Influence of The Expression Vector And of Its Elements On Recombinant hPRL Bacterial Synthesis: The Co-Directional Orientation of Replication And Transcription is Highly Critical

Regina Affonso (reginaffonso@yahoo.com.br)
Nuclear and Energy Research Institute: Instituto de Pesquisas Energeticas e Nucleares
https://orcid.org/0000-0002-9264-4262
Miriam Fussae Suzuki
IPEN: Instituto de Pesquisas Energeticas e Nucleares
Geraldo Santana Magalhães
Butantan Institute: Instituto Butantan
Paolo Bartolini
IPEN: Instituto de Pesquisas Energeticas e Nucleares

Original article

Keywords: human prolactin, Escherichia coli, λPL and tac promoters, co-directional collisions, head-on collisions

Posted Date: May 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-483600/v1

Influence of the expression vector and of its elements on recombinant hPRL bacterial synthesis: the co-directional orientation of replication and transcription is highly critical

Regina Affonso*, Miriam Fussae Suzuki, Geraldo Santana Magalhães, Paolo Bartolini

[1]Biotechnology Center, Instituto de Pesquisas Energéticas e Nucleares, IPEN - CNEN/SP, São Paulo, SP, Brazil
[2]Immunopathology Laboratory, Butantan Institute, São Paulo, SP, Brazil

* Address for correspondence: Dr. Regina Affonso, Biotechnology Center, IPEN-CNEN, Av. Prof. Lineu Prestes, 2242, Cidade Universitária, 05508-000, São Paulo, Brazil. Phone: 55 11 2810-5860, e-mail: reginaffonso@yahoo.com.br.

reginaffonso@yahoo.com.br	ORCID 0000-0002-9264-4262
mfsuzuki@ipen.br	ORCID 0000-0002-7467-3457
geraldo.magalhaes@butantan.gov.br	ORCID 0000-0001-9455-5493
pbartoli@ipen.br	ORCID 0000-0002-5258-1527

Abstract

The aim of the present work was to define a bacterial expression system that is particularly efficient for the synthesis of human prolactin (hPRL). In previous work, the synthesis of rec-hPRL by the p1813-hPRL vector in *E. coli* HB2151 was >500 mg/L, while it was much lower now (2.5-4-fold), in the strains RB791 and RRI. The highest positive influence on rec-hPRL synthesis was due to the transcription-replication co-orientation of *hPRL* cDNA and the *ori*/antibiotic resistance gene, responsible for up to a ~5-6-fold higher expression yield. In conclusion, this work confirmed that each bacterial strain of

E. coli has a genetic set that can allow a different level of heterologous protein synthesis. The individual study of each element indicated that its action critically depends on the reading orientation in which it is located inside the vector: co-directional orientation of replication and transcription, in fact, greatly increased the level of rec-hPRL expression.

Keywords: human prolactin, *Escherichia coli*, λP_L and *tac* promoters, co-directional collisions, head-on collisions

Introduction

Recombinant DNA technology for the expression of heterologous proteins began almost 40 years ago and this technology is still widely used and expanding. Bacterial expression systems continue to be the best choice for the synthesis of unmodified recombinant proteins due to their high level of expression, low production cost and suitability for laboratory investigation and pre-industrial scale production (Sørensen and Mortensen 2005; Chen 2012). Prolactin (PRL), as well as Growth Hormone (GH), based their production and commercialization on this technology. PRL and GH are members of the family of pituitary hormones, having similarity in their amino acid sequence, structural and biological features (Goffin et al. 2002; Ben-Jonathan et al. 2008). Human PRL (hPRL) is a single-chain protein which consists of 199 amino acids and contains three disulfide bonds (Bernard et al. 2015). This molecule is involved in diverse physiological functions, such as lactation and reproduction, among others. It has a single N-glycosylation site, whose N-glycan is not essential for its known functions (Price et al. 1994). The *Escherichia coli* (*E. coli*) strain and vector type used for hPRL expression are very important for an optimized synthesis and many options are available for research and production (Valdez-Cruz et al. 2010; Rosano and Ceccarelli, 2014), also considering

that this hormone is one of those most frequently determined in clinical assay laboratories (Arthuso et al. 2012).

E. coli strains express heterologous genes differently because of specific genomic characteristics that may limit recombinant protein yield and few studies have evaluated these (Makrides, 1996; Rosano and Ceccarelli, 2014). Different vectors also have intrinsic variables that can cause a higher or lower production of heterologous proteins. The most common vector elements are (*i*) transcriptional regulators: promoters with chemical or thermal induction, transcriptional terminators and transcriptional antiterminators; (*ii*) translational regulators: mRNA translation initiators, translation enhancers and translational terminators and (*iii*) other factors such as the origin of replication, codon bias and antibiotic resistance genes (Makrides, 1996; Rosano and Ceccarelli, 2014).

The present work was divided into two steps: analysis of an efficient p1813-hPRL vector/*tac*-promoter/*E. coli* strain system and evaluation of the expression-enhancing elements that are needed to efficiently synthesize rec-hPRL with basis on the λP_L promoter.

The efficient p1813-hPRL vector, containing an integrated repressor system and *hPRL* cDNA under control of the *tac* promoter (Affonso et al. 2018), was introduced in three different *E. coli* strains that had been used to express GH and PRL in other studies and its expression yields were determined (Morganti et al. 1998; Soares et al. 2008).

