The impact of *zwf* overexpression on plasmid biopharmaceutical production by *Escherichia coli*

Rita Azevedo Fernandes
Master Student in Biotechnology at Instituto Superior Técnico Universidade de Lisboa, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal

November, 2014

Abstract

Plasmid DNA vectors may find application as preventive or therapeutic DNA vaccines for viral, bacterial, or parasitic diseases or for other indications such as cancer, or gene therapy products. Therefore, the demand of pDNA is expected to rise significantly in the near future and the establishment of a high-yield and cost effective manufacturing process is one of the key challenges for the success of plasmid DNA market.

With the goal of creating an *E. coli* strain specifically adapted to meet the upstream and downstream processing challenges associated with large scale production of plasmid vectors, some mutations have been made in its genome. One main area of focus is modification of central carbon metabolism genes to increase flux toward nucleotide and amino acid precursor synthesis and reduction of byproducts, such as acetate. The *zwf* and *rpiA* are two important genes of the Pentose Phosphate Pathway and its overexpression was already tested and appeared to increase plasmid amplification rate. Therefore, the overexpression of the *zwf* and *rpiA* genes in GALG20 (MG1655*ΔendAΔrecAΔpgi*) is intended in this project, by replacing its original promoters with stronger ones.

The results obtained showed that *zwf* gene overexpression did not improve plasmid productivity in GALG20 strains, behaving similarly to GALG20rpiA.

1. Introduction

There exist several possible vectors to introduce the genetic information into human cells. The most relevant are virus (adenovirus or retrovirus) and plasmid DNA (pDNA), which can be used in aqueous solution (naked) or included in lipids or other formulations. Although expression of the therapeutic gene following transfection with viral vectors is significantly higher than that using pDNA, the direct use of plasmids has important advantages over viruses. They are relatively simple to generate and safe to administer, consisting only of DNA or RNA, with decreased risk of integration into the human chromosomes [1]. Unfortunately, the immune response of pDNA when used for vaccination is usually lower than that of viral vectors [2]. Hence, large amounts of pDNA are necessary for a successful immunization.

The demand for plasmid DNA (pDNA) is anticipated to increase significantly as DNA vaccines and non-viral gene therapies enter phase 3 clinical trials and are approved for use. This increasing demand, along with renewed interest in pDNA as a therapeutic vector, has motivated research targeting the design of high-yield, cost-effective manufacturing processes [3]. Much effort has been directed towards downstream processing and fermentation strategies including factors that can enhance process performance such as the effects of growth medium, feeding strategy and employment of high cell-density

cultures. But it is also important to try to improve the earlier stages of production namely the cell host/plasmid selection where the choice of the *E. coli* host strain used to propagate the product containing-plasmid may have a significant influence on the quality and quantity of the purified DNA [4].

Recently, many cell line engineering efforts have been made in order to improve plasmid DNA production by knockout or overexpression of rationally-selected genes [5]. One main area of focus is the modification of central carbon metabolism genes to increase flux toward nucleotide and amino acid precursor synthesis and reduction of byproducts, such as acetate [5].

Pentose Phosphate (PP) pathway is considered the second main destination for glucose, being responsible for provision of reducing power (NADPH) for biosynthesis of nucleic acids, amino acids and vitamins, among others [6]. Therefore, several strategies have been used in order to modify this pathway and increase the carbon flux towards nucleotide biosynthesis. A recent study has demonstrated that the knockout of the *pgi* gene redirects glycolytic flux, increasing fluxes in the pentose phosphate pathway and enhancing nucleotide synthesis and NADPH production [7]. Glycolysis is down-regulated, but proceeds via the formation of fructose 6-phosphate and glyceraldehyde 3-phosphate [7]. GALG20 strain (MG1655*ΔendAΔrecAΔpgi*) produced three-fold more pVAX1GFP pDNA (11.6 mg/g DCW) than MG1655ΔendAΔrecA in 5+10 g/L glucose (3.6 mg/g DCW) [7]. In glycerol, GALG20 produced similar amounts of pDNA to MG1655ΔendAΔrecA. Surprisingly, this strain had the best performance in 20 g/L of glucose, producing 25-fold more pDNA (19.1 mg/g DCW) than the parental strain (0.8 mg/g DCW) [7]. Besides *pgi* gene, *endA* and *recA* genes are also absent in this strain increasing plasmid stability by eliminating non-specific digestion of the plasmid and undesirable recombination, respectively.

rpiA gene codes for ribose-5-phosphate isomerase A. This enzyme converts Ru5P into R5P, and its activity decreases the flux into X5P (Figure 1). Therefore, a large part of the flux into Ru5P is used for synthesis of nucleotides. Theoretically, overexpression of *rpiA* would enhance synthesis of the nucleotide precursor ribose-5-phosphate (R5P), which would lead to an increase in nucleotide formation, and consequently, plasmid DNA production. Wang *et al*. [8] performed a study in which overexpression of the *rpiA* gene in *E. coli* BL21 was tested. This event caused a 3-fold increase in plasmid copy number of a ColE1-derived plasmid during continuous culture, using defined medium and glucose as the carbon source [8]. In a more recent work, a cultivation was performed in order to obtain a deeper understanding of plasmid replication in GALG20 (MG1655*ΔpgiΔendAΔrecA)* strains overexpressing the *rpiA* gene [9]. Increased production of Ribose-5-phosphate Isomerase (RpiA) allowed a faster plasmid replication until a certain point, where it became similar to GAL20 [9].

