Iranian J Biotech. 2017 April;15(2):e1467 **DOI:10.15171/ijb.1467**

Research Article

The Increase in Protein and Plasmid Yields of *E. coli* with Optimized Concentration of Ampicillin as Selection Marker

Sadegh Feizollahzadeh [1], Shirin Kouhpayeh [1], Ilnaz Rahimmansh [2], Hossein Khanahmad [2], Faezeh Sabzehei [2], Mazdak Ganjalikhani-hakemi [1], Alireza Andalib [1], Zahra Hejazi [2], Abbas Rezaei [1*]

[1] Department of Immunology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, 81746-73461 Iran
[2] Department of Molecular Biology and Genetics, School of Medicine, Isfahan University of Medical Sciences, Isfahan, 81746-73461 Iran

[*]*Corresponding author*: Abbas Rezaei, Department of Immunology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran. Tel: +98 31 37922460, Fax: +98 31 36688597, E-mail: rezaei@mui.ac.ir & rezaei1335@yahoo.com

Received: 1 March 2016; **Revised**: 10 Oct. 2016 ; **Accepted**:19 June 2017; **Published online:** 19 August 2017

Background: *Escherichia coli* is still the common host for ing and heterologous protein expression. Various strategies have been employed to increase protein expression in *E. coli*, but, it seems that external factors such as selection marker concentration can drastically affect the yield of protein and plasmid.
Objectives: Alterations of protein expression and plasmid yields of *E. coli* in different concentrations of ampicillin, as selection marker, will be determined. In order to improve heterologous expression, the system will be redesigned and optimized.
Materials and Methods: The expression cassette of codon optimized EGFP for *E. coli* was synthesized in pUC57. The pUC57-GFP was transformed into *E. coli Top10F'*. The expression of GFP was verified by SDS-PAGE and flow cytometry after induction by IPTG (0.5 mM) and incubation with 0, 100, 200 and 300 $\mu g.mL^{-1}$ ampicillin. Plasmid copy numbers of samples were determined by Real-Time PCR on AMP gene using regression line of diluted standard curve.
Results: GFP expressing clones formed fair green colonies on LB agar supplemented with 0.5 mM IPTG and showed fluorescence in FL1 filter of flow cytometry and an extra protein band on SDS-PAGE gel. The fluorescent intensity of GFP in 0, 100, 200 and 300 $\mu g.mL^{-1}$ ampicillin in medium were 549.83, 549.78, 1443.52, 684.87, and plasmid copy numbers were 6.07×10^{9}, 3.21×10^{9}, 2.32×10^{10} , 8.11×10^{8}, respectively. The plasmid yields were 55 $ng.\mu L^{-1}$, 69 $ng.\mu L^{-1}$, 164 $ng.\mu L^{-1}$ and 41 $ng.\mu L^{-1}$, respectively.
Conclusion: Protein and plasmid yields of *E. coli* are variable in different concentrations of ampicillin and need to be optimized in newly designed expression systems. Protein and plasmid yield in the optimized concentration (200 $\mu g.mL^{-1}$) was significantly ($P < 0.01$) higher than other doses.

Keywords: Ampicillin, *Escherichia coli*, Plasmid, Protein

1. Background

Escherichia coli is still one of the most widely used hosts for cloning and heterologous protein expression in research and industry (1-2). Various strategies have been employed to improve and increase protein expression in *E. coli* to cover the vast area of cloning and protein expression (3-5). In this regard, many improved strains and efficient expression vectors have been developed. Using strong promoters, suitable translation initiator and terminators, and origin of replications led to the generation of excellent protein expressing plasmids. Employing this improved strains and vectors resulting in the protein yields of up to 1 $g.L^{-1}$ with *E. coli* (2). However, external factors such as temperature, biophysicochemical character of expressing proteins, and other growth conditions are important and can drastically affect the result of the genetically well designed system. One of the key elements of culture condition is the antibiotic selection marker that is being used for selection pressure exertion to ensure that only bacterial cells with resistant plasmid(s) are being allowed to grow. As long been known, in cultures without selection pressure, resistant bacteria tend to decrease and lose their plasmid (6). Some

defined doses of selection marker are used in culture to restore a selection pressure (7). However, it seems that plasmid maintenance, its copy number and distribution to daughter cells altered in different concentrations of antibiotic. Low doses of antibiotic couldn't restore enough selection pressure to keep plasmids and then led to the high rate of plasmid free cell generation. Higher doses of antibiotics for high selection pressure may affect the growth rate of bacteria and increase the cost of production (8). Both conditions are undesirable for research and industrial scale protein and plasmid preparation. Thus, adjustment of cultures with effective concentrations of antibiotic to restore effective selection pressure can lead to high final protein and plasmid extraction yields that lowers the final cost.

