

The impact of polyadenylation signals on plasmid nuclease-resistance and transgene expression

Adriano R. Azzoni
Sofia C. Ribeiro
Gabriel A. Monteiro
Duarte M.F. Prazeres*

IBB – Institute for Biotechnology and Bioengineering, Centre for Biological and Chemical Engineering, Instituto Superior Técnico, 1049-001 Lisboa, Portugal

*Correspondence to:
Duarte M.F. Prazeres, IBB – Institute for Biotechnology and Bioengineering, Centre for Biological and Chemical Engineering, Instituto Superior Técnico, 1049-001 Lisboa, Portugal.
E-mail: miguelprazer@ist.utl.pt

Abstract

Background Efficient delivery and expression of plasmids (pDNA) is a major concern in gene therapy and DNA vaccination using non-viral vectors. Besides the use of adjuvants, the pDNA vector itself can be designed to maximize survival in nuclease-rich environments. Homopurine-rich tracts in polyadenylation sequences have been previously shown to be especially important in pDNA resistance.

Methodology The effect of modifications in the poly A sequence of a model pDNA vector (pVAX1GFP) on nuclease resistance and transgene expression was investigated. Four poly A sequences were studied: bovine growth hormone (BGH), mutant BGH, SV40 and a synthetic poly A. Plasmid resistance (half-life) was assessed through *in vitro* incubations with mammalian nucleases. The impact in transgene expression was studied by quantifying pDNA, mRNA, and GFP expression in CHO, hybridoma and HeLa cells.

Results and conclusions *In vitro* and cell culture studies indicate that plasmids containing the SV40 and the synthetic poly A sequences present significant improvements in nuclease resistance (up to two-fold increase in half-life). However, RT-PCR analysis demonstrated that significant reduction in mRNA steady-state levels were responsible for a decrease in transgene expression and detected transfection level of CHO and hybridoma cells when using the more resistant plasmids. Interestingly, transfection of HeLa cells demonstrated that both poly A efficiency and plasmid resistance interfere significantly in transgene expression. The results strongly suggest that the choice of the poly A is important, not only for mRNA maturation/stability, but also for pDNA resistance, and should thus be taken into consideration in the design and evaluation of pDNA vectors. Copyright © 2007 John Wiley & Sons, Ltd.

Keywords plasmid DNA vectors; polyadenylation signal; nuclease resistance; reporter protein expression; mRNA maturation; DNA vaccines; gene therapy

Introduction

Inefficient nuclear delivery of plasmid DNA (pDNA) remains a major bottleneck in DNA vaccination and non-viral gene therapy [1,2]. Nuclease degradation of pDNA after administration and during trafficking to the cell nucleus is one of the main causes of this inefficiency [2,3]. Experimental evidence shows that the majority of pDNA administered in animal models is degraded by a population of exo/endonucleases that are constitutively present in the plasma and cytosol of the receptor cells. These pharmacokinetic experiments in mice and rat, performed with naked or complexed



Received: 29 November 2006
Revised: 30 January 2007
Accepted: 19 February 2007

pDNA (with proteins or cationic lipids), have shown that intravenously administered supercoiled (SC) pDNA is converted to the open circular (OC) and linear (L) forms within a few minutes [3,4]. Clearance of pDNA taken up by tissues is slower, occurring within a few hours [4]. The most used approach to improve pDNA uptake involves the protection of pDNA from the surrounding nuclease-rich cell/tissue environment with adjuvants such as cationic lipids or polymers [2,5,6]. In these cases, even if the number of pDNA copies that enters the cells is high (up to hundreds of thousand copies [7]), a large majority is degraded in endo/lysosomal compartments, with only a few copies finding their way into the cell nucleus. Finally, once the pDNA-adjuvant complex escapes from endo/lysosomal entrapment, it must be dismantled in order to become transcriptionally active. Following this de-aggregation step, pDNA will still remain susceptible to nuclease attack.

Although the fate of the pDNA molecules that reach the cell nucleus is not yet completely understood [8], it is clear that successful expression requires plasmids to remain intact in the cell nucleus for a prolonged period of time. Finding alternatives to increase pDNA resistance during trafficking and after nuclear uptake with a minimum of disturbance to the cell metabolism and transgene expression is therefore an important challenge.