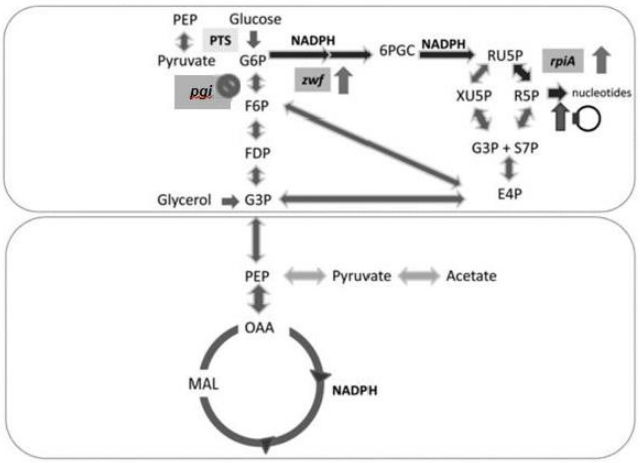

Figure 1 - Gene knockout and overexpression strategies to improve plasmid DNA production in *E. coli*. Overexpression of the *zwf* and *rpiA* genes is proposed as a means to increase fluxes in the pentose phosphate pathway and nucleotide synthesis. Knockout of *pgi* gene is suggested to reduce acetate formation, increase fluxes in the pentose phosphate pathway and TCA cycle. Adapted from Gonçalves *et al.* (2012) [5].

zwf gene codes for glucose-6-phosphate dehydrogenase (G6PDH) which is an enzyme responsible for the conversion of D-glucose 6-phosphate and $NADP^+$ into 6-phospho-D-glucono-1,5-lactone and NADPH (Figure 1). NADPH and nucleotides are required for biomass and pDNA production and they are intrinsically correlated in the PP pathway [5]. Overexpression of *zwf* has been investigated as a strategy to increase flux to the

pentose phosphate pathway. Williams *et al.* [10] verified that simultaneous overexpression of *zwf* and *rpiA* in DH5α appeared to increase plasmid amplification rate (mg pDNA/L/OD_{600}/hr) but not final specific yield (mg/L/OD_{600}), in fed-batch fermentations using complex medium with glycerol as the carbon source. In the same study, the overexpression of *zwf* alone as well as *zwf* and thioredoxin (*trxA)* did not show any effect on fermentation productivity [10].

These results indicate that engineering metabolic flux at the level of the G6P node can be used as a strategy to compensate for special metabolic demands of bacterial production strains. Hence, the goal of this study is to analyse the effect of overexpression of *zwf* and *rpiA* genes in plasmid DNA production by GALG20 (MG1655*ΔendAΔrecAΔpgi*) strains. In order to do that, the promoters of these genes will be replaced by new ones with higher strengths.

2. Materials and methods

2.1 Strains and plasmids

The host strain used as the basis of this work is MG1655*ΔendAΔpgi*. This strain is the result of an intermediate step in the GALG20 strain (MG1655*ΔendAΔrecAΔpgi*) construction, lacking the *recA* gene knockout. Strains used in this project and the corresponding genotype are represented in Table 1.

To construct the different strains three plasmids were used. The first one, pKD13 plasmid, contains a Kanamycin resistance gene [12]. The second one, pKD46 plasmid [13, 14], contains the Red recombinase gene from phage λ and an Ampicillin resistance gene. The third one, pCP20 plasmid [14, 15], was built by cloning of the FLP recombinase and carries Chloramphenicol and Ampicillin resistance genes. These constructions were performed at MIT and provided by Geisa Gonçalves, former PhD student at IST. To quantify plasmid DNA production by the different strains, a fourth plasmid, pVAX1GFP (3697 bp), was used. This plasmid contains a Kanamycin resistance gene for bacterial selection and the human cytomegalovirus (CMV) immediate-early promoter, for expression in human cells. It also has a pUC origin of replication lacking the *rop* gene, which exerts control mechanisms to keep the copy number low. With such modification, pUC plasmids can reach copy numbers of more than 500 copies per cell [2].

2.2 Promoter replacements

The new strains were constructed using a promoter replacement protocol adapted from the method described by Datsenko and Wanner [16]. The first step is the construction of the kan-cassette through a PCR reaction with pKD13 plasmid and specific primers containing the new promoters. The product obtained consists on a kan-cassette flanked by large stretches of homology to the area in the genome that is targeted, as well as the promoter sequence on the upstream side (Figure 2). It also contains a kanamycin resistance gene surrounded by FRT (FLP recognition target) sites to facilitate later excision of the Kanamycin resistance gene.