A $p\lambda P_L$ vector series that always contained the same $p\lambda P_L$-hPRL sequence was evaluated, in a second step, in comparison with the same elements that are present in the p1813-hPRL vector and that act in the transcription or translation phase. The *T1/T2* transcription terminators are composed of tandem sequences that make a clamp that interrupts mRNA transcription, while the *5S* sequence, located in front of these terminators, increases expression efficiency by forming a loop (Orosz et al. 1991; El Hage

et al. 2008). Translation enhancer elements that increase the efficiency of mRNA translation into the protein of interest by binding to the rRNA/tRNA complex, were also evaluated. The *g10* sequence, for example, is known for increasing the translation efficiency by as much as 40-fold (Olins and Rangwala 1989; Graham et al. 1995). Genes that confer resistance to a particular antibiotic were also considered to be very important for maintaining the expression of the protein of interest. Ampicillin and kanamycin are often used and act by entering to the bacterium periplasmic space and can be inactivated in this region by β-lactamase and aminoglycoside phosphotransferase, respectively (Sørensen and Mortensen 2005; Rosano and Ceccarelli, 2014).

These two vectors made possible, therefore, the study of vectors capable of providing different levels of recombinant protein synthesis, hPRL in particular, indicating the most efficient vector-promoter-strain combination system.

Materials and methods

Design and construction of the vectors

Fig. 1 (A) shows the strategy that was used to construct the pλP_L-hPRL vector series. All the elements that were inserted in the pλP_L-hPRL vector were selectively taken from the highly efficient p1813-hPRL vector (Affonso et al. 2018), to possibly increase rec-hPRL expression; when ligated, these parts formed the vectors pλP_L-hPRL I, pλP_L-hPRL II and pλP_L-hPRL III. The origin of replication (*ori*) for all vectors studied in this work came from plasmid pBR322 (Crowl 1986; Sambrook and Russel 2001) (Table 1)

Fig. 1

Table 1

The original λP_L plasmid (pλP_L) used in the construction did not have the *g10* and *5S/T1/T2* elements, lacked *hPRL* cDNA and encoded the Amp^R gene. *hPRL* cDNA was cleaved from p1813-hPRL by *Eco*RI and *Hin*dIII and inserted into pλP_L by using T4 ligase (New England Biolabs, USA) obtaining pλP_L-hPRL I, as shown in Fig. 1 A. To obtain pλP_L-hPRL II, the *g10* element was cleaved from p1813-hPRL by the enzymes *Xba*I and *Hin*dIII and inserted in pλP_L-hPRL I still using T4 ligase. The Amp^R gene, in pλP_L-hPRL II, was replaced with Kan^R gene and with the *5S/T1/T2* sequence obtained from p1813-hPRL by using enzymes *Aat*II, *Hin*dIII and T4 ligase, resulting in pλP_L-hPRL III. The pλP_L-hPRL IV vector was constructed from pλP_L-hPRL III withdrawing the *5S/T1/T2* sequence. The terminal region of the Kan^R gene which had been partly deleted, was reconstructed with the following primers:

forward 5'ttttatttaagcttcacgctgccgcaagc3'

reverse 5'cctcgtcctgcagttcattc3'

In these primers, the underlined regions correspond to the restriction sites for the enzymes *Hin*dIII and *Pst*I, respectively. The PCR conditions used were: 94°C for 1.5 min and 30 cycles of 94°C for 15 sec, 55°C for 15 sec and 72°C for 30 sec (Taq DNA polymerase, Invitrogen, USA). The enzymes *Hin*dIII and *Pst*I were used to cleave the PCR product and the pλP_L-hPRL III vector, while the final insertion (cloning) was done with T4 ligase.

Escherichia coli strains

Four different strains of *Escherichia coli* (HB2151, RRI, RB791 and BL21(DE3)) and two promoters (*tac* and λP_L) were used, since one of the elements that indirectly acts on protein synthesis, the promoter, needs a compatible strain.

Genotypes:

HB2151: [araD (lac pro AB), lacIq Δ(lacZ) M15, thi/F' $proA^+$ B^+ / fhua2Δ(thiΔ(hsdMSmerB) ($^-$ rk $^-$ mk $^-$ McrBC-) K12]. (Stratagene, Agilent Technologies, Inc., E.U.A).

RRI: *F^- mcrB mrr hsdS20(r_B^- m_B^-) recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20(Sm^R) glnV44 λ^- $recA^+$* (New England Biolabs, USA).

RB791: [*pQE30: adh*], BL21(DE3) [*pPal7:Dm_MFE-2, lacl'iLs*]. (cgsc.biology.yale.edu)

BL21(DE3): F- ompT hsdSB (rBmB-) gal dcm (DE3) (One shot BL21 catalogue, Invitrogen, USA).

Expression conditions of p1813-hPRL and of the pλP_L-hPRL and pET-hPRL vector series

The transformed *E. coli* HB2151 strain (Luck et al. 1986), RRI (Sambrook and Russel 2001) and RB791 (Fu et al. 1992), containing the p1813-hPRL plasmid, were grown in LB medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) supplemented with kanamycin (50 μg/mL) at 37°C, while shaking at 180 rpm. When the optical density (OD) reached 0.4–0.8 A_{600}, the cells were induced with 0.1 mM isopropyl-β-D-thiogalactoside (IPTG, Sigma, São Paulo, Brazil) and cultured for 9 h at 37°C. The expression of rec-hPRL in the p1813-hPRL/HB2151 system was induced by different concentrations of IPTG (0.1, 1.0, 1.5 and 2.0 mM) and at seven activation times (3, 5, 6, 7, 8, 9 and 16 h), in order to determine the most efficient conditions. The RRI strain, that was transformed with the pλP_L-hPRL vector series, was grown in LB medium supplemented with ampicillin (100 μg/mL) or kanamycin (50 μg/mL) at 30°C. When the OD_{600} reached 0.4–0.8 A_{600}, the cells were induced by increasing the temperature to 42°C for 9 h.