An alternative to adjuvant protection, which has been recently reported [9], relies on the construction of pDNA variants that are more resistant to nuclease action. We have demonstrated that pDNA molecules possessing base sequences that originate from secondary structures with single-stranded regions (hereafter named hot spots) are more susceptible to exo/endonucleases. In the case of the model pDNA vector used (pVAX1LacZ, Invitrogen), the hot spots were specifically located within the bovine growth hormone polyadenylation signal (BGH poly A) and within the pMB1 bacterial replication origin. Two pDNA variants were then created by replacing the original BGH poly A signal with the SV40 poly A or with a synthetic poly A signal. An *in vitro* pDNA degradation analysis using eukaryotic cell lysates and mice plasma demonstrated that the half-life of the SC isoforms of the new vectors were considerably higher when compared to the original pDNA [9].

The polyadenylation process is an intricate event by which the precursor RNA (pre-mRNA) is endonucleolytically cleaved 11 to 25 nucleotides downstream of the AAUAAA sequence [10]. As important players in this process, the cleavage and polyadenylation specific factor (CPSF) and the cleavage stimulation factor (CST) bind to the AAUAAA sequence and to the nearby GU/U-rich tract, respectively. These two proteins provide a platform for the assembly of the polyadenylation complex, that include the poly A polymerase and poly(A)-binding protein (PABP). This polyadenylation complex will then be responsible for the 3'-cut and addition of the poly A tail (200 to 250 adenine residues) [11,12].

The rational design of the plasmid vectors used in gene therapy (e.g. promoters, enhancers, CpG motifs,

gene switches) is pinpointed as a key factor to enhance specificity and modulate the therapeutic transgene expression [13,14]. Here we present evidence that indicate that plasmid poly A sequences also have a strong effect on transgene expression, not only due to the impact at the post-transcriptional level, but also by modifying pDNA resistance to nucleases. A main challenge in the study of pDNA resistance is the correct determination of the number of pDNA molecules inside the transfected cells. The most popular approach used for this purpose is fluorescence microscopy. However, despite the highly valuable information obtained from this technique, it lacks analytical precision or makes use of pDNA labelled with fluorophores. These fluorescent dyes may be toxic to the cells, modify pDNA conformation, or ultimately change the fate of the pDNA molecules under study [8]. For this purpose, we used a quantitative real-time polymerase chain reaction (PCR) method to quantify pDNA inside transfected cells at different times post-transfection. The effect of modifications in poly A was also assessed through the determination of mRNA transcripts using reverse-transcription (RT)-PCR and the levels of transgene expression (enhanced green fluorescent protein (eGFP)) by flow cytometry.

Materials and methods

Plasmid construction

The commercial plasmid vector pVAX1LacZ (6050 bp, Invitrogen), designed for use in the development of DNA vaccines, and the nuclease resistant variants pVAX1-SV40 and pVAX1-Synt, constructed as described previously [9], were used as starting model systems. In this work, the *lacZ* reporter gene was replaced in the three plasmids by the *eGFP* gene obtained from plasmid pFM20 [15]. The *eGFP* gene was cloned between the XbaI, EcoRI restriction sites. The new plasmids were then renamed pVAX1GFP-BGH (3697 bp), pVAX1GFP-SV40 (3690 bp), and pVAX1GFP-Synt (3488 bp). The plasmids were used to transform *E. coli* DH5 α and purified according to the High Pure plasmid isolation kit protocol (Roche, Germany). Automated DNA sequencing was performed on these constructions by Stab Vida Inc., Portugal.

The pVAX1GFP-BGH vector contains the human cytomegalovirus (CMV) immediate-early promoter, the BGH polyadenylation sequence, a kanamycin resistance gene for selection in *E. coli*, a pMB1 origin (pUC-derived), a multiple cloning site and a T7 promoter/priming site. In the variants pVAX1GFP-SV40 and pVAX1GFP-Synt, the BGH poly A sequence (225 bp) is replaced by two known poly A sequences: SV40 late polyadenylation signal (222 bp), and a synthetic polyadenylation sequence (49 bp), respectively. The synthetic poly A is based on the highly efficient poly A signal of the rabbit beta-globin gene [16].

Construction of plasmid pVAX1GFP-BGH-M1 with BGH poly A mutation

Plasmid pVAX1GFP-BGH-M1 was constructed by deleting a 61 bp sequence located downstream of the AATAAA sequence of BGH poly A. As represented in Figure 1, a restriction site for the enzyme BpiI (GAAGAC) was generated by *in vitro* site-directed mutagenesis 25 nucleotides downstream of the AATAAA signal. The mutagenesis was performed with the PfuTurbo DNA polymerase according to the QuickChange site-directed mutagenesis kit (Stratagene, UK) using the forward primer 5'-AAT TGC ATC GCA TTG TCT GAA GAC GGT GTC ATT CTA TTC TGG G-3' and the reverse primer 5'-CAC CCC CCA GAA TAG AAT GAC ACG TCT TCA GAC AAT GCG ATG C-3'. The mutagenesis created a second BpiI restriction enzyme site downstream of the AATAAA signal, enabling the digestion and deletion of the 61 bp sequence between the BpiI sites (Figure 1). The sequence of the modified, 164 bp BGH polyadenylation signal in the new pVAX1GFP-BGH-M1 plasmid was then confirmed by automated DNA sequencing.