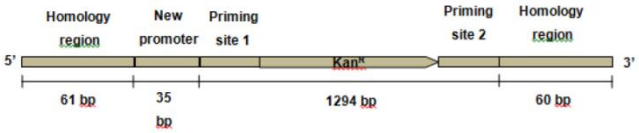

Figure 2 - Kanamycin cassette used in the promoter replacements (1450 bp).

The promoter sequences were obtained from the Anderson promoter collection [17], a well-characterized collection of *E. coli* σ^{70} constitutive promoters recovered from a library screen by Chris Anderson. The promoter strengths and sequences are represented in Table 2.

Table 1 - Strains used in this project and the respective genotype

Strains	Genotype	Reference
GALG20	MG1655*ΔendAΔrecAΔpgi*	[11]
MG1655*ΔendAΔpgi*	*F- λ –ilvG rfb-50 rph1ΔendAΔpgi*	[11]
GALG20rpiA5	MG1655*ΔendAΔrecAΔpgi rpiA*$^+$	[9]
GALG20rpiA10	MG1655*ΔendAΔrecAΔpgi rpiA*$^+$	[9]
GALG20zwf3	MG1655*ΔendAΔrecAΔpgi zwf*$^+$	This work
GALG20zwf5	MG1655*ΔendAΔrecAΔpgi zwf*$^+$	This work
GALG20zwf10	MG1655*ΔendAΔrecAΔpgi zwf*$^+$	This work

Table 2 - Sequences of the promoters used in the promoter replacement protocol [17].

Promoter sequence	Measured Strength
tttacggctagctcagtcctaggtacaatgctagc	0.33
ctgacagctagctcagtcctaggtataatgctagc	0.51
ttgacggctagctcagtcctaggtacagtgctagc	1

The insertion of the kan-cassette is accomplished by Red-mediated recombination in these flanking homologies achieved by the transformation of cells with a plasmid containing the λ-RED recombinase (pKD46 plasmid). After selection of the transformed colonies, the resistance gene can be eliminated by using a helper plasmid expressing the FLP recombinase (pCP20 plasmid), which acts on the directly repeated FRT (FLP recognition target) sites flanking the resistance gene. The Red and FLP helper plasmids can be simply cured by growth at 37°C because they are temperature-sensitive replicons [16]. The removal of the kan-cassette is then confirmed by PCR reaction and sequencing analyses.

2.3 P1 transduction

To perform the *recA* gene knockout P1 transduction protocol was used [18]. This gene was the last to be deleted, since all the previous modifications require an active RecA recombinase. This protocol is used to transfer genetic material such as genes or mutations from one strain of bacteria (the donor) to another (the recipient) [18]. In this case, the donor cells contain a kan-cassette with homology regions to the flanking regions of *recA* gene. Through controlled infection, the donor cells will introduce this kan-cassette into the recipient cells, being incorporated in the target genome and replacing the gene. Finally, the kan-cassette is removed and the cells are *recA*$^-$.

The first part of the protocol is the preparation of the donor phage stock. A 5 mL LB culture tube of the donor cells (JW2669-1 [Δ*recA*774::*kan*] [19]) with antibiotic was prepared, as well as a 5 mL LB culture tube with recipient cells without antibiotic, and incubated overnight at 37°C and 250 rpm. The next day, a 5 mL tube was prepared containing LB medium with 0.1% glucose, 5 mM CaCl2, and the antibiotic. This solution was then inoculated with 30 µL of the donor cells and incubated at 37°C until the OD_{600nm} was between 0.15-0.25. When the cells reached that OD_{600nm} value, P1 phage (ATCC #25404-B1) was added to the culture (5-10 µL) and incubated for 5-6 hours in order to allow the phage to infect and lyse the donor cells. After that, if cell debris were visible (if not, it must be incubated overnight proceeding with the protocol in the following day), 100 µL of chloroform were added to the culture and mixed by vortex for 1 minute. Then, 1 mL of the aqueous phase was transferred to an eppendorf and centrifuged at maximum speed for 2 minutes. The supernatant was then transferred to a new eppendorf. Next, 50 µL of chloroform were added and the culture was mixed vigorously with vortex. Then, the culture was centrifuged again for 2 minutes at maximum speed. Finally, the supernatant was stored as donor phage stock at 4°C, being stable for at least 3 weeks. The next step of the protocol is the preparation of the recipient cells. First, a 50 mL LB shake-flask was inoculated with 5 mL of stationary-phase recipient cell culture from the previous step and incubated at 37°C until the OD_{600nm} was approximately 0.8-1.2 (mid-exponential phase cells). After that, the cells were harvested by centrifugation at 4,000 g and 4°C for 15 minutes. The supernatant was discarded and the pellet was resuspended in 500 µL of 10 mM MgSO4 + 10 mM CaCl2. This solution was then divided into five 100 µL aliquots and 0, 30, 50, 70 and 120 µL of the donor phage stock were added to each aliquot. The aliquots were incubated at 30°C for 30 minutes, to allow the donor phage to infect the recipient cells. Next, 100 µL of 0.1 M sodium citrate (pH 7.0) were added to each aliquot to stop the infection process. Subsequently, the cells were centrifuged at 7000 rpm for 1 minute. This washing step was performed three times. After the final centrifugation, the cells were resuspended in 100 µL of 0.1 M sodium citrate (pH 7.0) and plated onto four LB plates with antibiotic (kanamycin, in this case) and incubated at 37°C overnight. The next day a colony PCR was performed in order to verify if the knockout was successful.