EGFP (Enhanced green florescent protein) is a commonly used reporter gene and it is widely used in flow cytometric applications, since it is efficiently being excited with a standard 488-nm argon laser. Applications of green fluorescent protein (GFP) as a reporter gene for the analysis of gene expression and protein localization have become prevalent in many aspects of science. The attributes of GFP could also be applied to the area of heterologous protein production.

Among several antibiotic selection markers, ampicillin (Amp), a beta-lactam antibiotic, has been commonly used as a selection marker in gene cloning and protein expression in *E. coli*. The Amp is an irreversible inhibitor of transpeptidase enzyme and can kill bacteria by inhibition of cell wall synthesis (8). Plasmid taking bacteria produces beta-lactamase enzyme encoded by plasmid and hydrolyzes the β-lactam ring and inactivate Amp (8). Then, colonies that have successfully taken up the plasmid will most likely develop a resistance against Amp and can grow in defining doses of Amp while others would eventually die. It mechanistically does not interfere with protein synthesis and is the best choice of selection marker for protein expression systems.

In the present study, a protein expression system is designed based on the Amp selection marker to evaluate its effect on protein and plasmid yield. The expression cassette of EGFP on the pUC57 plasmid backbone was used to evaluate the effect of Amp on protein expression and plasmid yield in *E. coli*. In this study, GFP was used as a tool to monitor protein production and reporter gene for scale-up and optimization of recombinant protein production.

2. Objectives

The effect of different concentrations of Ampicillin on the level of protein expression and plasmid yields in *E. coli Top10F* as an expression host was determined. The appropriate dose of Amp as selection marker was optimized and high protein and plasmid yields were obtained in that dose. This could be used for newly designed protein expression systems.

3. Materials and Methods

3.1. Bacterial Strains, Plasmids and Growth

The codon optimized sequence of Enhanced Green Fluorescent Protein (EGFP) (9) under control of an inducible Lacuv5 promoter was synthesized (Genecust Company) and cloned in pUC57 (Fig. 1). The construct (pUC-GFP) was transformed into chemically competent (10) *E. coli* strain Top10F' cells and spread on LB agar medium supplemented with 100 $\mu g.mL^{-1}$ Amp and 0.5 mM IPTG for clonal selection (Sigma-Aldrich, USA). GFP expression in green colonies was determined by flow cytometry and SDS-PAGE.

One of the selected clones was suspended in 5 mL of LB broth supplemented with 100 $\mu g.mL^{-1}$ Amp and incubated at 37 °C with shaking at 250 rpm for 4 h to obtain a homogeneous solution of bacteria. Homogeneous culture (100 µL) was inoculated into 10 ml LB broth supplemented with different concentrations (0, 100, 200 and 300 $\mu g.mL^{-1}$) of Amp. Cultures were induced with 0.5 mM IPTG at OD_{600} 0.8 and incubated for 16 h at 37 °C with shaking at 250 rpm. *E. coli* strain Top10F' without pUC-GFP was used as negative control.

3.2. SDS-PAGE Analysis

The standard sambrook method was used for SDS-PAGE analysis. Equal volume of each culture (based on OD_{600}) centrifuged at 7500 ×*g* for 3 min and pellets were resuspended in 70 µL 2× sample buffer and boiled for 10 min. Samples were loaded on 12% polyacrylamide gel and stained with Coomassie Brilliant Blue-R-250 or G-250 after electrophoresis.

3.3. Flow Cytometry

EGFP has the same excitation and emission range as FITC at 488 nm and 509 nm. Accordingly, normal FITC filter sets (488 nm, FL1) used for visualizing GFP. Flow cytometric analysis was performed to accumulate up to 100,000 events per tube in a set of triplicate for each sample by FACS Calibur Flow cytometer (BD Bioscience, USA). LB broth bacterial suspensions (1 mL each) were centrifuged after overnight incubation for 10 min at 7500 ×*g* and the pellets were resuspended in 1 ml of sterile PBS for flow cytometric analysis. The light scattered population of bacteria was gated and analyzed in FL1 filter for GFP expression.

