. The rationale for the implementation of this concept was based on the analysis of the natural occurring interactions, in cellular environment, where some target DNA or RNA sequences are recognized by specific protein domains, enabling the occurrence of specific cell mechanisms. Thus, the use of amino acids as ligands can mimic this specific interaction, the recognition of the sc pDNA based on affinity approaches [22,23]. Moreover, the authors have already demonstrated that with these affinity ligands it is possible to favor some particular interactions through the adjustment of the binding and elution conditions, such as the temperature and buffer composition (pH, ionic strength or presence of competitive agents), giving more versatility to the method [14,24].

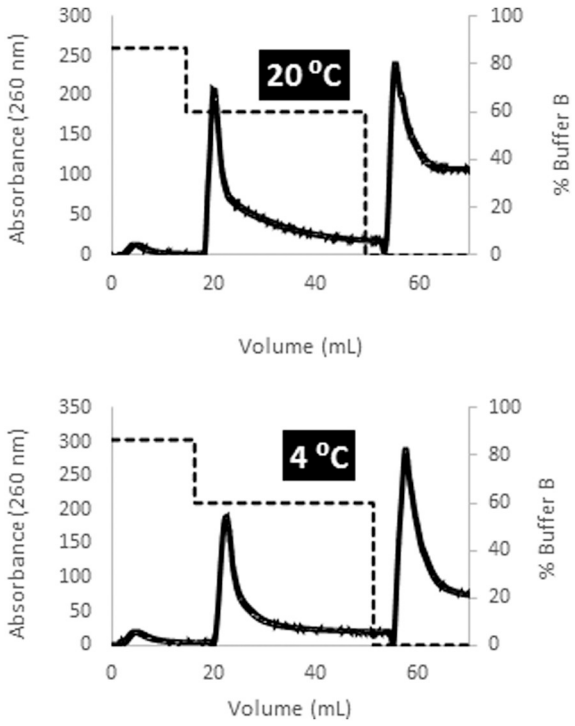

Fig. 1. Chromatographic profiles of the p53 encoding plasmid lysate loaded onto the *O*-phospho-L-tyrosine agarose column at 4 and 20 °C. The elution was performed by a stepwise gradient of 2.6 M of $(NH_4)_2SO_4$ in 10 mM Tris-HCl pH 8.0, then 1.8 M of $(NH_4)_2SO_4$ in in the same buffer and finally 10 mM Tris-HCl pH 8.0, as represented by the dashed line.

In this work, tyrosine amino acid was selected as specific ligand to isolate the sc p53-encoding pDNA isoform from a complex lysate. To accomplish this, different conditions, such as buffer composition, pH and Temperature were tested in a random experiment approach to define the first binding/elution settings. Thus, pDNA retention was first evaluated in low ionic strength conditions, but the loaded sample immediately eluted during the flowthrough (data not shown). Conversely, total pDNA retention was achieved by using ammonium sulphate in the binding buffer and its elution was verified by decreasing the salt concentration. This behaviour suggests that hydrophobic interactions between pDNA and *O*-phospho-L-tyrosine are favoured, mainly due to the existence of a benzyl aromatic ring in the ligand [25].

Temperature was also a parameter under evaluation, being verified that it definitely plays an important role in sc pDNA isolation in the *O*-phospho-L-tyrosine agarose matrix (Fig. 1). Concerning this, a lysate sample made of oc and sc pDNA and also RNA was injected in the matrix and, the results have shown a decrease on pDNA binding with the temperature increase, also indicating that the pDNA binding is not exclusively due to hydrophobic interactions, which typically are favoured by the temperature increase. In fact, the biorecognition results from the establishment and conjugation of multiple non-covalent elementary interactions, such as hydrogen bonds, van der Waals contacts and hydrophobic interactions [16]. The last peak of the chromatogram is supposed to be mostly made of RNA since this specie strongly binds to the matrix.

3.1. Design of experiments

The DoE model used in this work was the CCF, mainly because it is more inclusive and does not consider points outside of the ranges established for the inputs (starting points) [17]. Considering the aim of this work and expecting a pDNA quality in accordance with the regulatory agencies criteria, the responses (outputs) chosen were the % of

Table 1
Summary of the multiple regression equation (main effects of the factors).

Parameter	Recovery effect	Purity effect
sc binding (A)	Positive	Positive
sc Elution (B)	Negative	Negative
Temperature (C)	Negative	Positive
sc binding × sc elution	Positive	Negative
sc binding × temperature	Positive	Negative
sc elution × temperature	Negative	Positive
sc binding^2	Positive	Positive
sc elution^2	Negative	Positive
Temperature^2	Positive	Negative

pDNA purity, that should be higher than 97%, and simultaneously, the maximization of the % of recovery yield.