2.4 Medium and growth conditions

For cell growth, the medium used was LB (Luria-Bertani) from Sigma Aldrich (Molecular Biology tested). Some of the cell growths need the addition of antibiotics: Ampicillin (100 µg/mL), Kanamycin (30 µg/mL) and Chloramphenicol (50 µg/mL). These were obtained from Calbiotech. The agar plates were made using LB-agar from NZYtech.

In the P1 transduction protocol three solutions were used: 0.1M $CaCl_2$ (Merck ®), 10 mM $MgSO_4$ (Merck ®), 0.1 M sodium citrate (Merck ®) pH 7.

To perform the shake-flask cultivations a complex medium was used. This medium contains a basal cultivation medium [Glucose (Panreac ®), 20g/L; Bacto peptone (BD ®), 10 g/L; yeast extract (BD ®), 10 g/L; $(NH_4)_2SO_4$ (Panreac

®), 3 g/L; K_2HPO_4 (Panreac ®), 3.5 g/L; KH_2PO_4 (Panreac ®), 3.5 g/L] with a final pH of 7.1. After autoclaving the basal cultivation medium the following components are added: Seed supplement solution, 8.3 mL/L (thiamine (Sigma Aldrich ®), 199 mg/L; $MgSO_4$ (Sigma Aldrich ®), 1.99 g/L); Trace elements solution, 1 mL/L ($FeCl_3$•$6H_2O$, 27 g/L; $ZnCl_2$, 2 g/L; $COCl_2$•$6H_2O$, 2 g/L; Na_2MoO_4•$2H_2O$, 2 g/L; $CaCl_2$•$2H_2O$, 1 g/L; $CuCl_2$•$2H_2O$, 1.3 g/L; H_3Bo_3, 0.3 g/L; 1.2N HCl, 100 mL/L) [11]. This complex medium is then supplemented with 30 µg/mL of kanamycin.

The first step in the shake-flask cultivations was the preparation of an overnight culture. The next day, the OD600nm was measured and an appropriate volume of cells (in order to start cultivation with an OD600nm = 0.1) was transferred into a shake-flask containing 45 mL of basal cultivation medium, 415 µL of seed supplement solution, 50 µL of trace elements solution (1 mL/L), 5 mL of Glucose (20 g/L) and 50 µL of kanamycin (30 µg/mL). The cells were incubated at 37°C and 250 rpm. This culture was grown for 24 hours and the OD600nm was measured every hour during the first 12 hours and the final sample taken at 24 hours, in order to construct a growth curve. Sampling for DNA quantification was performed after 12 hours of growth. This experiment was performed four times.

2.5 Plasmid DNA quantification

Plasmid DNA quantification was performed using High Pure Plasmid Isolation Kit from (Roche®) and NanoVue Plus Spectrophotometer (GE Healthcare ®) equipment. During the cell growth, 1 mL of the culture was collected after 12 hours of growth and used to purify PVAX1GFP plasmid. The sample was then centrifuged at 12,000 rpm for 10 minutes and the supernatant discarded. The pellet was then stored at -20°C. To perform the isolation of the plasmid with Roche kit, an appropriate volume of water was added to the cell pellet in order to start with an OD600nm=4 with 1mL of sample. After the isolation protocol, the plasmid concentration was quantified using NanoVue Plus Spectrophotometer (GE Healthcare ®).

3. Results and discussion

After the construction of the different strains it is necessary to evaluate cell growth behavior as well as pDNA production. In order to do that, all of the strains (Table 1) were electroporated with pVAX1GFP plasmid (Kan^R) and cultivated in shake-flasks.

The growth curves of the different strains are represented in Figure 3. All the strains showed a similar growth behavior, reaching stationary phase after approximately 6 hours. GALG20rpiA5 presented the highest growth rate (0.71 ± 0.02 h^{-1}) while GALG20zwf10 presented the lowest (0.54 ± 0.15 h^{-1}) (Table 3). Taking into consideration the standard deviation, the values obtained are reasonable and close to the ones described in previous studies [9,11].