Figure 1. Expression cassette of GFP. This construct contains Lacuv5 promoter, CDS of EGFP and rrnB-T1 terminator inserted in PUC57 vector..

3.4. Plasmid Extraction and Its Copy Number

Plasmids were extracted with commercial GeneJET™ Plasmid Miniprep kit after 18 h incubation at 37 ºC with shaking for each concentration of Amp. The extracted plasmids of each group were evaluated on agarose gel and plasmid concentration measured by Nanodrop. The copy number of plasmid in 1 µL was determined according to the number of nucleotides of pUC-GFP (11) with online copy number calculator (SciencePrimer.com). Standard curve was generated with different concentrations of the pUC57 plasmid (from 10^7 to 10^{14} plasmid molecules.μL^{-1}) prepared by serial dilutions. The prepared plasmids from cells cultured in different Amp concentration and standard samples were used for quantitative real-time PCR with Amp primers (F amp: 5′-CAGTGCTGCAATGATACCG-3′- R amp: 5′-AATAATAGACTGGATGGAGGC-3′) using ABI Step One Plus in a set of triplicate. The reaction mixture containing 10 µL of SYBR Green PCR master mix (Thermo Scientific, Waltham, MA, USA), 0.5 nM of each forward and reverse primer,1 µL of each plasmid and standard samples (template DNA) was adjusted to 20 µL final volume with deionized water. The samples were heated at 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s, and 60 °C for 1 min. After amplifying the standards, a linear relation was observed between the copy number and cycles of threshold (Ct) of real-time PCR. The regression line of diluted standard curve was used to determine the copy number of samples (0, 100, 200 and 300 $\mu g.mL^{-1}$).

3.5. *Statistical Analysis*

The mean GFP intensity and plasmid copy number in different concentrations of AMP were compared using independent t Test. Statistically significant level was set at 0.05.

4. Results

4.1. GFP Expression

Transformed clones formed green appearance colonies on LB agar supplemented with 0.5 mM IPTG. The fair green appearances of colonies were attributed to GFP expression of colonies.

4.2. SDS-PAGE Analysis

Green clonies were evaluated by SDS-PAGE for extra or strong band close to GFP protein molecular weight (27 kDa). The protein banding pattern of transformed clones on 12% polyacrylamide gel showed an extra protein band near 27 kDa in comparisons to un-transformed clones of *E. coli TOP10F'* (Fig. 2). The intensity of GFP bond was strong after induction by IPTG (Fig. 3). The intensity of GFP bond was variable

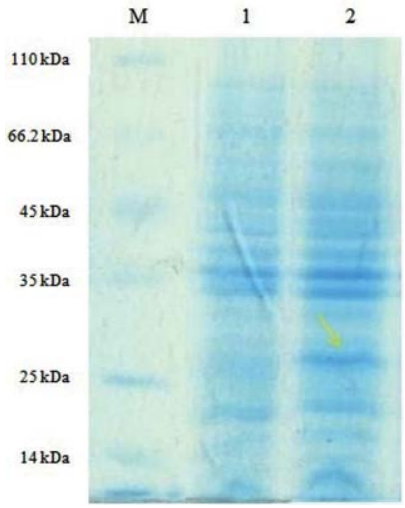

Figure 2. SDS-PAGE analysis of transformed and un-transformed clones. (M: protein marker, lane 1 un-transformed and lane 2 transformed clones).

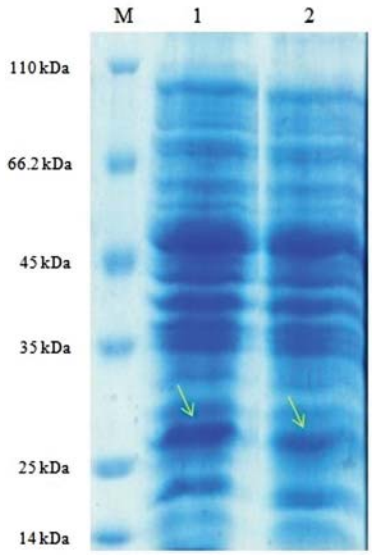

Figure 3. The protein banding pattern on SDS-PAGE after induction with IPTG. (M: protein marker, lane 1 induced and lane 2 un-induced clones).