Secondary structure mapping of the new plasmids

Secondary structures of the plasmids containing the *eGFP* gene were mapped by performing a sequential double digestion with the single-stranded specific nuclease S1 from *Aspergillus oryzae* (Roche, Germany) and restriction enzymes (Promega, USA) with a unique restriction site in the plasmids. After nuclease S1 digestion (50 min, 37°C, 0.93 U/μg plasmid), the plasmid samples (4 μg) were incubated (3 h, 37°C) with one of the following restriction enzymes: XbaI, EcoRI or NheI. In each case, a control was prepared by adding the restriction enzyme buffer to the S1-digested plasmid in order to assure that restriction fragments were not the result of a persistent nuclease S1 attack. The digestions were stopped by adding an EDTA solution up to a

final concentration of 30 mM on ice. The samples were then analyzed by agarose gel electrophoresis as described below. The approximate location of the hot spot sequence can be found once the molecular weight (MW) of the restriction fragments obtained is determined and the restriction site of the enzyme is known.

In vitro nuclease resistance experiments

The resistance of the plasmids pVAX1GFP-BGH, pVAX1GFP-SV40, pVAX1GFP-Synt, and pVAX1GFP-BGH-M1 was tested through *in vitro* incubations with Chinese hamster ovary (CHO) cell intracellular nucleases at pH 7.0 and 4.7, and with mice plasma nucleases as described previously [9]. Samples were examined by 1% agarose gel electrophoresis run for 4 h at 60 V and stained for 30 min in a 0.5 μg/mL ethidium bromide solution. In order to assess the variation of the SC and OC plasmid conformations during incubations, the intensities of the corresponding bands in the agarose gels were determined using densitometry software (TotalLab v2.0, Phoretix, UK). The percentage of the SC isoform was calculated from at least three independent experiments by normalizing band intensities to incubation time zero. The resistance of the plasmids stored at 4°C in ultra-pure (and sterile) water was also evaluated as previously described [9].

Culture and transfection of mammalian cells

CHO, mouse hybridoma (5367G, αCD8), and human epitheloid carcinoma (HeLa) cells were grown in F-12 (Ham's) nutrient mixture (Gibco, UK) containing 10% (v/v) foetal bovine serum (Gibco, UK). All cell cultures were performed in 75 cm² culture flasks and incubated in 5% CO₂ humidified environment at 37°C. Following

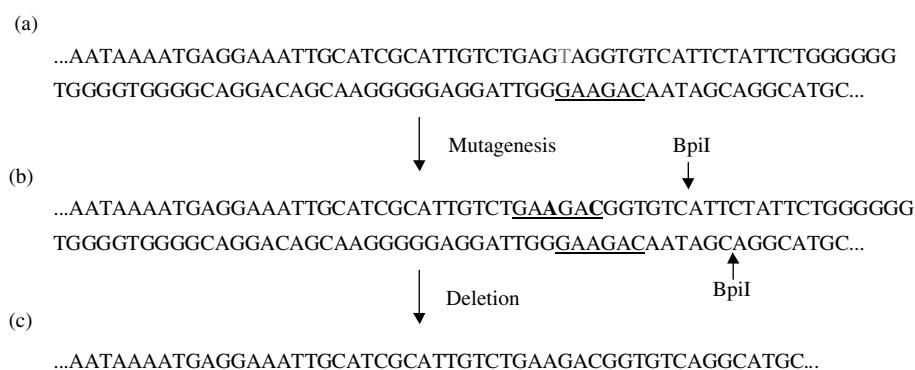


Figure 1. Mutation and partial deletion of the 3' flanking sequence of the BGH polyadenylation signal, generating the plasmid pVAX1GFP-BGH-M1: (a) original BGH poly A sequence. The nucleotide in grey (T) was deleted during mutagenesis; (b) mutated BGH poly A with new BpiI restriction enzyme site (sites are underlined and introduced nucleotides appear in bold); and (c) final poly A sequence of the pVAX1GFP-BGH-M1. Arrows indicate the exact site of the restriction enzyme cut

growth up to confluence, cells were trypsinized and seeded in 24-well culture plates (5×10^4 cells per well). The cells were incubated for 48 h and then transfected with the different plasmids. Transfection was carried out using the Lipofectamine 2000TM reagent (Invitrogen, USA), according to the manufacturer's instructions (1 μ g pDNA plus 1.5 μ L lipid in 100 μ L of medium per well). The medium containing the transfection solution remained in the cell-containing wells for 12 h and was then replaced by regular medium.