All cultures were centrifuged at 4,000 g for 5 min at 4°C, and the pellets were processed or stored at -20°C (Carvalho et al. 2014). The induced and non-induced cultures were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting (Carvalho et al. 2014).

The synthesis of pET3a-hPRL and pET3d-hPRL vectors (Fig. 1, B) was carried out by Biomatik, Wilmington, Delaware, USA; transformed *E. coli* BL21(DE3) strain, containing these plasmids, was cultivated as done for p1813-hPRL in HB2151. In this case, the BL21(DE3) is the strain that carries the gene for the T7 RNA polymerase, under control of the lacUV5 promoter (activated by IPTG) in bacterial chromosomal DNA.

SDS-PAGE and Western blot analysis

E. coli, transformed with p1813-hPRL or pλP_L-hPRL, were analyzed by SDS-PAGE under reduced conditions, and these analyses were conducted using a 12% denaturing polyacrylamide gel and Coomassie Blue (Sigma, São Paulo, Brazil) staining (Carvalho et al. 2014).

For Western blotting, the separated protein bands from the polyacrylamide gel were transferred to a nitrocellulose membrane, which was incubated with polyclonal anti-hPRL rabbit antiserum (Santa Cruz Biotechnology, Dallas, USA) at a 1:500 dilution (Carvalho et al. 2014). The samples were then incubated with ^{125}I-labeled Protein A that had been obtained via the Chloramine-T method (Morganti et al. 1996). All samples were analyzed at a concentration of $4.8x10^7$ cells/mL (A_{600}=0.06).

Protein quantification

The intensity of the bands of reduced pituitary hPRL standards, of the corresponding electrophoretic bands of the rec-hPRL samples and of total bacterial proteins were

determined using a computerized laser scanning densitometer (Model CS-9301PC Dual Wavelength, Shimadzu, Japan). The amount of rec-hPRL in each extract was thus estimated against a reference standard curve that had been constructed using different amounts of pituitary hPRL. Triplicate determinations were performed for all experiments.

Plasmid DNA quantification

The plasmid level was determined as follows: 125 µL of competent cells were transformed with 1 ng of each one of the two plasmid types (pET3a-hPRL and pET3d-hPRL), by thermal shock, adding 200 µL of LB medium and incubating for 1 h at 180 rpm at 37°C. Based on the methodology of Stueber and Bujard (1989), these cultures were inoculated into 200 mL of LB medium containing ampicillin (100 µg/mL). When the OD reached ~1.6 A_{600}, samples were collected from each of the vector sets, these samples being called Initial (I). Then, these cultures were induced for 9 hours as described in item 2.3 (Induced- Ind), cultivated in parallel with non-induced control cells (Control – C). All samples were adjusted to a total of 3 x 10^9 cells (considering that A_{600} = 1 corresponds to 8 x 10^8 cells/mL) and then centrifuged at 4,000 g for 10 min at 4°C. Plasmid DNA was then extracted from pellets using a Mini prep kit (Model 2000, Thermo Scientific, Uniscience, SP, Brazil) and quantified on a NanoDrop Spectrophotometer (Model 2000, Thermo Scientific, Uniscience, SP, Brazil). All determinations were carried out in triplicate.

Statistical analysis

Data were expressed as the mean ± S.E. (S.E., standard error) of at least three independent experiments. Statistical significance was determined by using the unpaired Student's *t* test. A $p \leq 0.05$ was conventionally considered statistically significant.

Results

Determination of the expression efficiency of p1813-hPRL, based on the tac promoter, in three different E. coli strains: HB2151, RRI and RB791

The p1813-hPRL vector, considering that the characteristics of its elements allow high expression of recombinant proteins (Pereira et al. 2014; Affonso et al. 2018), had its efficiency evaluated in three strains, including RRI. The most efficient conditions were obtained with 0.1 mM IPTG and 9 h induction time (data not shown).

The expression levels obtained with p1813-hPRL in the three strains are shown in Table 2, the volumetric yield of rec-hPRL in HB2151 being 4.1- and 2.6- fold higher than in RRI and RB791, respectively. The RB791 and HB2151 strains are used in general for the expression of proteins with basis on the *tac* promoter and, in the present study, produced significantly more prolactin than the RRI strain ($p=0.02$). The difference in mass fraction between RRI and RB791 was, however, not significant ($p=0.05$). As observed, the synthesis of rec-hPRL in HB2151, harboring the p1813-hPRL vector, provided the highest specific (132.7 μg/mL.A_{600}) and volumetric (> 500 mg/L) yields; the mass fraction was also the highest, suggesting that this strain, whose expression is based on the *tac* promoter, is the most suitable for the expression of this recombinant protein.

Table 2

Analysis of rec-hPRL expression via p1813 and pλP_L vectors, considering their different structural components

Description of the structural components

Strong promoters, such as λP_L and *tac*, are widely used in basic research and industrial production. However, the use of IPTG in the large-scale production of human therapeutic proteins is undesirable, mainly because of its cost and potential toxicity (Makrides 1996; Li 2018). Thermal induction by the λP_L promoter is the most desirable in research and industry due to its lower cost and practicality.