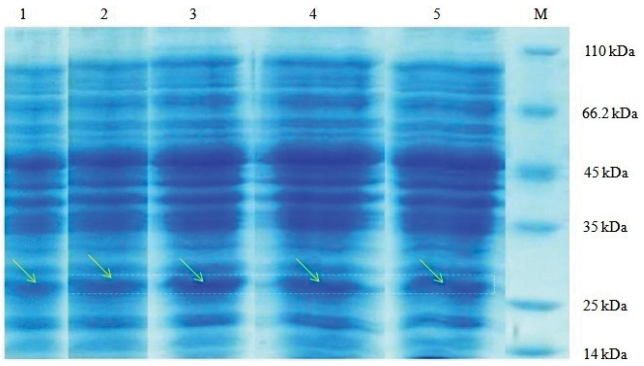

Figure 4. The protein band pattern on SDS-PAGE after induction with IPTG and incubation with different concentrations of Amp. Lane 1 un-induced, lanes 2, 3, 4 and 5 in the presence of 0, 100, 200 and 300 µg.mL^{-1} Amp, respectively, M= protein marker.

in different concentrations of Amp and was stronger in 200 µg.mL^{-1} Amp than 0, 100 and 300 µg.mL^{-1} (Fig. 4).

4.3. Flow Cytometry

The untransformed cell population (*E. coli Top10F'*) was located on the edge of the first log decade in FL1 filter without any emission and GFP expression (Fig. 5A$_{1-3}$). Transformed cells with GFP expression was shifted toward the second log decade in FL1 based on the intensity of GFP expression level (Fig. 5B$_{1-3}$).

The ratio of GFP expressing and non-expressing cells and its intensity as protein expression level in positive population were determined (Figs. 6A-D) (12). The fluorescent intensity of GFP in 0, 100, 200 and 300 µg.mL^{-1} Amp condition were 549.83, 549.78, 1443.52, 684.87, respectively (Fig. 7E). The fluorescent intensity of GFP was significantly higher ($p < 0.01$) in 200 µg.mL^{-1} than 0, 100 and 300 µg.mL^{-1} concentrations. The variation in GFP intensity (CV%) in the GFP positive population considered as the degree of heterogeneity in protein expressing cells. The variation of GFP expression was 53, 56, 36 and 61% in 0, 100, 200 and 300 µg.mL^{-1} Amp condition, respectively. Homogeneously (CV = 36%) high GFP expression in 2× Amp showed that under this condition, cells uniformly kept a high copy number of plasmid than without, 100 and 300 µg.mL^{-1}Amp concentrations. The GFP negative cells considered the plasmid losing cells in the provided culture condition. The GFP negative cells were 2.28, 2.74, 0.29, and 5.25% in 0, 100, 200 and 300 µg.mL^{-1} Amp condition. The low percent of GFP negative cells in 200 µg.mL^{-1} Amp condition showed that this condition is more compatible with plasmid maintenance than 0, 100 and 300 µg.mL^{-1} conditions.