Real-time PCR for determination of plasmid content

Quantitative real-time PCR for plasmid determination was carried out as recently described [15]. Firstly, transfections of the mammalian cells were performed as described above. As a negative control, medium containing pDNA but no transfection lipid was added to specific control wells. After appropriate periods, transfected and non-transfected cells were washed four times with 1 mL phosphate-buffered saline (PBS) to remove non-specifically bound pDNA (as described by Zabner *et al.* [7]), trypsinized from wells, and centrifuged at 160 g for 8 min. Cell pellets were then washed with PBS, resuspended in 0.6 mL of the same saline solution, and counted on a Neubauer chamber (counting was performed in duplicate). Finally, cells were centrifuged again, the supernatant solution was removed, and pellets were stored at -80°C for PCR analysis. Prior to analysis, cell pellets were ice-thawed and resuspended in PCR-grade deionized water to a cellular concentration of 6250 cells/ μ L. Volumes of 2 μ L of such suspensions were used in each PCR. Negative controls containing the same amount of non-transfected cells, exposed to the pDNA but not to LipofectamineTM (control cells), were always included. Determination of pDNA concentration in transfected cells was done in triplicate. Standard curves were constructed by adding different amounts of purified pDNA (pVAX1GFP-BGH) to the control cells (2 μ L per reaction, with 6250 cells/ μ L). Negative controls using PCR-grade water instead of control cells were also always run to detect undesired contamination.

Quantitative real-time PCR for determination of pDNA was performed by amplification of a 108 bp sequence from the *eGFP* gene using the 5'-TCG AGC TGG ACG GCG ACG TAA A-3' forward primer, and 5'-TGC CGG TGG TGC AGA TGA AC-3' reverse primer. PCRs were carried out in a Roche LightCyclerTM detection system using the FastStart DNA Master SYBR Green I kitTM (Roche, Germany) as recommended by the manufacturer. Each 20 μ L of final reaction volume contained 2.0 μ L of the $10\times$ SYBR Green mixture, 0.4 μ L of each primer (0.4 μ M final concentration), 1.6 μ L of MgCl_2 solution (3.0 mM final concentration), 13.6 μ L PCR-grade water, and 2.0 μ L of sample. Reactions were incubated at 95°C for 10 min to activate FastStart DNA polymerase, followed by 30 cycles of 10 s at 95°C , 5 s at 55°C , and 7 s at 72°C . Following

the final cycle, reaction mixtures were kept at 70°C for 30 s, and heat-denatured over a temperature gradient at 0.1°C/s from 70 to 95°C .

Quantitative RT-PCR

Transfection, washing and collection of mammalian cell samples were performed as described above. Cell samples for RNA analysis were immediately stored at -80°C . Total cellular RNA was isolated from these samples using the High Pure RNA isolation kit (Roche, Germany) according to the manufacturer's instructions. The total RNA content of each sample was quantified by absorption at 260 nm. The reverse transcription reactions were normalized to contain equivalent amounts (300 ng) of total RNA. Reverse transcription was performed using the 1st Strand cDNA synthesis kit for RT-PCR (AMV) (Roche, Germany). The cDNA synthesis was performed using the Oligo-p(dT)₁₅ primer, according to the manufacturer's instructions. Volumes of 2 μ L of each reverse-transcription reaction were then used in each quantitative real-time PCR reaction, performed as described above for plasmid DNA determination. The RT-PCR determination for GFP mRNA content in transfected cells was done in triplicate. Standard curves were constructed by adding different amounts of purified pDNA (pVAX1GFP-BGH) to the cDNA of control cells (non-transfected cells). Negative controls using PCR-grade water instead of control cells were also run to detect any undesired contamination.

Detection of GFP expression by flow cytometry

Cultivation and transfection of the mammalian cells with the plasmids were performed as described above. After selected periods post-transfection, cells contained in 24-well plates were trypsinized and centrifuged at 160 g for 8 min. Cell pellets were then washed with PBS and resuspended in 1 mL of the same solution. Transfected and non-transfected cells were analyzed in a FACScan Scalibur (Becton-Dickinson, USA), which recorded forward scatter (FSC), side scatter (SSC), and green fluorescence (FL1). From these data, cells were first isolated from debris by their FSC versus SSC characteristics, and transfection efficiency was determined by subtracting those cells displaying background FL1 (non-transfected cells) from the total cell population. Average green fluorescent intensities corresponding to the expression level of GFP were calculated using the CellQuest Pro software (Becton-Dickinson, USA). In this work we define 'transfection level' as the percentage of cells expressing green fluorescence above the threshold level defined by the natural autofluorescence of the control non-transfected cells.