According to preliminary studies, a chromatographic method was outlined by considering three steps. First, it was established a sc pDNA binding step, ranging between 2.3 M and 2.6 M of $(NH_4)_2SO_4$, that intends to efficiently bind this target molecule to the column and at the same time to eliminate some of the impurities in the flowthrough. Secondly, a sc elution step, ranged between 1.8 M and 2.1 M of $(NH_4)_2SO_4$, was performed to elute most of sc pDNA with the required purity degree. Finally, the last step at 10 mM of Tris buffer pH 8 was used as a washing step to elute the remaining bound impurities. Given that, the outputs are the sc pDNA purity and recovery yield, the binding step and sc elution step were chosen as inputs. In addition, it was also included the Temperature as an input (ranging between 4 and 20 °C), as the initial results indicated a significant influence of this parameter on pDNA retention and selectivity. The different experiments designed by UNICORN™ 6.3 software and the respective responses are available in supporting information (SI). The three central points are presented in grey (Table in SI), which were tested in the same conditions in order to access the model reproducibility. The regression Eqs. (3) and (4) are resultant from the Design of Expert 7.0 and provide the level of responses for recovery yield and purity as a function of the different variables chosen ("sc binding step", "sc elution step" and "Temperature"). The signal of each factor designates a positive or negative effect on the response.

$$\text{Recovery yield} = +40.53 + 13.63 \times A - 4.33 \times B - 3.4 \times C + 0.89 \times AB + 0.49 \times AC - 4.04 \times BC + 4.07 \times A^2 - 11.83 \times B^2 + 0.52 \times C^2 \quad (3)$$

$$\text{Purity} = +98.58 + 1.96 \times A - 1.56 \times B + 5.68 \times C - 0.69 \times AB - 3.56 \times AC + 1.59 \times BC + 1.56 \times A^2 + 2.06 \times B^2 - 9.24 \times C^2 \quad (4)$$

In Table 1 it is possible to observe the information extracted from the multiple regression equation, corresponding to the main effects of the chosen parameters.

From these results, it is possible to verify that the "sc binding step" condition is positive for both responses, which can indicate that higher ammonium sulphate concentrations intensify the interactions established with the sc pDNA, strengthening its retention on the matrix. Thus, the sc pDNA recovery yield and purity could be increased in the subsequent elution step. Concerning the "sc elution step", it is possible to verify negative responses either for recovery yield as well as for purity. This mean that if low ammonium sulphate concentrations were applied the matrix will retain not only the sc pDNA but also impurities and the opposite could also happen (when high concentrations of ammonium sulphate were applied the elution of the sc and impurities will be promoted). Which means that high recovery and purity levels are only obtained when a very specific salt concentration is applied, being not allowed large variations.

On the other hand, the input "Temperature" presents a different contribution for each response (negative for recovery yield and positive for purity). Actually, in accordance with what was previously observed in preliminary experiments, the decrease of Temperature seems to intensify the interactions between the ligand and the nucleic acids, which are only eluted in the washing step, favouring the purity but also decreasing the recovery yield of the sc pDNA in the elution step. These results show the influence of the salt concentration and Temperature in the purity and the recovery yield of sc pDNA for the elution step, being fundamental to determine the optimal point by a combination of these factors through an experimental design tool, aiming to maximize the intended responses.

Table 2
ANOVA table for CCF design model.

Source	Sum of squares	Degrees of freedom	Mean square	F-value	*p*-Value
Recovery ANOVA analysis					
Model	2724.25	9	302.69	12.95	0.0014
A - sc binding	1857.77	1	1857.77	79.5	< 0.0001
B - sc elution	187.49	1	187.49	8.02	0.0253
C - Temperature	115.6	1	115.6	4.95	0.0615
AB	6.3	1	6.3	0.27	0.6196
AC	1.9	1	1.9	0.081	0.7837
BC	130.41	1	130.41	5.58	0.0502
A^2	44.34	1	44.34	1.9	0.2108
B^2	375.06	1	375.06	16.05	0.0051
C^2	0.72	1	0.72	0.031	0.8657
Residual	163.57	7	23.37		
Lack of fit	71.86	5	14.37	0.31	0.8721
Pure error	91.71	2	45.85		
Cor total	2887.82	16			
Purity ANOVA analysis					
Model	761.76	9	84.64	17.78	0.0005
A - sc binding	38.49	1	38.49	8.09	0.0249
B - sc elution	24.4	1	24.4	5.13	0.058
C - Temperature	322.85	1	322.85	67.83	< 0.0001
AB	3.81	1	3.81	0.8	0.4008
AC	101.39	1	101.39	21.3	0.0024
BC	20.1	1	20.1	4.22	0.079
A^2	6.54	1	6.54	1.38	0.2793
B^2	11.4	1	11.4	2.4	0.1656
C^2	228.6	1	228.6	48.03	0.0002
Residual	33.32	7	4.76		
Lack of fit	32.51	5	6.5	16.12	0.0594
Pure error	0.81	2	0.4		
Cor total	795.08	16			
Model summary statistics					
Response	R^2	Adj R^2	Adeq Precision	Q^2	
Recovery	0.9434	0.8705	13.949	0.7837	
Purity	0.9581	0.904	13.684	0.7	

From the interaction of the parameters "sc binding × sc elution", it is possible to observe a positive effect for recovery and a negative effect for purity which means that a decrease in the sc pDNA binding and an increase in its elution will allow an increase in the sc pDNA recovery. However, this behavior will be followed by the co-elution of impurities with the sc pDNA, thereby decreasing the purity of the sample. Also, when the "sc binding × Temperature" parameters were combined, a positive response for recovery and a negative response for purity was observed. This can indicate that a Temperature increase will increase the retention of all species (sc pDNA and impurities) which could be responsible for a decreased purity of the recovered sample. Finally, the interaction between "sc elution × Temperature" is negative for the recovery and positive for the purity. This means that an increase of Temperature in the elution step will increase sample retention, which will decrease the recovery of the target molecule and, consequently, increase the sample purity. Regarding the quadratic effect, it is positive for recovery in the case of "sc binding^2" or "Temperature^2" and negative for "sc elution^2", while the purity effect is positive for "sc binding^2" and "sc elution^2" and negative for "Temperature^2".