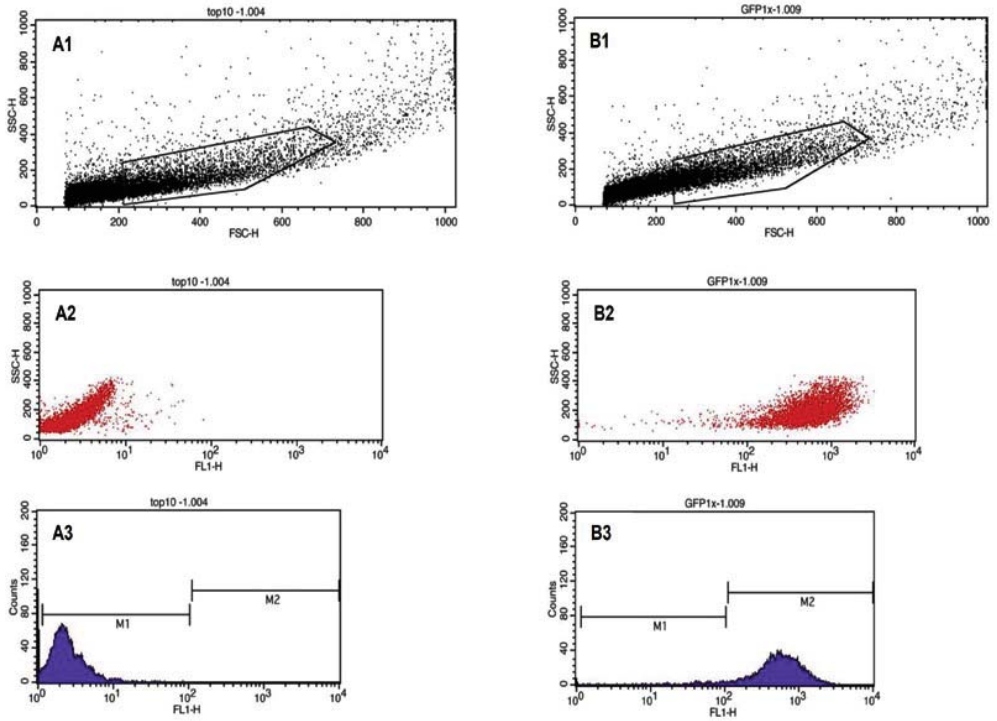

Figure 5. Flow cytometric analysis of GFP expression in the *E. coli*. Forward and side scatter graph of *E. coli TOP10* f (A1), GFP expressing *E. coli (B1)*, graphs of *E. coli TOP10* f as negative control (A2, 3) and GFP expression by *E. coli (B2,3)* in FL1 filter.

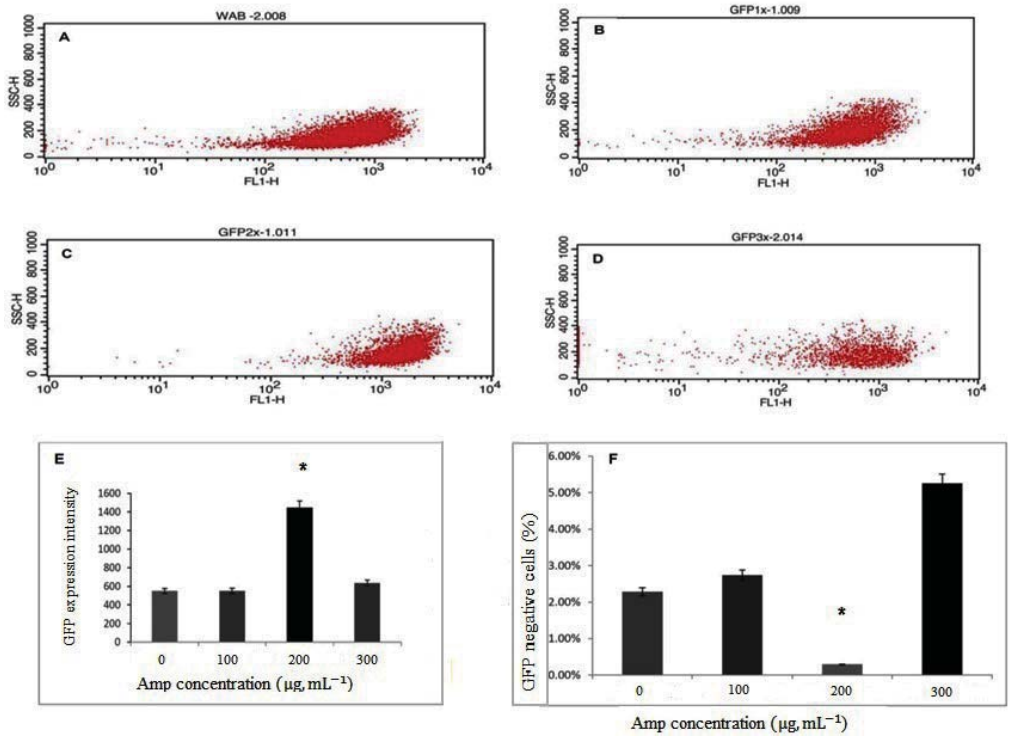

Figure 6. Flow cytometric representation of GFP expression intensity in different concentrations of Amp. Panels A, B, C and D: flow cytometric graphs of cultures in the presence of 0, 100, 200 and 300 μg.mL^{-1} Amp, respectively, graphs of GFP expression intensity, panel E and GFP negative cells (Lane F) in 0, 100, 200 and 300 μg.mL^{-1} Amp.