The final OD_{600nm} value achieved was lower than expected, in all the strains. This might be due to problems in the cultivation conditions, such as temperature or agitation shifts, different lots of the components or different shake-flasks which can cause

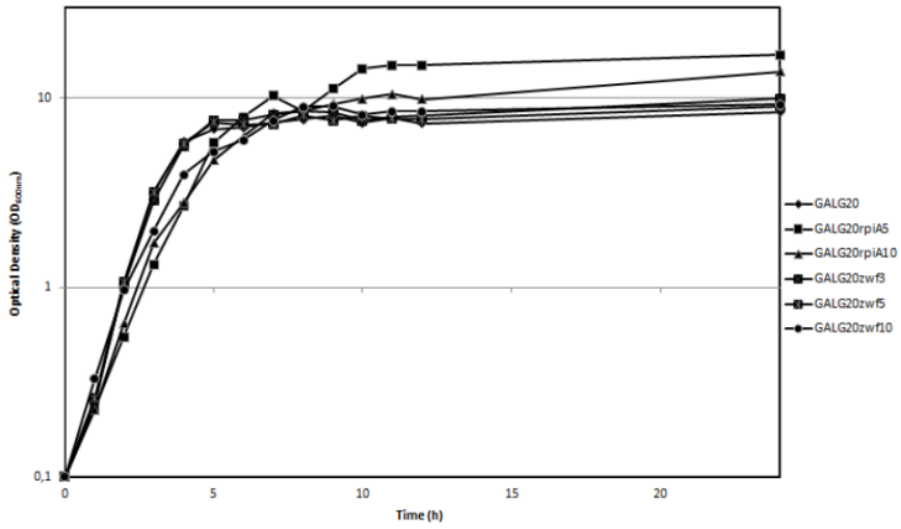

Figure 3 - Growth curves of the different *E.coli* strains used in the shake-flask cultivations.

Table 3 - Growth rate of the strains used in the shake-flask cultivations in comparison with values obtained in previous studies.

Strain	Growth rate (h^{-1})	Growth rate from previous studies (h^{-1})
GALG20	0.68±0.05	0.77±0.05 [11]; 0.85 [9]
GALG20rpiA5	0.71±0.02	0.90 [9]
GALG20rpiA10	0.64±0.09	0.89 [9]
GALG20zwf3	0.68±0.05	-
GALG20zwf5	0.67±0.04	-
GALG20zwf10	0.54±0.15	-

alterations in the aeration, among others. These parameters are hard to control and can be responsible for differences between studies. In addition, although the strains used have the same genotype as the ones from previous studies, in this project they were constructed again starting from MG1655 strain. This might also add some variability to the study.

After the shake-flask cultivations pDNA production by the different strains was quantified using NanoVue Plus Spectrophotometer (GE Healthcare ®). pVAX1GFP plasmid was purified from the different strains using Roche kit and then its concentration was measured.

In order to obtain relevant information about pDNA productivity two parameters were determined. Specific plasmid yield (mg/g DCW) indicates plasmid productivity per gram of cells while volumetric plasmid yield (mg/L) is relative to the total productivity of the cell culture. The first parameter is very important since it demonstrates which cells produce higher amounts of DNA, while the second is dependent on the growth behavior of the cell (more biomass formation, more plasmid DNA produced). In Figure 4 and Figure 5 is represented the volumetric plasmid yield (mg/L) and the specific plasmid yield (mg/g DCW), respectively, of each strain after 12 hours of growth.

As observed in Figures 4 and 5, all the strains produced similar amounts of DNA. In order to have a better understanding of the values obtained in this study, in Table 4 these values are compared with the ones obtained in previous studies.

Relative to volumetric plasmid yield, GALG20rpiA5 appears to have higher productivity in comparison with the other strains. This value was expected since this strain also had a higher growth rate (Table 3). Therefore, with more cell biomass, more total pDNA is produced. GALG20rpiA10 had a lower volumetric yield, since it had a lower growth rate. Concerning specific plasmid yield, both GALG20rpiA strains show a similar productivity. Wang *et al.* [8] performed a study in which overexpression of the *rpiA* gene in *E. coli* BL21 caused a 3-fold increase in plasmid copy number of a ColE1-derived plasmid during continuous culture, using defined medium and glucose as the carbon source [8]. However, since the strain used was different, the values obtained are not directly comparable. In a more recent study, it was demonstrated that increased production of Ribose-5-phosphate Isomerase permitted a faster plasmid replication until a certain point, where it becomes similar to GAL20 [9]. The calculated values were higher than the ones in this project (Table 4), however these were obtained after 24 hours of growth, while the ones relative to this study were obtained only after 12

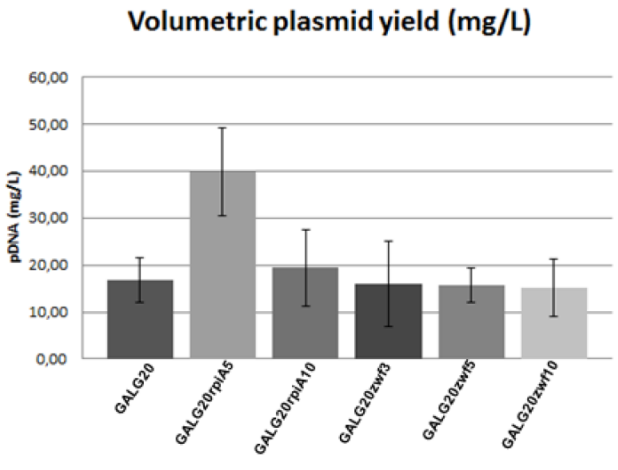

Figure 4 - Volumetric plasmid yield (mg/L) calculated for the different strains used in shake-flask cultivations.