To evaluate the presence of potential outliers the Cooks Distance analysis was performed concluding their inexistence [26]. Moreover, the evaluation of the significance of the model was based on the analysis of the variance (ANOVA), as it is presented in Table 2. Closely looking for this analysis, it is possible to confirm that the model data are statistically significant for both responses, since the probability value (*p*-value) was always lower than 0.05 for purity and recovery yield [27,28]. Also, for recovery yield, the "Model F-value" was 12.95, with only 0.14% of chance of it occur due to the noise and, in case of purity, the "Model F-value" was 17.78 having 0.05% of chance to occur due to the noise.

The Lack of Fit is responsible to describe if the model is appropriate to correctly express data or if a more complex model should be applied. To achieve this parameter, a comparison between the variability of the current model residuals and the variability between observations and replicative settings of the factors was used [29]. The "Lack of Fit F-value" to the output Recovery yield was 0.31, implying that the Lack of Fit is not significant relatively to the pure error, and there is 87.21% of chance that a "Lack of Fit F-value" this large could occur due to noise. This Lack of fit showed to be non-significant which is a positive result since we want that the model fits. In case of the output Purity, the "Lack of Fit F-value" was 16.12, which implies 5.94% of chance for this large "Lack of Fit F-Value" occur due to the noise. Also, in this case, this parameter showed to be non-significant.

The coefficient of determination (R^2) shows if data points fit the statistical model, indicating the high significance of the model. Although, being acceptable a value ranging between 0 and 1, ideally it should be closer to 1 [28]. In this case, the R^2 were 0.9434 and 0.9581 for recovery yield and purity respectively, suggesting that both responses fit the models. Adjusted R^2 (Adj R^2) comprises the variation of the ordinary R^2 and is adjusted for the "size" of the model (the number of the factors) [17]. The results obtained from the ANOVA analysis present good Adj R^2 for both responses.

Moreover, the usefulness of the model to predict new data is associated with positive predicted values variation (Q^2), while negative Q^2 is related to models that do not have a predicted power [30]. The Q^2 data for purity (0.7) and for recovery yield (0.7837) responses presented positive values, suggesting that the model can predict new data. However, the value of Q^2 for purity was lower than the value of Q^2 for recovery yield, indicating that the prediction will be better for the recovery yield response. This result does not demerit the present work, since it is intended to maximize the sc pDNA recovery yield, maintaining the pDNA purity higher than 97%. Actually, this affinity chromatographic methodology based on amino acids as specific ligands already proved to be efficient on sc pDNA purification, while the major drawback is usually the recovery yield. For example, in a previous purification strategy of sc p53-encoding pDNA by using the methionine matrix, the final pDNA recovery was very low, because it was necessary to reduce the recovery in order to guarantee the pDNA purity [16]. So, the present results indicate that the use of DoE tool can help in the definition of the best experimental conditions, to improve the recovery yield and purification of pDNA. Finally, the "Adequate Precision" is the parameter responsible for measuring the signal-to-noise ratio, being desirable a value greater than 4. The results obtained in this field were 13.949 and 13.684 for the recovery yield and purity respectively, suggesting that both models can be used to navigate the design space.

3.2. Three-dimensional response surface plots

The surface and contour plots show the three-dimensional representation of the multiple regression equations [29]. By the analysis of the plots illustrated in Fig. 2, it is easily observed the relation between the chosen parameters (sc binding step, sc elution step and Temperature) and the outputs (recovery yield and purity). The different colours intensity represents the range for optimal points. Thus, blue represents the lowest response followed by green, yellow and finally red, which represents the largest interaction. At the same time, the ellipticity obtained in the plots is indicative of the interaction order that occurs between the chosen variables. In general, by analysing the colour and ellipticity, it is visible that Temperature has more influence in the purity response (Fig. 3) than in the recovery yield response (Fig. 2). Also, through the ellipticity and colour of the recovery yield contour plots, it is possible to verify that the interaction between sc elution step versus sc binding step (Fig. 2A) or Temperature versus sc binding step (Fig. 2B) is more pronounced than in the case of Temperature versus sc elution step (Fig. 2C). These results indicate that Temperature has a less pronounced effect on the sc elution and may influence in a more effective way its binding to the matrix. In purity plots (Fig. 3), it is depicted that all the variables chosen for the design have a crucial role, being red the most observed colour, indicating a very expressive interaction between sc binding step versus sc elution step (Fig. 3A). In fact, to maximize the sc pDNA purity it should be found a compromise between the salt concentration of the sc binding step and the sc elution step.