Effect of individual vector components on rec-hPRL expression yields

SDS-PAGE and Western blot analyses confirmed the presence of a protein that was expressed in inclusion bodies; it had the same molecular weight as expected for rec-hPRL and presented the expected immunological activity (Fig. 2).

Fig. 2

Rec-hPRL expression yields from the p1813-hPRL and pλP_L-hPRL vector series and their important components are shown in Table 3 as determined by SDS-PAGE densitometry (see Fig. 2). Vectors that have a comparative element are listed in this table two-by-two. Although the p1813-hPRL vector highly expressed rec-hPRL in the HB2151 strain with high mass fraction, its expression in the RRI strain was markedly lower.

Table 3

The volumetric yield of rec-hPRL synthesized in BL21(DE3) by the two vectors illustrated in Fig. 1B, was analyzed, providing 178 mg/L for pET3a-hPRL and 399 mg/L

for pET3d-hPRL, the difference being 124%. In fact, the pET3d-hPRL vector synthesized a 2.24-fold higher amount of recombinant protein than pET3a-hPRL. These results support our hypothesis that, when the DNA sequences of the Amp^R gene and *ori* are not in frame with the *hPRL* expression cassette, this can negatively influence the transcription process, as better described in the Discussion.

Plasmid DNA quantification (Fig. 3) was also obtained after 9 h induction, as described in Materials and methods (item 2.6). The plasmid levels of the non-induced cultures (Controls) were 23-40 % higher than the initial measurements. In induced cultures, the level of plasmid DNA was higher in pET3a-hPRL than in pET3d-hPRL according to the different properties of each vector. The pET3a-hPRL produced more than 2.5-fold plasmid DNA in relation to control, while the pET3d-hPRL vector produced less than 2.0-fold. Plasmid DNA of induced pET3d-hPRL was, therefore, much lower (~40%) when compared to that obtained with pET3a-hPRL.

Fig.3

Discussion

Our research group has studied the expression of recombinant pituitary hormones: human growth hormone (hGH), prolactin (hPRL), thyrotropin (hTSH), folliculotropin (hFSH) and luteotropin (hLH), some of them specifically expressed in bacterial systems (Morganti et al. 1998; Oliveira et al. 1999; Soares et al. 2008; Suzuki et al. 2012; Affonso et al. 2018). One of the objectives of the present study was to compare the effects of *tac* and λP_L promoters on the cytoplasmic expression of rec-hPRL, recalling that the *tac*

promoter is activated by IPTG, while the λP_L promoter is activated by a temperature increase (42°C). The repression of these promoters is also specific: *tac* is inactivated by a protein expressed by the *lacI*^{q} gene, which is encoded in the same p1813-hPRL vector (Sambrook and Russel, 2001), while λP_L is repressed by a temperature-sensitive protein that is expressed by the *ci857* gene, contained in a specific plasmid (pRK248cIt2) which can be present in the RRI strain (Crowl 1986). The λP_L promoter works specifically in the RRI strain and will be used for comparing the expression level obtained with the *tac* promoter in the same strain. The evaluation of p1813-hPRL vector efficiency in this strain will therefore be fundamental for this comparison because this vector in strain HB2151 had high expression (Affonso et al 2018).

The bacterial strain can be a decisive factor for the high expression of recombinant proteins (Makrides 1996; Sørensen and Mortensen 2010). In fact, in our hands, hGH secretion in the periplasm of four different strains presented specific yields of 2.8 (RRI), 1.2 (HB2151) and 3.9 μg/mL.A_{600} (RB791 and W3110), all carrying a λP_L vector (Soares et al. 2003). In the present study, the RRI strain with p1813-hPRL vector produced 2 times less bacterial protein and almost 3 times less rec-hPRL in relation the production in the HB2151, confirming the literature data.

The analysis of the individual components of the vector in the expression yields of rec-hPRL, under the control of the $p\lambda P_L$ promoter, showed that pλP$_L$-hPRL III was the best arrangement between components and reading of the DNA sequence.

Comparing rec-hPRL expression by p1813-hPRL and pλP$_L$-hPRL III vectors with two different promoters in the RRI strain (comparison # 1), we can observe that λP_L is significantly more efficient than the *tac* promoter (p=0.003), although the difference in mass fraction is not significant (p=0.36). The volumetric yield (171 mg/mL and 131 mg/mL for the λP_L and the *tac* promoters, respectively) was ~30% higher for λP_L.

Valdez-Cruz et al. (2010) reported cytoplasmic yields of twenty-five different recombinant proteins that were obtained with the *$\lambda P_L/\lambda P_R$/cI857* thermoinduced expression system. A wide range of yields was found among the proteins that were analyzed. Some of these yields were: 0.95 g/L for β-galactosidase, 3 mg/L for Integration Host Factor -α and β, while for Green Fluorescent Protein the yields were 7, 30, 45, 50, 68 and 273 mg/L, due to different activation conditions. These data demonstrate that there is no constant expression level, not even for the same protein.