Results

Plasmid DNA construction and secondary structure mapping

The commercial plasmid vector pVAX1LacZ and the poly A modified variants pVAX1LacZ-SV40 and pVAX1LacZ-Synt [9] were used as starting model systems. To facilitate detection of the transgene expression, the *lacZ* reporter gene was replaced by the *eGFP* gene in these three plasmids. The vectors were then renamed pVAX1GFP-BGH (3697 bp), pVAX1GFP-SV40 (3690 bp), and pVAX1GFP-Synt (3488 bp). The plasmid yield and isoform distribution obtained with the three new constructs after amplification in *E. coli* and purification were similar to those obtained with the LacZ plasmids (results not shown).

The secondary structure of the new plasmids was mapped to check if any new hot spot had been introduced after the replacement of the reporter gene. Table 1 shows the results of the sequential double digestion with the single-stranded specific nuclease S1 and restriction enzymes with a unique restriction site in the plasmid pVAX1GFP-BGH. An analysis of the molecular mass of the fragments resulting from sequential double digestion indicated the presence of only two hot spots in the new plasmids, in the same locations found previously for pVAX1LacZ (average positions 1621 bp and 3051 bp, corresponding to the BGH poly A sequence and pMB1 origin of replication, respectively) [9]. These results show that the replacement of the *lacZ* gene by the *eGFP* gene did not introduce new hot spots in the plasmids.

Plasmid resistance to nucleases

The resistance of each plasmid towards nucleases present in mice plasma, and CHO cell lysates at pH 4.7 and 7.0, was assessed by *in vitro* incubations at 37°C. Tests were also done using the original pVAX1LacZ as a control plasmid. The time course of nuclease degradation of each plasmid in terms of loss of the SC isoform was quantified by scanning the corresponding bands in the agarose gel (Figure 2).

Plasmids pVAX1GFP-SV40 and pVAX1GFP-Synt were significantly more resistant than the pVAX1GFP-BGH (and pVAX1LacZ) containing the original BGH poly A

sequence. This was more evident when using mice plasma and CHO cell lysate nucleases at pH 4.7, representative of the endo/lysosomal nucleases. In mice plasma, the half-lives of pVAX1GFP-SV40 and pVAX1GFP-Synt were approximately 2-fold higher when compared with pVAX1GFP-BGH. The results also indicate that the replacement of the *lacZ* gene by the *eGFP* gene slightly increased the overall plasmid resistance to nucleases, probably because plasmids containing the *eGFP* gene (3697 bp) are significantly smaller than those containing the *lacZ* gene (6050 bp). Overall, plasmids containing the *eGFP* gene presented a nuclease degradation pattern similar to that presented by Ribeiro *et al.* for the 'LacZ' plasmids [9].

The stability of the plasmids during storage at 4°C in MilliQ water (pH around 7.0) was also evaluated (Figure 3). The results found indicate impressive differences in the isoform distribution of the four tested plasmids after 5 months of storage. Almost all the SC isoforms of plasmids pVAX1LacZ and pVAX1GFP-BGH were converted into the OC and L isoforms, while most of the SC isoforms of plasmids pVAX1GFP-SV40 and pVAX1GFP-Synt remained intact, indicating a significant improvement in storage stability.

Plasmid DNA decay in transfected CHO cells

The resistance of the poly A modified plasmids in cultured cells was evaluated by detecting the decay of the different plasmids after transfection of CHO cells, using real-time PCR. Firstly, a standard curve for the determination of the number of plasmid copies was generated by spiking non-transfected CHO cells with increasing amounts of pDNA. A linear working range (from 5 pg to 2.5 ng of pDNA per PCR reaction) was obtained, enabling the quantification of a minimum of 350 copies/cell (LOQ) and detection of a minimum of 140 copies/cell (LOD). Figure 4 presents the copy numbers of the different plasmids inside CHO cells at different times post-transfection. In agreement with the *in vitro* nuclease resistance assays, plasmids pVAX1GFP-SV40 and pVAX1GFP-Synt presented higher resistance to degradation when compared with the original plasmid. The results also indicate that a high number of pDNA copies (tens of thousands) were internalized by each cell during transfection. This is in accordance with results obtained by other authors using lipid-mediated transfection [7,17], which also showed that the large majority of pDNA copies remain entrapped in endosomal vesicles and never reach the cell nucleus. The observed rate of pDNA decay was not linear, decreasing for longer post-transfection times. An average value of 1300 copies of the plasmid pVAX1GFP-BGH was detected 4 days post-transfection. The calculated half-life of this pDNA (after 24 h post-transfection) was 19.5 h, a value coherent with those found in literature [1,7]. The half-lives of the SV40 and Synt plasmids were 26.4 and 29.3 h, respectively. It is important to note that since the cells continue to divide