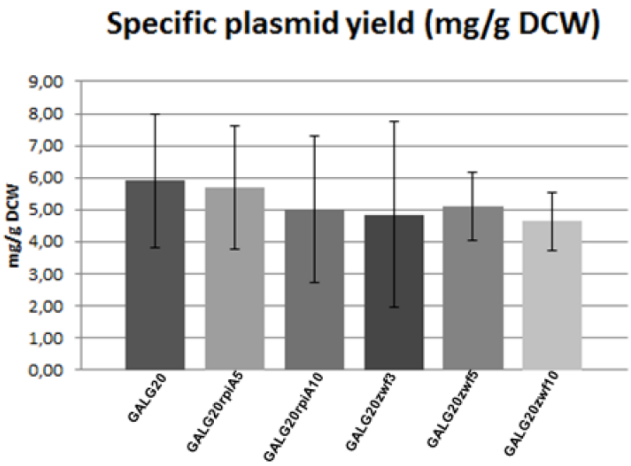

Figure 5 - Specific plasmid yield (mg/g DCW) calculated for the different strains used in shake-flask cultivations

Table 4 - Plasmid DNA productivity of the different strains used in shake-flask cultivations in comparison with values obtained in previous studies.

Strain	pDNA (mg/L)	pDNA (mg/L), previous studies	pDNA (mg/g DCW)	pDNA (mg/g DCW), previous studies
GALG20	16.8 ± 4.8	90.1 ± 1.9 [11] 125 [9]*	5.9 ± 2.1	10.3 ± 0.5 [11]; 13.3 [9]*
GALG20rpiA5	39.8 ± 9.3	30 [9]*	5.7 ± 1.9	3.8 [9]*
GALG20rpiA10	19.4 ± 8.2	90 [9]*	5.0 ± 2.3	9.6 [9]*
GALG20zwf3	16.0 ± 9.1	-	4.9 ± 2.9	-
GALG20zwf5	15.8 ± 3.7	-	5.1 ± 1.1	-
GALG20zwf10	15.2 ± 6.1	-	4.6 ± 0.9	-

*values obtained after 24 hours of growth

hours of growth. In the previous study it was observed a 3.5 fold increase in the specific pDNA yield of GALG20 between 12 and 24 hours, while GALG20rpiA5 and GALG20rpiA10 displayed a 2.6 and 1,022 fold increase [9]. Regarding the productivity after 12 hours of growth, all the strains presented a value of approximately 4 mg/g DCW and 30 mg/L. Taking this into consideration, it is possible to conclude that the values obtained in the two studies are similar and the small differences can be explained by variability in the experimental conditions, as mentioned before.

GALG20zwf3, GALG20zwf5 and GALG20zwf10 strains had very similar volumetric and specific productivities. The values presented in the literature are not comparable since the strain used was DH5α, but it is documented that the overexpression of *zwf* gene alone had no effect on fermentation productivity [10]. In this case it is possible to verify that there is no significant increase in specific plasmid productive. Relative to volumetric yield, these strains were the ones with the lower values in comparison with the other strains.

GALG20 strain was expected to have high specific and volumetric plasmid DNA productivities. Since the main via of utilization of carbon in this strain is the PPP, glycolysis is down-regulated, and low amount of acetate is formed, which can be an advantage for pDNA production. Using a fed-batch strategy, this strain achieved productivities of ~160 mg/L [11]. However, the plasmid productivities of this strain in this study were lower than expected. As mentioned before, control of the cultivation parameters is difficult in the case of shake-flask cultivations and so differences between experiments are expected.

The results obtained in this study were, in general, between acceptable values. However, there were some changes in the protocol used that might explain the observed differences. Although the growth conditions were the same as in previous studies [9, 11] small unintended differences can have a big influence on the final results. Perhaps some of the components used in the growth media are from a different brand or a different lot which can influence its activity and consequently cell growth. Another important aspect is the plasmid DNA purification method. In this case, plasmid DNA purification and quantification was performed using Miniprep Kit from Roche® and NanoVue Plus Spectrophotometer (GE Healthcare ®), respectively. In the previous studies used as comparison [9, 11] plasmid DNA was quantified from crude alkaline lysates prepared from cell pellets (OD_{600nm} = 10) using the hydrophobic interaction HPLC method described before by Diogo *et al.* [20]. This method is much more precise and reduces plasmid DNA losses during the procedure, therefore resulting in higher amounts of DNA. Another possible reason related to plasmid purification method might be the capacity of the Miniprep column. It was hypothesized that the column might become clogged or saturated and was losing pDNA. To test that hypothesis, instead of purifying a sample with OD_{600nm}=4, samples with OD_{600nm}=1 and OD_{600nm}=2 were used. But the difference in the results was not significate which indicates that this might not be the problem. Also, the use of NanoVue Plus Spectrophotometer is associated with the occurrence of errors in the measurements. To test that, some plasmid DNA samples were quantified using a standard spectrophotometer and OD_{260nm} was

measured. The concentration of pDNA was calculated and the values obtained were superior relative to Nanodrop measurements with a ratio between 1 and 2. This demonstrates that the use of NanoVue Plus Spectrophotometer was not the best approach to quantify plasmid DNA, being important to compare results from the two methods. Finally, it is also known that shake-flask cultivations are not a very precise method to evaluate plasmid DNA productivity due to the lack of control over important parameters such as pH and dissolved oxygen.