3.3. Model validation

After the analysis of the responses given by the experiments provided by DoE, it was predicted the optimal point, consisting in the best conditions to achieve the global aim of this study, which is the sc p53-encoding pDNA preparation with a purity degree higher than 97% and maximizing the recovery yield. So, the best conditions, as well as the obtained responses, are presented in Table 3. Therefore, the optimal conditions suggested by the software (2.58 M $(NH_4)_2SO_4$ in 10 mM Tris-HCl buffer pH 8, for the sc binding step, 1.92 M of $(NH_4)_2SO_4$ in 10 mM Tris-HCl buffer pH 8 for the sc elution step and 9.5 °C for the Temperature) were used to perform the experiment in triplicate. All chromatograms revealed the same profile, suggesting that the method was suitable and reproducible. Fig. 4 shows one representative chromatogram from the triplicate assays that illustrate the total retention of the target molecule in the binding step, the selective isolation of sc p53-encoding plasmid in the second step (peak 2, lane 2) and finally the elution of impurities such as RNA (peak 3, lane 3) in the washing step at 10 mM Tris-HCl buffer pH 8. It is important to notice that the purification of sc pDNA was achieved from a complex lysate containing gDNA, oc isoform, sc isoform, RNA and also proteins and endotoxins. From the analysis of the electrophoretic profile of Fig. 4 it is also visible that the sc pDNA is recovered in peak 2 with high purity level, but a small amount is also eluting in the final step, which justifies the obtained recovery yield. Triplicates samples of peak 2 were analysed by the CIMac pDNA analytic column and the mean value for purity was 98.2% with a standard deviation of 1.3% and for recover was 49.7% with a standard deviation of 1.9%.

In Table 4 is represented the confidence intervals (CI) for both responses analysed. The 95% CI represent the range in which responses should lie 95% of the time. Regarding this information and considering the obtained results, the 95% CI for the recovery yield ranged from 49.6 to 63.3% and for purity ranged from 97.4 to 103.6%. Comparing these intervals with the obtained values for the recovery yield and also purity it is possible to conclude that they are between the predicted CI. The purity values obtained are in accordance with the recommendations of regulatory agencies and, comparing the recovery yield of sc p53-encoding pDNA with the values obtained with other matrices, like L-methionine that presented 8% of sc pDNA recovery yield, it is possible to conclude that the *O*-phospho-L-tyrosine matrix was significantly more efficient [16]. These results suggest a huge potential of this matrix to be integrated into a downstream platform in order to obtain this biopharmaceutical product for further application in gene therapy studies.