Table 1. Mapping of secondary structures (hot spots) in pVAX1GFP-BGH by double digestion with nuclease S1/restriction enzyme

Restriction enzyme	Secondary structure location (bp)	
XbaI	*	3066 ± 35
EcoRI	1629 ± 22	3092 ± 29
NheI	1613 ± 15	2996 ± 48
Average location	1621 ± 30	3051 ± 50
Location within the plasmid	BGH polyA	pMB1 origin

*Not possible to quantify fragment size.

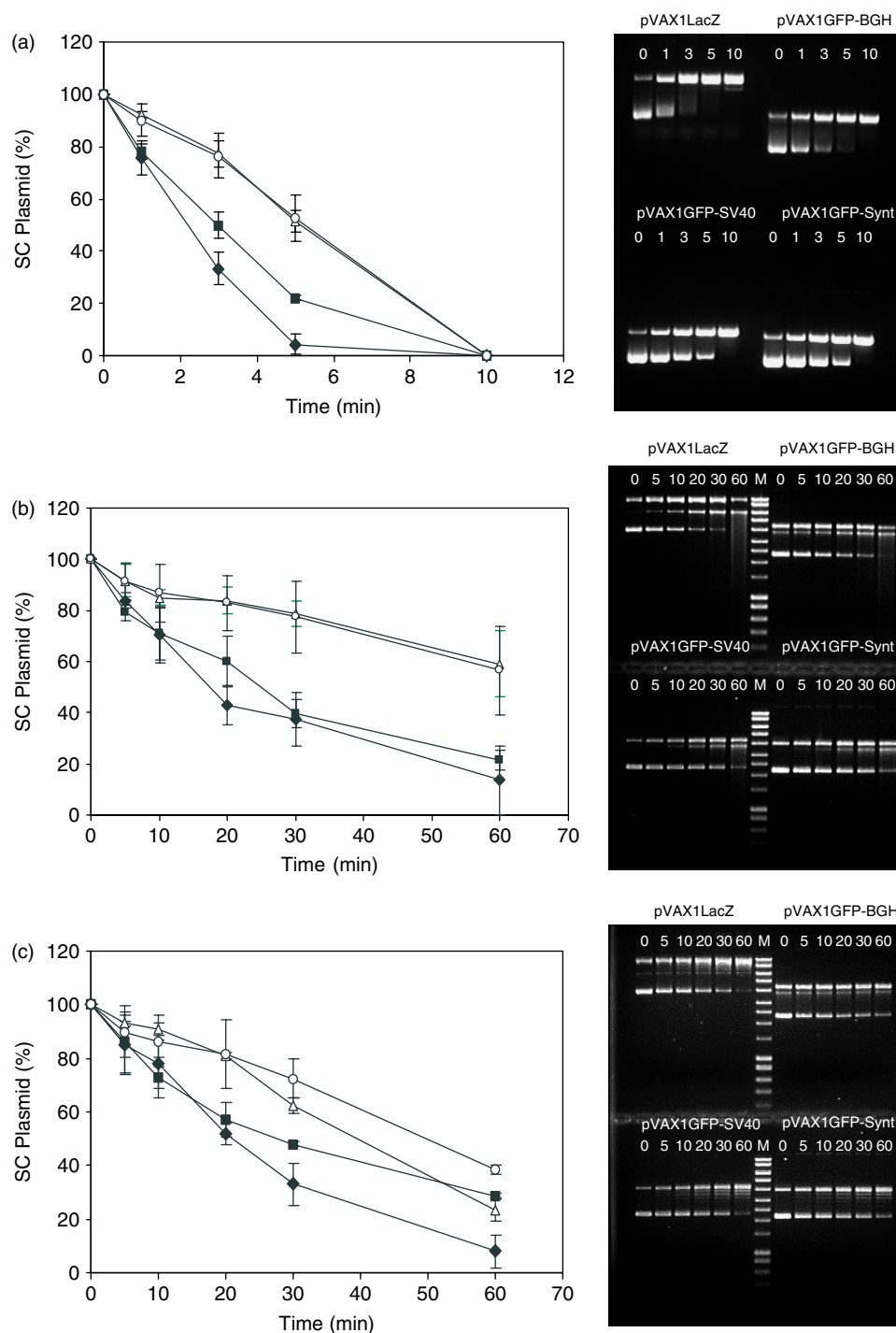


Figure 2. Time-course evolution of SC plasmid isoforms during incubation with nucleases at 37°C, and agarose gel electrophoresis presenting the corresponding nuclease degradation patterns: (a) mice plasma, (b) CHO cell lysates at pH 4.7, and (c) CHO cell lysates at pH 7.0. Symbols: pVAX1LacZ (◆), pVAX1GFP-BGH (■), pVAX1GFP-SV40 (△) and pVAX1GFP-Synt (○). Error bars indicate standard deviation between three independent experiments. Numbers over each gel lane correspond to incubation time in minutes. Masses of 0.8 μ g and 0.4 μ g of each pDNA were loaded in the agarose gel lanes for mice plasma and CHO lysates, respectively

after the transfection procedure (although at a lower rate after reaching confluence), part of the reduction in plasmid copies per cell should be attributed to pDNA partition between daughter cells and not exclusively to degradation. Also important is the fact that the PCR quantification method determines the total amount of pDNA inside the cells, with no distinction between those

entrapped within endosome vesicles (the majority) or those that reached the cell nucleus. Nevertheless, it is reasonable to assume that most of the pDNA decay measured accounts for degradation of plasmids entrapped within the vesicles. Plasmids that reach the cell nucleus may have a different fate than immediate degradation, remaining in the cell nucleus for a longer period [8].