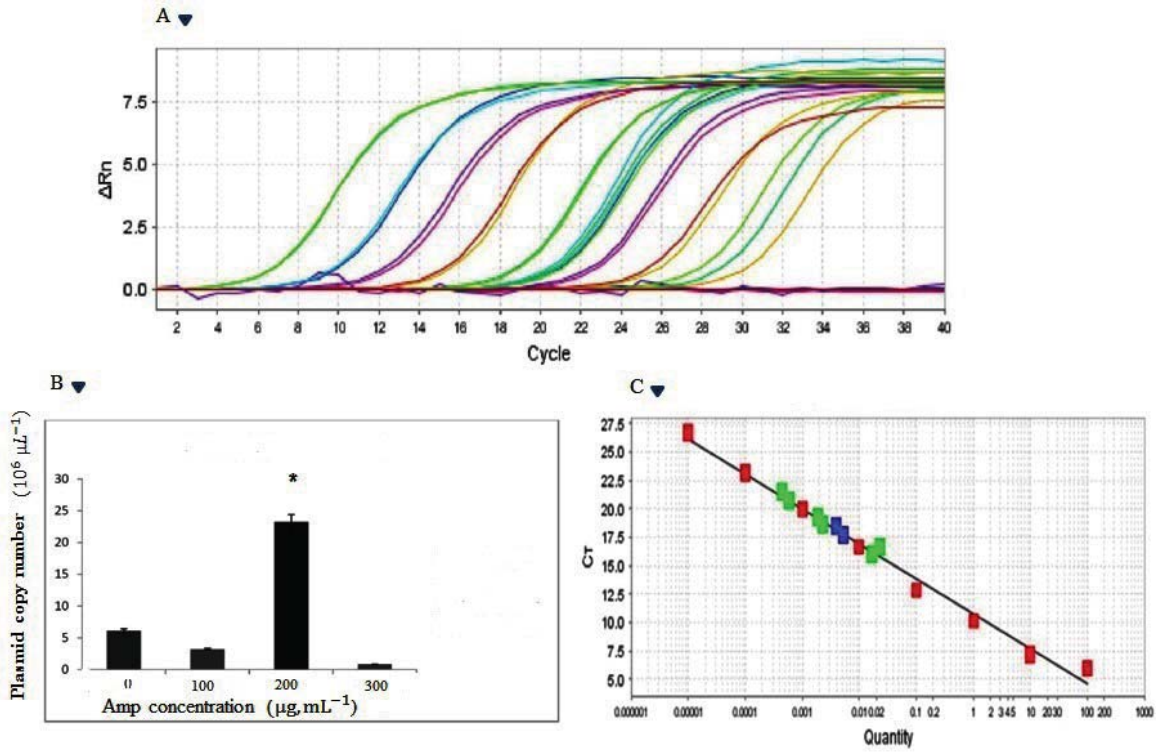

Figure 7. Real-time PCR and plasmid copy number estimation. As shown, the plasmid copy number in 200 μg.mL^{-1} Amp was considerably higher ($p < 0.01$) than 0, 100 and 300 μg.mL^{-1} . *$p < 0.01$(Amplification plot (A), Plasmid copy number (B) and Standard curve (C).

Table 1. The effect of ampicillin concentrations on plasmid yield and copy number. The concentration and copy number of plasmid in 200 µg.mL^{-1} was higher ($p < 0.01$) than100, 200, 300 µg.mL^{-1} and without Amp condition.