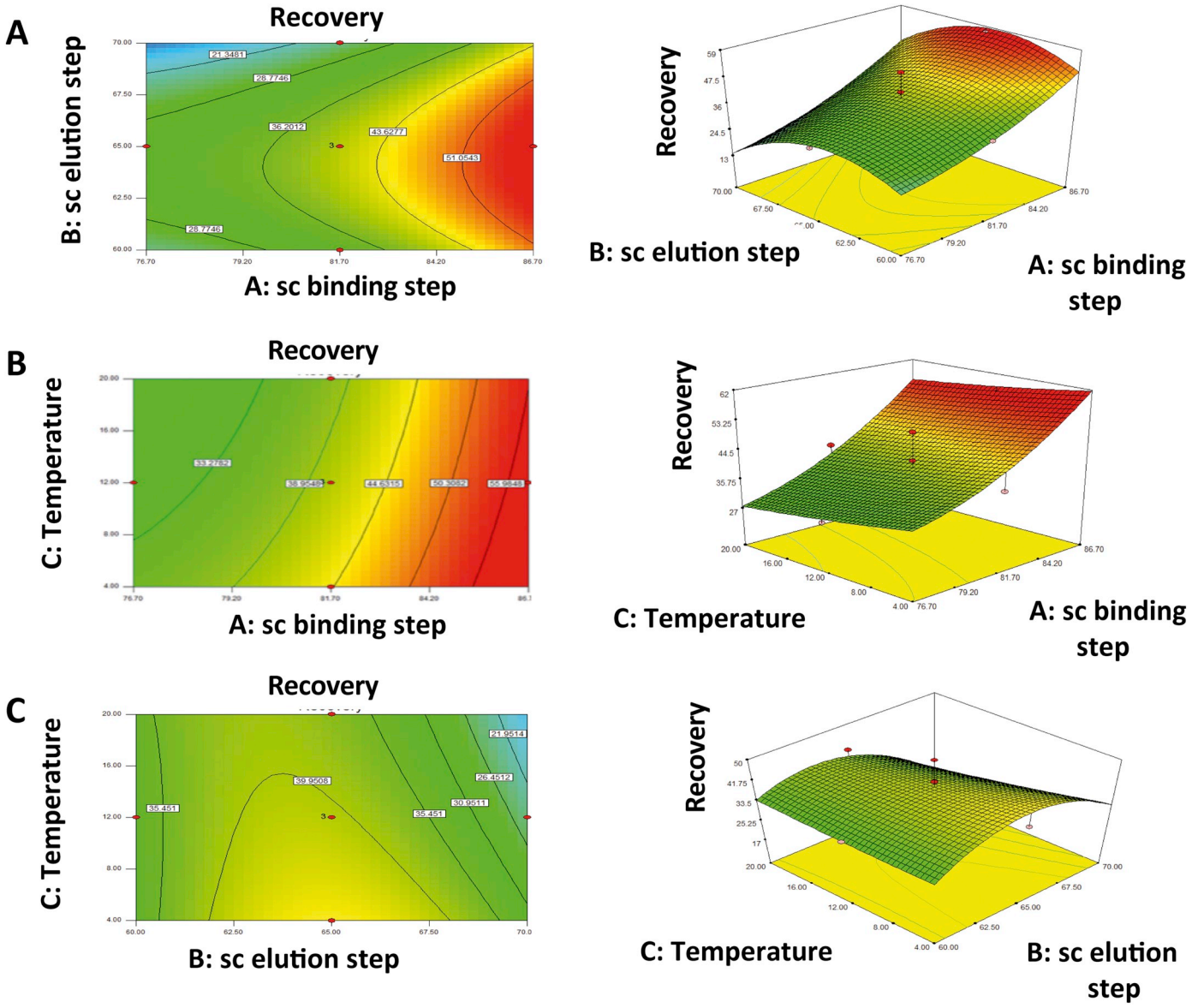

Fig. 2. Contour and three-dimensional response surface plot of the interaction of the different variables (binding sc; elution sc; Temperature) and its effect on % of recovery yield.

3.4. Dynamic binding capacity

Determining the DBC of a stationary phase is an important step for a complete characterization of the support. Regarding this, breakthrough experiments were performed at 1 mL/min, with 0.025 mg/mL pDNA solution containing 90% of sc conformation, pre-purified with NZYtech kit. The obtained results showed that the DBC of the *O*-phospho-L-tyrosine agarose column was 0.35 ± 0.02 mg/mL of the matrix at 50% of the breakthrough. In fact, comparing these results with previous studies of conventional amino acids-agarose matrices, where ammonium sulphate was applied to perform the sc pDNA isolation, it is possible to verify that this *O*-phospho-L-tyrosine matrix presents a similar binding capacity. For example, Sousa and co-workers (2007) characterized the capacity of the histidine-agarose matrix at 50% of breakthrough, using similar chromatographic conditions to the ones applied in this research work, obtaining a DBC near to 0.22 mg pDNA/mL of matrix [33]. The higher DBC of *O*-phospho-L-tyrosine matrix can be seen as an advantage if a scale-up of the pDNA purification process is intended, as an increased capacity could reduce the production runs and costs, increasing the amount of pDNA recovered per experiment and the feasibility of the downstream process.

3.5. p53-encoding plasmid quality assessment

Based on the fact that sc pDNA preparations have proved to be more efficient for cells transfection than other isoforms and, considering that the efficient access of pDNA to the cell nucleus improves gene expression in eukaryotic cells, it is extremely important to guarantee that the final sample is enriched in sc pDNA [13].

However, besides the requirement of 97% of sc homogeneity, the regulatory agencies also indicate other specifications that must be accomplished by pDNA to be suitable for therapeutic applications. In particular, pDNA solution must be clear and colorless, the RNA and proteins content should not be detectable, the amount of gDNA needs to stay below 2 μg gDNA/mg of pDNA and also the level of endotoxins should not exceed the 10 EU/mg of pDNA [31]. In this way, the second peak of the replicate chromatograms containing sc pDNA was analysed regarding these impurities. The results (Table 5) indicated that this new chromatographic method enables the recovery of sc pDNA free from proteins. Moreover, real-rime PCR was performed in order to quantify the gDNA present in these samples. This assessment is extremely important because some studies demonstrated that the existence of gDNA