The most interesting result was, however, comparison # 4, in which ampicillin and kanamycin antibiotic resistance genes are compared. The pλP_L-hPRL IV vector produced in fact almost 6-fold more rec-hPRL than pλP_L-hPRL II. It is of note that the mass fractions from all vectors containing the *Kan^R* gene were approximately 2.0 or 2.6-fold higher than those related to the *Amp^R* gene. It should also be mentioned that the RRI strain produced the same amount of bacterial protein in the same three vectors based on *Kan^R* gene (p1813-hPRL, pλP_L-hPRL III and IV): an average of 156.7 ± 10.4 μg/mL.A_{600}. This indicates that the constant and limited effect of the elements analyzed in the RRI strain may be due to intrinsic characteristics of the strain.

The insertion of specific sequences into expression vectors markedly enhances the production of recombinant proteins in *E. coli*. Some of these DNA elements (*g10* translation enhancer and *5S/T1/T2* transcription terminators) were analyzed here under the control of the *λP_L* promoter. When used in the pλP_L-hPRL vector series, the *g10* element (comparison # 2) seems, however, to be responsible for a lower (24%) rec-hPRL synthesis, contrary to results described in the literature (Olins and Rangwala 1989). Transcription terminators *5S*, followed by the *T1* and *T2* sequences in tandem were, instead, able to increase the specific rec-hPRL expression yield by 13%. It is known that transcription terminators improve the stability of mRNA by forming a stem loop and can

substantially increase the level of protein production (Sambrook and Russel 2001; Sørensen and Mortensen 2010).

As already pointed out, the most unexpected result obtained in this study was with pλP_L-hPRL II and pλP_L-hPRL IV, apparently showing that the orientation of the *Amp^R/ori* or *Kan^R/ori* cassettes may have a significantly positive or negative influence on the expression of rec-hPRL (Table 3, comparison # 4).

To confirm this hypothesis, the DNA sequence that confers resistance to ampicillin was reversed via two specifically constructed pETs vectors. pET3a and pET3d have the same elements, both being based on the *tac* promoter, but the reading frame for the *Amp^R* gene and for the *ori* is opposite to that of the introduced foreign cDNA in the case of pET3a and co-directional in the case of pET3d. The cDNA for *hPRL* was thus inserted into each one of these two vectors (Fig. 1, B) in order to prove the above mentioned hypothesis.

The levels of plasmids DNA in the two pETs vectors *Amp^R* gene→*ori*→*hPRL* cDNA versus *Amp^R* gene→*ori*←*hPRL* cDNA were determined at the maximum level of absorbance (1.6 A_{600}) and, before induction, practically provided equivalent values: 68.3±2.4 ng/μL and 67.8±1.8 ng/μL, i.e. 1.37 μg/mL.A_{600} and 1.36 μg/mL.A_{600}, respectively (Fig. 3). One can speculate that the higher plasmid levels of non-induced cultures was due to the lack of heterologous protein expression: the machinery, therefore, did not prevent the replication process. Upon induction, pET 3a-hPRL machinery favored the maintenance of bacterial life (in the antibiotic medium) more than heterologous protein expression. In conclusion, the difference in rec-hPRL expression, between these two vectors and, consequently, between pλP_L-hPRL II and pλP_L-hPRL IV, was not due to any direct influence of the antibiotic resistance gene but to the head-on replication-transcription collision between *ori* and *hPRL* cDNA. Chromosomal and bacterial DNA

topology, especially in the transition region between *ori* replication forks and DNA transcription machineries when these are working simultaneously, have shown that there is a conflict that can lead to DNA stopping, instability or breakdown (Mirkin and Mirkin 2005; Deepak et al. 20015; Achar and Foiani 2017; Wein et al. 2019). Deepak et al. (2015) explained this conflict as collisions and defined two ways of how these collisions occur: co-directional collisions, when the gene is transcribed in the same direction as the replicon machinery, and head on collisions, when the gene is transcribed in the opposite direction of the replicon machinery.

Achar and Foiani (2017) data on the positive and negative supercoiling formed during co-directional and head-on collision can support our results. We believe that, in head-on collision, competition between replication and transcription depend on many factors, such as replicon and promoter power, antibiotic pressure and nutrient availability, among others. This mechanism can, therefore, explain the low expression presented by pλP_L-hPRL II, in which the choice between replication and transcription was for more replication (Mirkin and Mirkin 2005).

The choice of a perfect match between bacterial strain and the components of a given vector is difficult to be determined *a priori*. In the present work, the highest production of rec-hPRL is apparently due to the optimized vector/strain system: p1813-hPRL, under control of the *tac* promoter, in the HB2151 strain. Analysis of the elements that provide an increase in protein expression, such as the enhancer elements, also depends on the co-orientation between transcription and replication machineries. The replication process opens the DNA double strand, leading to the formation of a negative supercoil in the reading direction of the DNA double strand and a positive supercoil in the opposite direction. This was observed by comparing the pλP_L-hPRLIII and pλP_L-hPRLIV vectors

concerning the effect of transcription terminators, which might have minimized the co-directional collision.

We can conclude that the thermo-inducible *λPL* promoter, frequently considered to be quite inefficient when compared to the *tac* promoter, became much more competitive and even more efficient, especially if the goal is to optimize cost / benefit. The elements that allow an increase in the synthesis of recombinant proteins must be utilized with an appropriate bacterial strain, especially considering that the co-directional orientation of the replication and transcription elements can be largely responsible for increasing the level of protein expression.

Acknowledgements

We thank the members of the Biotechnology Center for experimental support.

Authors' contributions

MFS and GFM intellectual contribution, RA experimental procedures, and PB and RA data analysis, drafting the manuscript, critical intellectual contribution, and thorough and final approval of the version to be published.