Ampicillin concentration	Plasmid concentration	Theoritical copy number	Experienced copy number
100 µg.mL^{-1}	69 ng.µL^{-1}	1.62 ×10^{10}	3.21 ×10^{9}
200 µg.mL^{-1}	164 ng.µL^{-1}	3.95 ×10^{10}	2.32 ×10^{9}
300 µg.mL^{-1}	41 ng.µL^{-1}	9.87 ×10^{10}	8.11 ×10^{9}
Without antibiotic	55 ng.µL^{-1}	1.32 ×10^{10}	6.07 ×10^{9}

4.4. Real-Time PCR

Amplification plot for samples obtained from cultures with different concentrations of Amp showed different C_T values (Fig. 7). The early C_T value for 200 µg.mL^{-1} indicated that plasmid copy number in this condition is higher than other concentrations. The absolute copy number for each sample was determined by comparison of samples C_T value with standards curve and it was higher for 200 µg.mL^{-1} than 0, 100 and 300 µg.mL^{-1} conditions. The absolute copy number of 0, 100, 200 and 300 µg.mL^{-1} Amp condition were 6.07×10^{9}, 3.21×10^{9}, 2.32×10^{10} and 8.11×10^{8}.µL^{-1}, respectively (Fig. 7).

4.5. Plasmid Extraction and Copy Number

Evaluation of plasmid extracts on 1% agarose gel electrophoresis and spectrophotometry showed that the yield of plasmid is significantly high ($p < 0.01$) in 200 µg.mL^{-1} Amp than others (Table 1). The more copy number and plasmid yield of 200 µg.mL^{-1} condition showed that 200 µg.mL^{-1} Amp is the optimum condition to obtain a high plasmid yield.

5. Discussion

These results showed how protein and plasmid yield of *E. coli* is affected in different concentrations of Amp and the necessities of optimization. One of the bottlenecks of protein and plasmid yield from *E. coli* can be attributed to the adjustments of cultures with the proper concentrations of Amp. Using the optimized concentration of Amp selection marker is an approach to decrease final cost of protein and plasmid extraction yields.

The high plasmid copy number and its maintenance are essential factors to achieve high protein and plasmid yields (13). As recently discussed by Gaimster and Summers (year is missing), the copy number and replication of ColE1 derived plasmids are controlled by small regulatory RNA. Such RNA molecules ensure the high copy numbers of the plasmid in the following bacterial generations(14). But Plasmid dimer or multimer (dimer catastrophe) is the main cause of plasmid instability and the generation of plasmid free cells (15). Selection pressure by adding some doses of selecting antibiotic inhibits and kills the plasmid free cells and prevents their accumulation. Although the generation of plasmid free cells in natural plasmids is low (10^{-5}-10^{-6}), it is unclear in manipulating plasmids (16). As the plasmid metabolic burden removes in plasmid free cells, they grow faster than plasmid containing heterologous protein expressing cells in the weak or without selection pressure. After a time period they constitute considerable part of culture and decrease the final yield of protein and plasmid. There are other mechanisms that can exacerbate the results from Amp as a selection marker; 1) Amp is chemically more sensitive than other antibiotics like carbenicillin and kanamycin and usually inactivates after overnight incubation. 2) The betalactamase enzymes secrete to the medium by resistant bacteria containing beta lactamase gene (plasmid) cleave and diminish the Amp concentration and plasmid free bacteria can grow and accumulate in this condition (17).

As mentioned, plasmid dimer or multimer formation is the main cause of plasmid instability. This problem was resolved by dimer resolution system developed in the dimer containing bacteria. Dimer resolution system converts plasmid multimers to monomers and helps to the uniform distribution of plasmids to both daughter cells, decreasing number of plasmid free cells (18). However, the dimer resolution system acts relatively slow in ColE1 derived pUC plasmids and needs to be completed before cell division. Moreover, timing between cell division and completion of dimer resolution is critical for the generation of plasmid free cells. Therefore, effective selection pressure presumably delays cell division and give enough time to completion of dimer resolution in extensively multimeric pUC plasmids. Determining the optimum dose and the effective selection pressure in each new designed system to inhibit generation and the overgrowth of plasmid free cells can increase

productivity of the system. Our results are consistent with the above mentioned, as the plasmid loss in low concentrations of Amp or media without Amp were the same (0 = 2.28% and 100 $\mu g.mL^{-1}$ Amp = 2.74%) and higher than 200 $\mu g.mL^{-1}$ Amp (0.29%).

Although theoretically high concentration of Amp ensures plasmid losing cells growth containment in long term cultures, it can increase the mortality rate and final cost. The bacterial mortality inclines in higher doses of Amp that affects the final yield. In 300 $\mu g.mL^{-1}$ Amp, the plasmid yield and GFP expression was lower and plasmid loss was higher than other doses. This is probably due to the high mortality rate.