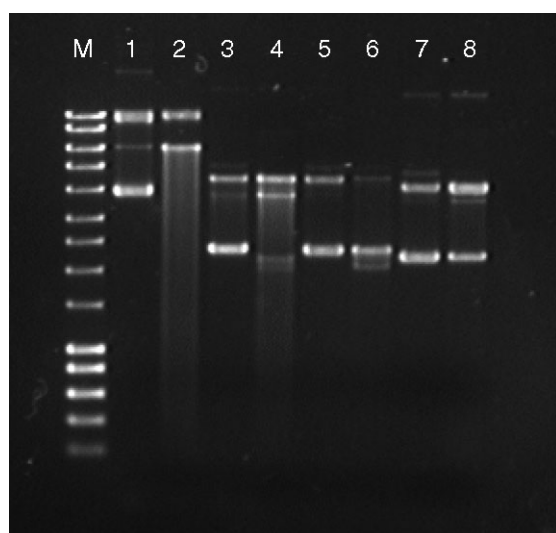


Figure 3. Agarose gel electrophoresis presenting the plasmid stability during storage in ultra-pure water, pH 7, at 4°C. M, DNA ladder; lanes 1, 3, 5 and 7, control plasmids before storage (pVAX1LacZ, pVAX1GFP-BGH, pVAX1GFP-SV40 and pVAX1GFP-Synt, respectively); lanes 2, 4, 6 and 8, plasmids stored for 5 months (pVAX1LacZ, pVAX1GFP-BGH, pVAX1GFP-SV40 and pVAX1GFP-Synt, respectively)

Transgene expression in mammalian cells

The efficiency of transfection and transgene expression of each plasmid vector was analyzed through the quantification of the number of cells presenting *eGFP* fluorescence after transfection experiments. Cells expressing fluorescence were quantified by flow cytometry using non-transfected cells as negative controls. Figure 5 shows the percentage of transfected cells showing green fluorescence at different times post-transfection in a typical transfection experiment, whereas Table 2 presents the values obtained 24 h after transfection in three independent experiments. Despite the fact that the transfection level varied between independent experiments, transfection with pVAX1GFP-BGH always resulted in a higher number of detected fluorescent cells when compared with transfections using the pVAX1GFP-SV40 and pVAX1GFP-Synt variants. The differences amongst transfections are explained by a sum of factors including small differences in plasmid batches, cationic-lipid/plasmid complex preparation, and cell confluence during transfection. Transfection experiments conducted in CHO cells using electroporation led to results (not shown) similar to those obtained by lipofection. Thus, a direct correlation between the efficiency of transgene expression in CHO cells and the resistance of the plasmids to nuclease degradation (*in vitro* and in cultured cells) was not observed.