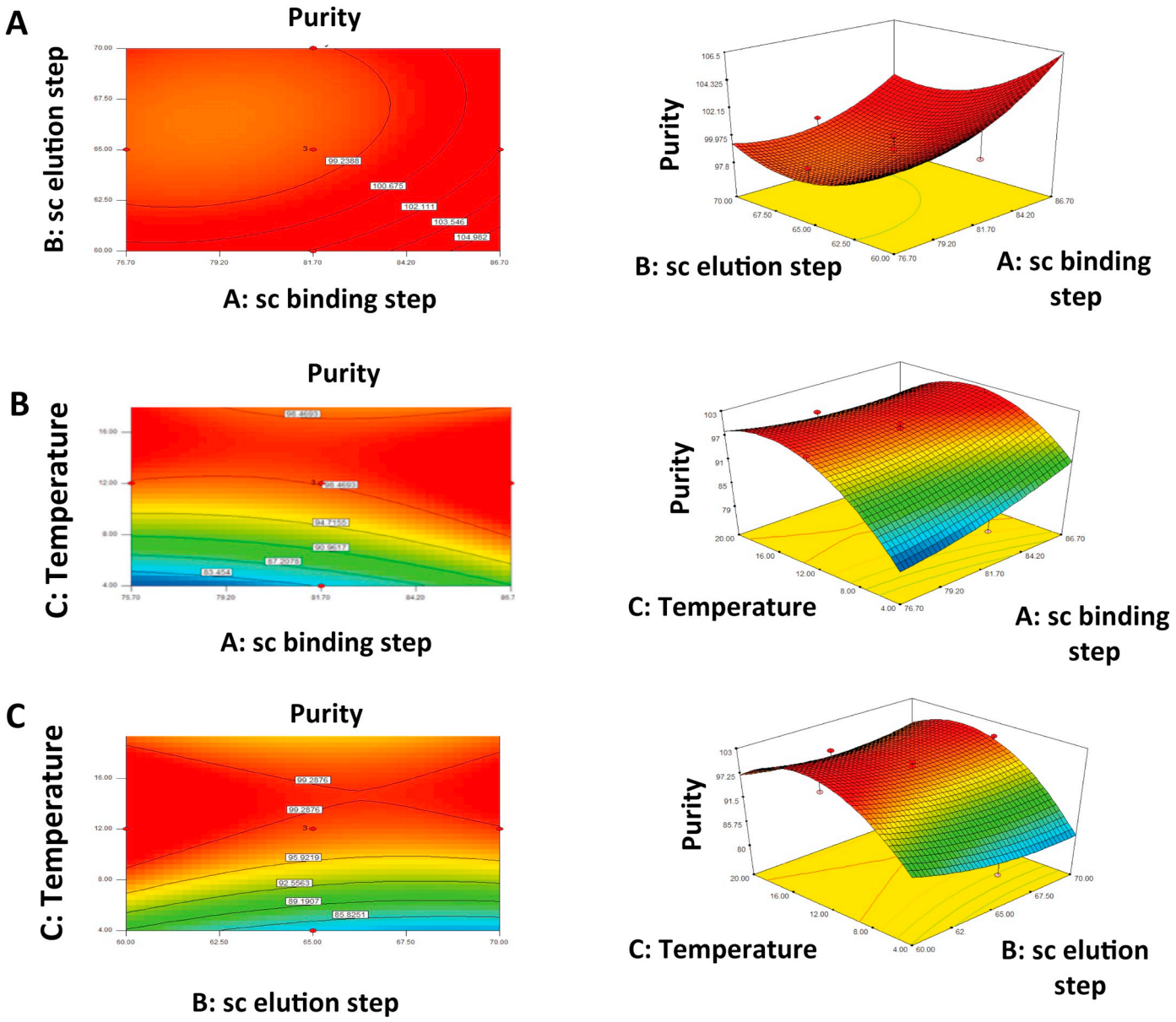

Fig. 3. Contour and three-dimensional response surface plot of the interaction of the different variables (binding sc; elution sc; Temperature) and its effect on % of purity.

Table 3
Optimal conditions selected by DoE to recover and purify sc p53-pDNA.

Predicted inputs	Value
Sc binding ($[(NH_4)_2SO_4]$ in 10 mM Tris-HCl buffer, pH 8)	2.58 M
Sc elution ($[(NH_4)_2SO_4]$ in 10 mM Tris-HCl buffer, pH 8)	1.92 M
Temperature	9.5 °C
Predicted output	Value
Recovery	56.5%
Purity	100%

in a purified plasmid sample for clinical application was directly related with necrosis of the muscle cells [32]. The gDNA quantification demonstrated that the amount of gDNA in the sc peak was below the limit given by the regulations (Table 5). Finally, the endotoxins content in the sc pDNA preparations was assessed through an LAL endotoxin kit, and once again the values were below the recommended limit given by the regulatory agencies.

4. Conclusions

This research work used *O*-phospho-L-tyrosine combined with CCF design in order to achieve the optimal conditions for the isolation of the sc p53-encoding plasmid. To perform that, three parameters were used and combined ("binding sc", "elution sc" and "Temperature") and finally the percentage of recovery yield and also purity was evaluated as responses. It was possible to see during the experiments the interactions of all inputs and their influence on the final outputs. The model revealed to be statistically significant (p-value < 0.05), with coefficient of determination of 0.9434 for recovery yield and 0.9581 for purity and the central point was successful validated. The central point was repeated three times being achieved 49.7% of recovery yield and 98.2% of purity. Moreover, the DBC achieved for this matrix was 0.35 ± 0.02 mg pDNA/mL matrix at 50% of the breakthrough. Finally, the quality of the resulting sc pDNA was assessed, being confirmed that the levels of impurities are in agreement with all the parameters required by the regulatory agencies such as EMA and FDA.