4. Conclusions and future work

Plasmid DNA market is expected to increase in the next years, being necessary to create a system with high productivity and low costs. Several studies have been performed with the goal of improving both upstream and downstream processes of plasmid DNA production process. However there is still no consensus on which strategy is the best in all parameters.

The goal of this project was to verify the influence of *zwf* gene overexpression on plasmid DNA production on GALG20 strains. In order to that, three promoter replacements were performed using three different strengths (0.33, 0.51 and 1). There are a few previous studies that indicate that overexpression of this gene would have a positive impact on pDNA production by increasing the reducing power available for pDNA [21]. However, in this study a high productivity by these strains was not observed. On the contrary, they showed a similar or worse behavior when compared to the other strains evaluated. Williams *et al.* [10] verified that simultaneous overexpression of *zwf* and *rpiA* in DH5α appeared to increase plasmid amplification rate (mg pDNA/L/OD600/hr). On the other hand, overexpression of *zwf* alone as well as *zwf* and thioredoxin (*trxA*) did not show any effect on fermentation productivity [10]. Therefore, as future work, it is necessary to investigate the influence of simultaneous overexpression of *rpiA* and *zwf* genes in GALG20 strains. It was previously observed that GALG20rpiA strains have an increased plasmid amplification rate [8,9] so it would be useful to test whether these strains overexpressing also *zwf* gene would produce higher amounts of plasmid DNA relative to GALG20.

The other strains evaluated in this study did not present any significant result. Highest volumetric plasmid yield was obtained by GALG20rpiA5 strain (39.8 mg/L) while the lowest was obtained by GALG20zwf10 (15.2 mg/L). Relative to specific plasmid yield, GALG20 presented a productivity of 5.9 mg/g DCW followed by GALG20rpiA5 with 5.7 mg/g DCW. Again, GALG20zwf10 gave the lowest values with only 4.6 mg/g DCW. In comparison with previous studies all the strains had lower productivities than expected, in particular GALG20. The differences observed might be explained by the difficulty in controlling certain growth parameters. Although the growth medium was the same, certain parameters are difficult to control and small changes to the protocol can have big influences on the final plasmid DNA productivity.

There are several feeding strategies that can be used in the process of pDNA production. However it is known that productivity data obtained from shake flask experiments often fail to predict the outcome of pDNA production in bench-scale bioreactors [10]. The lack of control of important parameters such as pH and dissolved oxygen can lead to high concentrations of accumulated acetate which can be toxic and inhibit biomass formation [22]. In addition, it was observed that plasmid degradation can occur at the end of the fermentation in shake flasks as well as the formation of nicked plasmids and multimers [23, 24]. In this study lower productivities were observed relative to previous studies, which might be due to the limitations of shake-flask cultivations and difficulties in controlling important parameters. Therefore, it would be important to test the strains constructed in this study using batch and fed-batch strategies. Growth conditions should also be optimized in order to try to obtain higher plasmid DNA productivities. Finally, precise plasmid DNA purification

and quantification methods must be used in order to obtain reliable values.

In conclusion, there are still a lot of studies to perform in order to identify a strategy that increases plasmid productivity and quality, reducing the production costs. In addition to modifying genes related to plasmid productivity it is also necessary to increase stability and safety. It is known that plasmid productivity reaches a plateau in which the cells cannot produce higher quantities. Williams *et al.* [10] hypothesized that it was probably determined by the efficiency of the replication origin and the percent of initiated replication cycles that are completed. The amplification slope may be set by a limiting factor, perhaps a protein or nucleotide necessary for plasmid DNA replication [10]. Given that, the genes responsible for these mechanisms have been analysed, such as genes involved in DNA replication. For instance, *topA* (Topoisomerase I), *polA* (DNA Polymerase I), *ligA* (DNA ligase) and *gyrAB* (DNA gyrase) are some of the genes that have been tested by overexpression or knockout. Therefore, a combination of several strategies must be performed in which different upstream and downstream processes are evaluated, namely fermentation strategy, ideal host strain and effective purification methods, with the goal of creating the best system for plasmid DNA production.

5. References

[1] NP Restifo, H Ying, L Hwang, and W. L. (2000). The promise of nucleic acid vaccines. *Gene Therapy, 7*, 89–92.