Funding

This work was supported by: The National Nuclear Energy Commission (CNEN); The National Council for Scientific and Technological Development (CNPq); The State of São Paulo Research Foundation (FAPESP), São Paulo, Brazil (project number 2017/50332).

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Reference

Achar YJ, Foiani M (2017) Coordinating Replication with Transcription. Adv Exp. Med Biol 1042:455-487.

Affonso R, Soares CR, Ribela MTC, Bartolini P (2018) High production and optimization of the method for obtaining pure recombinant human prolactin. Protein Expr Purif 152:131-136.

Arthuso FS, Bartolini P, Soares CR (2012) Laboratory production of human prolactin from CHO cells adapted to serum-free suspension culture. Appl Biochem Biotechnol 167(8):2212-2224.

Ben-Jonathan N, LaPensee CR, LaPensee EW (2008) What can we learn from rodents about prolactin in humans? Endocr Rev 29(1):1-41.

Bernard V, Young J, Chanson P, Binart N (2015) New insights in prolactin: pathological implications. Nat Ver Endocrinol 11(5):265-275.

Carvalho CV, Ricci G, Affonso R (2014) Guia de práticas em Biologia Molecular, 2nd edn Yends, São Paulo, Br.

Chen R (2012) Bacterial expression systems for recombinant protein production: *E. coli* and beyond. Biotechnol Adv 30(5):1102-1107.

Crowl R (1986) Expression of human interferon genes in *E. coli* with the lambda P_L promotor. Method Enzimol 119:376-383.

Deepak K, Yashvantha KJ, Sathisha TN, Veeresh BA, Desai SA (2015) The Conflict of Transcription-Replication Coordination: Understanding in the Perspective of Genome Integrity. Clon Transgen 4(2):1-5.

El Hage A, Koper M, Kufel J, Tollervey D (2008) Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. Gene Dev 22(8):1069-1081.

Fu J, Togna AP, Shuler ML, Wilson DB (1992) *Escherichia coli* host cell modifications in continuous culture affecting heterologous protein overproduction: A population dynamics study. Biotechnol Prog 8:340-346.

Goffin V, Binart N, Touraine P, Kelly PA (2002) Prolactin: the new biology of an old hormone. Annu Rev Physiol 64:47-67.

Graham RW, Greenwood JM, Warren RA, Kilburn DG, Trimbur DE (1995) The pTug A and pTug AS vectors for the high-level expression of cloned genes in *Escherichia coli.* Gene 158:51-54.

Li X (2018) Bioengineering of FGFs and New Drug Developments. Fibroblast Growth Factors 477-558. Elsevier Inc. https://doi.org/10.1016/B978-0-12-816142-5.00008-4.

Luck DN, Ngsee JK, Rottman FM, Smith M (1986) Synthesis of bovine prolactin in *Escherichia coli.* DNA 5:21-28.

Makrides SC (1996) Strategies for Achieving High-Level Expression of Genes in *Escherichia coli.* Microbiol Rev 60(3):512-538.

Mirkin EV, Mirkin SM (2005) Mechanisms of transcription-replication collisions in bacteria. Mol Cell Biol 25(3):888-895.

Morganti L, Huyer M, Gout PW, Bartolini P (1996) Production and characterization of biologically active Ala-Ser-(His)6-Ile-Glu-Gly-Arg-human prolactin (tag-hPrl) secreted in periplasmic space of *Escherichia coli*. Biotechnol Appl Biochem 23:67-75.

Morganti L, Soares CRJ, Affonso R, Gout PW, Bartolini P (1998) Synthesis and characterization of recombinant, authentic human prolactin secreted into the periplasmic space of *Escherichia coli*. Biotechnol Appl Biochem 27:63-70.

Olins PO, Rangwala SH (1989) A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of lacZ gene in *Escherichia coli.* J Biol Chem 264 (29):16973-16976.

Oliveira JE, Soares CR, Peroni CN, Gimbo E, Camargo IM, Morganti L, Bellini MH, Affonso R, Arkaten RR, Bartolini P, Ribela MT (1999) High-yield purification of biosynthetic human growth hormone secreted in *Escherichia coli* periplasmic space. J Chromatogr A 852(2):441-450.

Orosz A, Boros I, Venetianer P (1991) Analysis of the complex transcription termination region of the *Escherichia coli* rrnB gene. Eur J Biochem 201(3):653-659.

Pereira LM, Silva LR, Alves JF, Marin N, Silva FS, Morganti L, Silva IDCG, Affonso R (2014) A simple strategy for the purification of native recombinant full-length human RPL10 protein from inclusion bodies. Protein Expres Purif 101:115-120.

Price AE, Logvinenko KB, Higgins EA, Cole ES, Richards SM (1995) Studies on the microheterogeneity and in vitro activity of glycosylated and nonglycosy-lated

recombinant human prolactin separated using a novel purification process. Endocrinology 136:4827–4833.

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:172.

Sambrook J, Russel DW (2001) Cold Spring Harbor Laboratory Press, 3rd edn New York, USA.

Soares CR, Gomide FI, Ueda EK, Bartolini P (2003) Periplasmic expression of human growth hormone via plasmid vectors containing the lambda PL promoter: use of HPLC for product quantification. Protein Eng16(12):1131-8. doi: 10.1093/protein/gzg114

Soares CRJ, Ueda EKM, Oliveira TL, Gomide FIC, Heller SR, Bartolini P (2008) Distinct human prolactin (hPRL) and growth hormone (hGH) behavior under bacteriophage lambda PL promoter control: Temperature plays a major role in protein yields. J Biotechnol 133(1):27-35.

Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli.* J Biotechnol 115:113-128.

Stueber D, Bujard H (1982) Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. EMBO J 1(11):1399-1404.

Suzuki MF, Arthuso FS, Oliveira JE, Oliveira NA, Goulart HR, Capone MV, Ribela MT, Bartolini P, Soares CR (2012) Expression, purification, and characterization of authentic mouse prolactin obtained in *Escherichia coli* periplasmic space. Biotechnol Appl Biochem 59(3):178-185.

Valdez-Cruz NA, Caspeta L, Pérez NO, Ramírez OT, Trujillo-Roldán MA (2010) Production of recombinant proteins in *E. coli* by the heat inducible expression system

based on the phage lambda pL and/or pR promoters. Microb Cell Fact 9:18. doi: 10.1186/1475-2859-9-18

Wein T, Hülter NF, Mizrahi I, Dagan T (2019) Emergence of plasmid stability under non-selective conditions maintains antibiotic resistance. Nat Commun 10(1):2595.

Figure Legends

Fig. 1 Schematic representation of the vectors. A) p1813-hPRL vector and pλP_L-hPRL vector series construction, showing the location of all elements; B) Schematic representation of the vectors pET3a-hPRL and pET3d-hPRL. Abbreviations: *hPRL*, Prolactin hormone cDNA sequence; amp^R, ampicillin resistance gene; kan^R, kanamycin resistance gene; *ori*, origin of replication; λP_L, lambda promotor P_L; *tac*, *tac* promotor; *g10*, translator enhancer; *5S/T1/T2*, transcription terminators; $LacI^q$, lac repressor protein DNA sequence; bp, base pair.

Fig. 2 SDS-PAGE (A and B) and Western blot (C) analysis of rec-hPRL expression by p1813-hPRL and pλP_L-hPRL vector series, in either HB2151 or RRI *E. coli* strains. Lanes M3, M2 and M1 contained 3, 2 and 1 μg of protein markers, respectively. The 23 kDa protein is pituitary PRL that was used as standard. Induced (+) and non-induced (-) conditions are shown.

Fig. 3 Quantifications of plasmid DNA from pET3a-hPRL and pET3d-hPRL, obtained in cultured *E. coli* BL21(DE3) strain. A) graph and B) table with plasmid DNA levels: (I), initial cultures when reaching 1.6 A_{600}; (C), non-induced control cultures after 9 h of incubation; (Ind**),** IPTG-induced cultures after 9 h of incubation. (n=3).

Figures

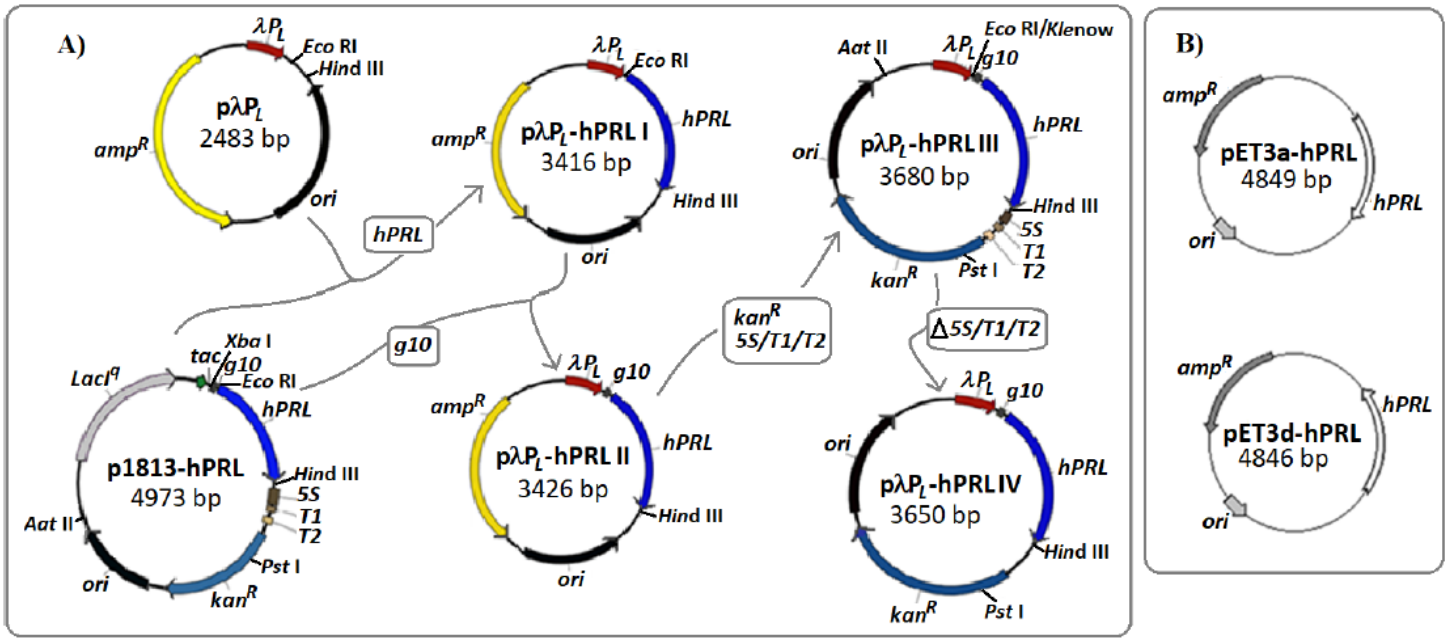


Figure 1

Schematic representation of the vectors. A) p1813-hPRL vector and pλPL-hPRL vector series construction, showing the location of all elements; B) Schematic representation of the vectors pET3a-hPRL and pET3d-hPRL. Abbreviations: hPRL, Prolactin hormone cDNA sequence; ampR, ampicillin resistance gene; kanR, kanamycin resistance gene; ori, origin of replication; λPL, lambda promotor PL; tac, tac promotor; g10, translator enhancer; 5S/T1/T2, transcription terminators; Laclq, lac repressor protein DNA sequence; bp, base pair.