As expected, the plasmid losing cells accumulate in the culture in the low/no selection pressure, decreasing protein and plasmid yields. These results showed that 200 $\mu g.mL^{-1}$ Amp is the optimum condition that can keep stable selection pressure for 18 h in this expression system, increasing final yields. Therefore, 200 $\mu g.mL^{-1}$ Amp is recommended to use in routine cultures to obtain better results when *E. coli Top10F* is being used as cloning and expression host. However, it seems that the optimization of Amp concentration is needed for any newly designed cloning and expression systems for protein production by *E. coli Top10F.*

Acknowledgment

This study was registered by grant number 192062, and funded by the deputy of Research in Isfahan University of Medical Sciences. The authors would appreciate for their official and financial supports.

References

1. Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol*. 1999 Oct;**10**(5):411-21.DOI: 10.1016/s0958-1669(99)00003-8
2. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;**5**:172. DOI: 10.3389/fmicb.2014.00172
3. Francis DM, Page R. Strategies to optimize protein expression in *E. coli*. *Curr Protoc Protein Sci*. 2010;29:5-24. DOI: 10.1002/0471140864.ps0524s61
4. Hannig G, Makrides SC. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol*. 1998 Feb;**16**(2):54-60. DOI: 10.1016/s0167-7799(97)01155-4
5. Jana S, Deb JK. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl Microbiol Biotechnol*. 2005 May;**67**(3):289-98. DOI: 10.1007/s00253-004-1814-0
6. Gerdes K, Larsen JE, Molin S. Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol*. 1985 Jan;**161**(1):292-8. DOI: 10.1016/0022-2836(86)90001-x
7. Sambrook J, Russell D. Molecular cloning: a laboratory manual. Q Rev Biol. 2001;76(3):348-9. DOI: 10.1086/394015.
8. Wilke MS, Lovering AL, Strynadka NC. β-Lactam antibiotic resistance: a current structural perspective. C*ur. Opin.Microbiol*. 2005;**8**(5):525-33. DOI: 10.1016/j.mib.2005.08.016
9. Chalfie M. Green fluorescent protein. *Photochem Photobiol*. 1995 Oct;**62**(4):651-6. DOI: 10.1111/j.1751-1097.1995.tb08712.x
10. Mandel M, Higa A. Calcium-dependent bacteriophage DNA infection. *J Mol Biol*. 1970 Oct 14;**53**(1):159-62. DOI: 10.1016/0022-2836(70)90051-3
11. http://scienceprimer.com/copy-number-calculator-for-realtime-pcr.
12. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. *Science*, 1994 Feb 11;**263**(5148):802-5. DOI: 10.1126/science.8303295
13. Sorensen HP, Mortensen KK. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol*. 2005 Jan 26;**115**(2):113-28. DOI: 10.1016/j.jbiotec.2004.08.004
14. Gaimster H, Summers D. Plasmids in the driving seat: The regulatory RNA Rcd gives plasmid ColE1 control over division and growth of its *E. coli* host. *Plasmid*. 2015 Mar;**78**:59-64. DOI: 10.1016/j.plasmid.2014.11.002
15. Summers DK, Beton CW, Withers HL. Multicopy plasmid instability: the dimer catastrophe hypothesis. *Mol Microbiol*. 1993 Jun;**8**(6):1031-8. DOI: 10.1111/j.1365-2958.1993.tb01648.x
16. Summers DK, Sherratt DJ. Multimerization of high copy number plasmids causes instability: *ColE1* encodes a determinant essential for plasmid monomerization and stability. *Cell*, 1984 Apr;**36**(4):1097-103. DOI: 10.1016/0092-8674(84)90060-6
17. Korpimaki T, Kurittu J, Karp M. Surprisingly fast disappearance of beta-lactam selection pressure in cultivation as detected with novel biosensing approaches. *J Microbiol Methods*. 2003 Apr;**53**(1):37-42. DOI: 10.1016/s0167-7012(02)00213-0
18. Summers D. Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol Microbiol.*1998 Sep;**29**(5):1137-45. DOI: 10.1046/j.1365-2958.1998.01012.x