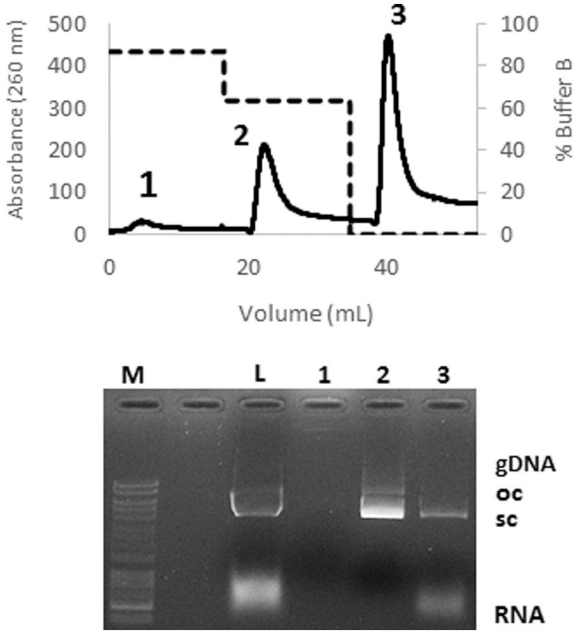

Fig. 4. Chromatographic profile of the sc pDNA purification from an *E. coli* lysate by using the *O*-phospho-L-tyrosine agarose column, at 9.5 °C. The different species were eluted from the column by a decreasing stepwise gradient of ammonium sulphate, as it is presented by the dashed line in the chromatogram. In the agarose gel electrophoresis, the lane M represents the molecular weight marker; lane L represents the lysate sample; and lanes 1, 2 and 3 represent the first, second and third peaks of the chromatogram, respectively.

Table 4
Confidence intervals for the two responses.

Response	95% CI low	95% CI high
Recovery	49.6	63.3
Purity	97.4	103.6

Table 5
Quantitative analysis of proteins, endotoxins, and gDNA in the purified sc fractions.

	Proteins (μg/mL)	Endotoxins (EU/mg pDNA)	gDNA (μg gDNA/mg DNA)
sc elution peak	Undetectable	8.5	1.1

Overall, DoE demonstrated to be a very powerful and quick tool for the establishment of the optimal conditions for the sc p53-encoding plasmid purification. In addition, it was established a purification methodology based on the *O*-phospho-L-tyrosine matrix that allows an improvement of the recovery yield of this plasmid, being a promising tool for the downstream process of this biopharmaceutical product.

Declarations of interest

None.

Acknowledgment

The authors would like to thank Dr. Thomas Roberts for providing the pcDNA3–FLAG–p53 construct through Addgene, ref.: 10838 and to Patricia Pereira for her help in the real-time PCR study. This work was supported by FEDER funds through the POCI - COMPETE 2020 - Operational Programme Competitiveness and Internationalization in Axis I - Strengthening research, technological development and innovation (Project POCI-01-0145-FEDER-007491) and National Funds by FCT - Foundation for Science and Technology (Project UID/Multi/00709/2013). J.F.A. Valente and A. Sousa also acknowledge PhD and Postdoctoral fellowships (Ref SFRH/BD/96809/2013 and Ref SFRH/BPD/102716/2014, respectively).