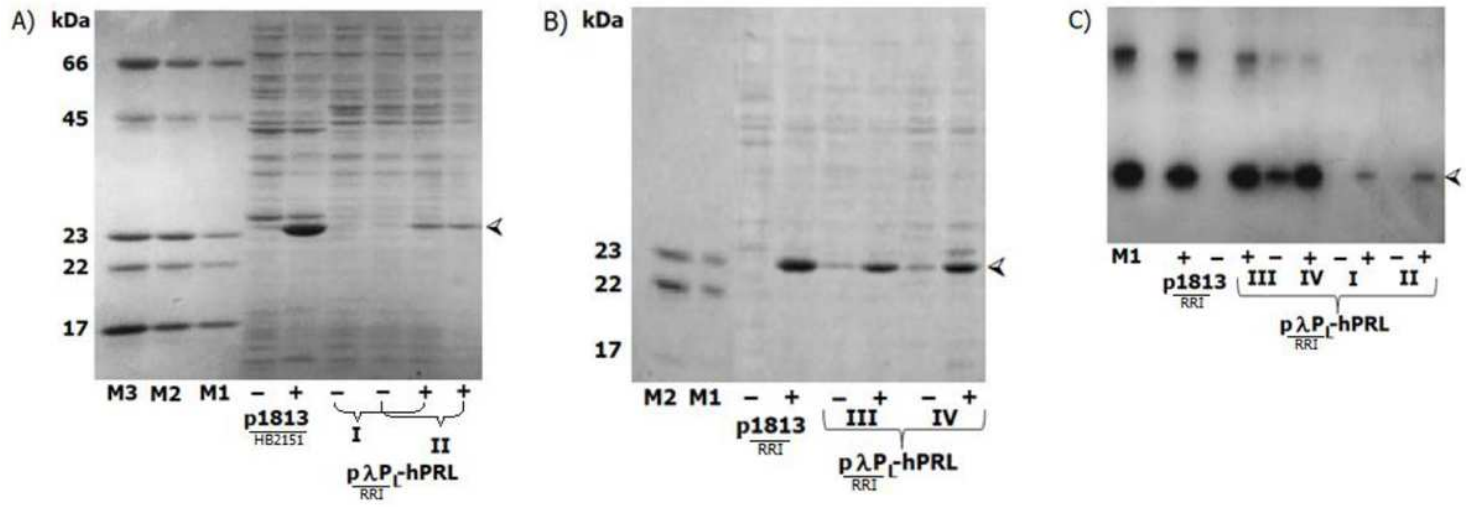
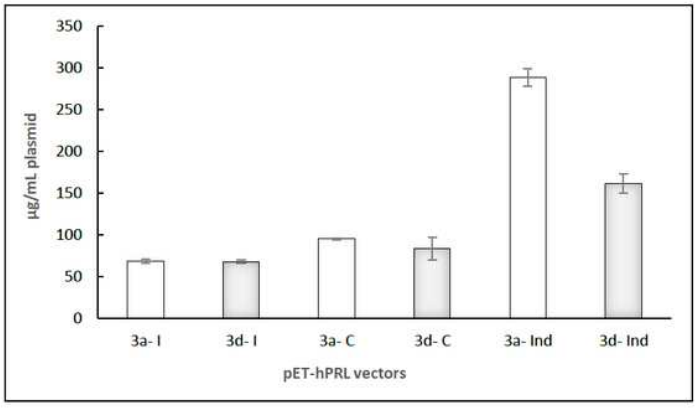


Figure 2

SDS-PAGE (A and B) and Western blot (C) analysis of rec-hPRL expression by p1813-hPRL and pλPL-hPRL vector series, in either HB2151 or RRI E. coli strains. Lanes M3, M2 and M1 contained 3, 2 and 1 µg of protein markers, respectively. The 23 kDa protein is pituitary PRL that was used as standard. Induced (+) and non-induced (-) conditions are shown.

A)



B)

Conditions	Vectors-hPRL	µg/mL±SE	$A_{260/280}$ *
Initial (I)	pET3a	68.3 ± 2.4	2.03
	pET3d	67.8 ± 1.8	2.00
Control (C)	pET3a	95.3 ± 1.1	1.96
	pET3d	83.5 ± 13.7	1.99
Induced (Ind)	pET3a	268.0 ± 10.1	1.94
	pET3d	161.4 ± 11.3	1.93

*ratio of 1.8 - 2.0 is generally accepted as pure for DNA

Figure 3

Quantifications of plasmid DNA from pET3a-hPRL and pET3d-hPRL, obtained in cultured E. coli BL21(DE3) strain. A) graph and B) table with plasmid DNA levels: (I), initial cultures when reaching 1.6 A600; (C), non-induced control cultures after 9 h of incubation; (Ind), IPTG-induced cultures after 9 h of incubation. (n=3).