[2] Lara, A. R., & Ramírez, O. T. (2012). Plasmid DNA Production for Therapeutic Applications. (A. Lorence, Ed.)*Methods in Molecular Biology, 824,* 271–303. doi:10.1007/978-1-61779-433-9

[3] Bower, D. M., & Prather, K. L. J. (2009). Engineering of bacterial strains and vectors for the production of plasmid DNA. *Applied Microbiology and Biotechnology, 82*(5), 805–13. doi:10.1007/s00253-009-1889-8

[4] Yau, S. Y., Keshavarz-Moore, E., & Ward, J. (2008). Host strain influences on supercoiled plasmid DNA production in *Escherichia coli*: Implications for efficient design of large-scale processes. *Biotechnology and Bioengineering, 101*(3), 529–44. doi:10.1002/bit.21915

[5] Gonçalves, G. a L., Bower, D. M., Prazeres, D. M. F., Monteiro, G. a, & Prather, K. L. J. (2012). Rational engineering of *Escherichia coli* strains for plasmid biopharmaceutical manufacturing. *Biotechnology Journal, 7*(2), 251–61. doi:10.1002/biot.201100062

[6] Ow, D. S.-W., Nissom, P. M., Philp, R., Oh, S. K.-W., & Yap, M. G.-S. (2006). Global transcriptional analysis of metabolic burden due to plasmid maintenance in *Escherichia coli* DH5α during batch fermentation. *Enzyme and Microbial Technology, 39*(3), 391–398. doi:10.1016/j.enzmictec.2005.11.048

[7] Gonçalves, G. a L., Prazeres, D. M. F., Monteiro, G. a, & Prather, K. L. J. (2013). De novo creation of MG1655-derived *E. coli* strains specifically designed for plasmid DNA production. *Applied Microbiology and Biotechnology, 97*(2), 611–20. doi:10.1007/s00253-012-4308-5

[8] Wang, Z., Xiang, L., Shao, J., Wegrzyn, A., & Wegrzyn, G. (2006). Effects of the presence of ColE1 plasmid DNA in *Escherichia coli* on the host cell metabolism. *Microbial Cell Factories, 5*, 34. doi:10.1186/1475-2859-5-34

[9] Carreira, L. (2013). Evaluation of *rpiA* promoter strength in plasmid DNA production. Master thesis, Instituto Superior Técnico da Universidade de Lisboa, Lisboa

[10] Williams, J. a, Luke, J., Langtry, S., Anderson, S., Hodgson, C. P., & Carnes, A. E. (2009). Generic plasmid DNA production platform incorporating low metabolic burden seed-stock and fed-batch fermentation processes. *Biotechnology and Bioengineering, 103*(6), 1129–43. doi:10.1002/bit.22347

[11] Gonçalves, G. (2013). Rational engineering of Escherichia coli strains for plasmid biopharmaceutical manufacturing. PhD thesis, Instituto Superior Técnico da Universidade de Lisboa, Lisboa

[12] Template plasmid pKD13, complete sequence. [Online]. [cited 2014 October 25. Available from: http://www.ncbi.nlm.nih.gov/nuccore/15554335

[13] Red recombinase plasmid pKD46, complete sequence. [Online]. [cited 2014 October 25. Available from: http://www.ncbi.nlm.nih.gov/nuccore/AY048746.1

[14] Doublet, B., Douard, G., Targant, H., Meunier, D., Madec, J.-Y., & Cloeckaert, A. (2008). Antibiotic marker modifications of lambda Red and FLP helper plasmids, pKD46 and pCP20, for inactivation of chromosomal genes using PCR products in multidrug-resistant strains. *Journal of Microbiological Methods, 75*(2), 359–61. doi:10.1016/j.mimet.2008.06.010

[15] Cherepanov, P. P., & Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene, 158*(1), 9–14.

[16] Datsenko, K. a, & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America, 97*(12), 6640–5. doi:10.1073/pnas.120163297

[17] Registry of Standard Biological Parts. [Online]. [cited 2014 October 25. Available from: http://parts.igem.org/Promoters/Catalog/Anderson

[18] Thomason, L. C., Costantino, N., & Court, D. L. (2007). *E . coli* Genome Manipulation by P1 Transduction. *Current Protocols in Molecular Biology, 1.17*(July), 1–8. doi:10.1002/0471142727.mb0117s79

[19] Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., … Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology, 2*. doi:10.1038/msb4100050

[20] Diogo, M. M., Queiroz, J. a, & Prazeres, D. M. F. (2003). Assessment of purity and quantification of plasmid DNA in process solutions using high-performance hydrophobic interaction chromatography. *Journal of Chromatography. A, 998*(1-2), 109–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12862377

[21] Cunningham, D. S., Koepsel, R. R., Ataai, M. M., Domach, M. M., (2009). Factors affecting plasmid production in *Escherichia coli* from a resource allocation standpoint. *Microbial Cell Factories, 8,* 1–27.

[22] Gonçalves, G. a L., Prather, K. L. J., Monteiro, G. a, & Prazeres, D. M. F. (2014). Engineering of *Escherichia coli* strains for plasmid biopharmaceutical production: scale-up challenges. *Vaccine, 32*(24), 2847–50. doi:10.1016/j.vaccine.2014.02.023

[23] Tejeda-Mansir, A., & Montesinos, R. M. (2008). Upstream processing of plasmid DNA for vaccine and gene therapy applications. *Recent Patents on Biotechnology*, 2(3), 156–72.

[24] Kennedy, R. D. O., Ward, J. M., Keshavarz-moore, E., Street, G., & Wce, L. (2003). Effects of fermentation strategy on the characteristics of plasmid DNA production. *Biotechnology and Applied Biochemistry*, *37*, 83–90.