References

[1] A. Rufini, P. Tucci, I. Celardo, G. Melino, Senescence and aging: the critical roles of p53, Oncogene 32 (2013) 5129.
[2] R. Petty, A. Evans, I. Duncan, C. Kurbacher, I. Cree, Drug resistance in ovarian cancer—the role of p53, Pathol. Oncol. Res. 4 (1998) 97–102.
[3] G.-X. Chen, S. Zhang, X.-H. He, S.-Y. Liu, C. Ma, X.-P. Zou, Clinical utility of recombinant adenoviral human p53 gene therapy: current perspectives, OncoTargets Ther. 7 (2014) 1901.
[4] E.M. Dijkgraaf, S.J. Santegoets, A.K. Reyners, R. Goedemans, H.W. Nijman, M.I. van Poelgeest, A.R. van Erkel, V.T. Smit, T.A. Daemen, J.J. van der Hoeven, A phase 1/2 study combining gemcitabine, Pegintron and p53 SLP vaccine in patients with platinum-resistant ovarian cancer, Oncotarget 6 (2015) 32228.
[5] K. Kojima, J. Ishizawa, M. Andreeff, Pharmacological activation of wild-type p53 in the therapy of leukemia, Exp. Hematol. 44 (2016) 791–798.
[6] G.L. Clayman, A.K. El-Naggar, J.A. Roth, W.-W. Zhang, H. Goepfert, D.L. Taylor, T.-J. Liu, In vivo molecular therapy with p53 adenovirus for microscopic residual head and neck squamous carcinoma, Cancer Res. 55 (1995) 1–6.
[7] T. Fujiwara, E.A. Grimm, T. Mukhopadhyay, W.-W. Zhang, L.B. Owen-Schaub, J.A. Roth, Induction of chemosensitivity in human lung cancer cells in vivo by adenovirus-mediated transfer of the wild-type p53 gene, Cancer Res. 54 (1994) 2287–2291.
[8] F.R. Spitz, D. Nguyen, J.M. Skibber, J. Cusack, J.A. Roth, R.J. Cristiano, In vivo adenovirus-mediated p53 tumor suppressor gene therapy for colorectal cancer, Anticancer Res. 16 (1996) 3415–3422.
[9] J. Kim, E.S. Hwang, J.S. Kim, E.H. You, S.H. Lee, J.H. Lee, Intraperitoneal gene therapy with adenoviral-mediated p53 tumor suppressor gene for ovarian cancer model in nude mouse, Cancer Gene Ther. 6 (1999) 172–178.
[10] S.L. Ginn, I.E. Alexander, M.L. Edelstein, M.R. Abedi, J. Wixon, Gene therapy clinical trials worldwide to 2012–an update, J. Gene Med. 15 (2013) 65–77.
[11] A. Almeida, J. Queiroz, F. Sousa, A. Sousa, Optimization of supercoiled HPV-16 E6/E7 plasmid DNA purification with arginine monolith using design of experiments, J. Chromatogr. B 978 (2015) 145–150.
[12] V.M. Gaspar, I.J. Correia, Â. Sousa, F. Silva, C.M. Paquete, J.A. Queiroz, F. Sousa, Nanoparticle mediated delivery of pure P53 supercoiled plasmid DNA for gene therapy, J. Control. Release 156 (2011) 212–222.
[13] D. Costa, A.J. Valente, M.G. Miguel, J. Queiroz, Plasmid DNA microgels for drug/gene co-delivery: a promising approach for cancer therapy, Colloids Surf. A Physicochem. Eng. Asp. 442 (2014) 181–190.
[14] A. Sousa, F. Sousa, J.A. Queiroz, Differential interactions of plasmid DNA, RNA and genomic DNA with amino acid-based affinity matrices, J. Sep. Sci. 33 (2010) 2610–2618.
[15] A. Sousa, F. Sousa, J. Queiroz, Impact of lysine-affinity chromatography on supercoiled plasmid DNA purification, J. Chromatogr. B 879 (2011) 3507–3515.
[16] J. Valente, A. Sousa, J. Queiroz, F. Sousa, Selective purification of supercoiled p53-encoding pDNA with l-methionine–agarose matrix, Anal. Biochem. 459 (2014) 61–69.
[17] A. Afonso, P. Pereira, J.A. Queiroz, Â. Sousa, F. Sousa, Purification of pre-miR-29 by a new *O*-phospho-L-tyrosine affinity chromatographic strategy optimized using design of experiments, J. Chromatogr. A 1343 (2014) 119–127.
[18] O. Gjoerup, D. Zaveri, T.M. Roberts, Induction of p53-independent apoptosis by simian virus 40 small t antigen, J. Virol. 75 (2001) 9142–9155.
[19] F. Sousa, D.M. Prazeres, J.A. Queiroz, Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography, J. Gene Med. 11 (2009) 79–88.
[20] É. Mota, Â. Sousa, U. Černigoj, J.A. Queiroz, C.T. Tomaz, F. Sousa, Rapid quantification of supercoiled plasmid deoxyribonucleic acid using a monolithic ion exchanger, J. Chromatogr. A 1291 (2013) 114–121.
[21] A. Sousa, A. Almeida, U. Černigoj, F. Sousa, J. Queiroz, Histamine monolith versatility to purify supercoiled plasmid deoxyribonucleic acid from *Escherichia coli* lysate, J. Chromatogr. A 1355 (2014) 125–133.
[22] F. Sousa, D.M. Prazeres, J.A. Queiroz, Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA, Trends Biotechnol. 26 (2008) 518–525.
[23] Â. Sousa, F. Sousa, J.A. Queiroz, Advances in chromatographic supports for pharmaceutical-grade plasmid DNA purification, J. Sep. Sci. 35 (2012) 3046–3058.
[24] L.F. Amorim, R. Gaspar, P. Pereira, U. Černigoj, F. Sousa, J. Queiroz, Ã. Sousa, Chromatographic HPV-16 E6/E7 plasmid vaccine purification employing L-histidine and 1-benzyl-L-histidine affinity ligands, Electrophoresis 38 (2017) 2975–2980.
[25] J.-B. Ma, Y.-R. Yuan, G. Meister, Y. Pei, T. Tuschl, D.J. Patel, Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein, Nature 434 (2005) 666.
[26] T. Zhang, A. Sharma, Y. Li, Y. Zhou, X. Ding, Orthogonal Array composite design to study and optimize antioxidant combinations in the prevention of UVB-induced HSF damage, J. Photochem. Photobiol. B Biol. 178 (2018) 568–576.
[27] A. Gelman, Analysis of variance—why it is more important than ever, Ann. Stat. 33

(2005) 1–53.
[28] J.L. Myers, A. Well, R.F. Lorch, Research Design and Statistical Analysis, Routledge, 2010.
[29] D. Montgomery, Design and Analysis of Experiments, Fifth edition, By John Wiley & Sons, Inc., 2001 All rights reserved.
[30] L. Eriksson, E. Johansson, N. Kettaneh-Wold, C. Wikström, S. Wold, Design of Experiments: Principles and Applications Third Revised and Enlarged Edition, Umetrics Academy, Sweden, 2008.
[31] J. Stadler, R. Lemmens, T. Nyhammar, Plasmid DNA purification, J. Gene Med. 6 (2004) S54–S66.
[32] C.I. Wooddell, V.M. Subbotin, M.G. Sebestyén, J.B. Griffin, G. Zhang, M. Schleef, S. Braun, T. Huss, J.A. Wolff, Muscle damage after delivery of naked plasmid DNA into skeletal muscles is batch dependent, Human Gene Ther. 22 (2010) 225–235.
[33] F. Sousa, D.M.F. Prazeres, J.A. Queiroz, Dynamic binding capacity of plasmid DNA in histidine–agarose chromatography, Biomed. Chromatogr. 21 (9) (2007) 993–998.