After determining that the plasmid containing the BGH poly A signal led to higher *eGFP* expression, despite its lower nuclease resistance *in vitro* and in cultured cells, time-course transfection experiments were conducted to quantify *eGFP* mRNA by RT-PCR (Figure 6). Absolute quantification of the copy number of the target mRNA

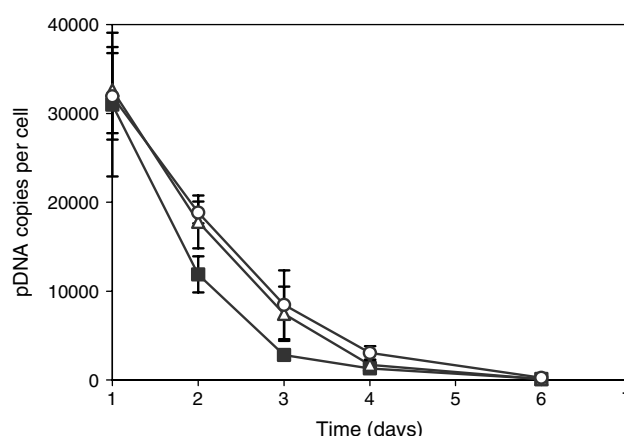


Figure 4. Kinetic of pDNA decay inside CHO cells quantified by real-time PCR after transfection with the pVAX1GFP-BGH (■), pVAX1GFP-SV40 (△), and pVAX1GFP-Synt (○) using Lipofectamine™. Error bars indicate standard deviation between triplicates

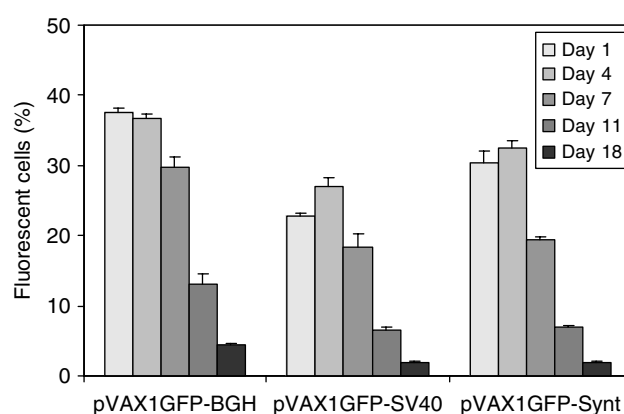


Figure 5. Time-course of transfected CHO cells presenting green fluorescence. The cells were transfected with the different pDNA vectors using Lipofectamine™, and the expression of the enhanced green fluorescent protein (*eGFP*) was detected by flow cytometry. Error bars indicate standard deviation between duplicates

Table 2. Percentage of CHO cells presenting fluorescence 24 h after transfection with different plasmids in three distinct experiments

Plasmid	Exp. 1	Exp. 2	Exp. 3	Average	SD*
pVAX1GFP-BGH	28.8	47.1	37.5	35.5	12.8
pVAX1GFP-SV40	13.4	37.4	22.8	24.5	12.1
pVAX1GFP-Synt	17.9	36.9	30.3	28.4	9.6

*SD: standard deviation.

transcript was done by relating the PCR signal to a standard curve with detection (LOD) and quantification (LOQ) limits of 1850 and 18 000 copies per ng of total RNA, respectively. The mRNA results indicated that the number of mRNA copies in cells transfected with pVAX1GFP-BGH was significantly higher when compared with the number of copies obtained with cells transfected with the pVAX1GFP-SV40 and pVAX1GFP-Synt plasmid

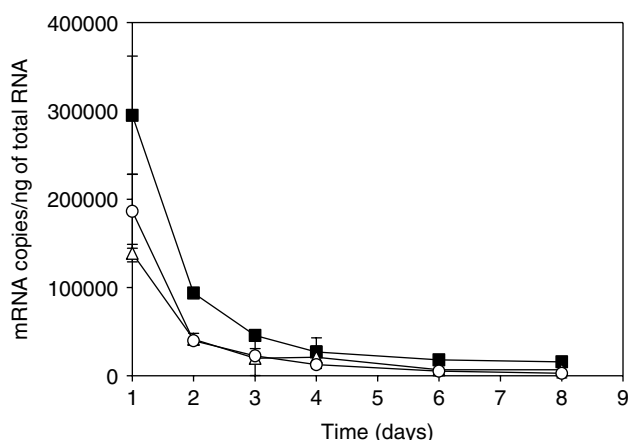


Figure 6. Time-course of eGFP mRNA level quantified in CHO cells by RT-PCR after transfection with the pVAX1GFP-BGH (■), pVAX1GFP-SV40 (△), and pVAX1GFP-Synt (○) using Lipofectamine[™]. Error bars indicate standard deviation between triplicates (days 1 and 2 post-transfection) or deviation between duplicates (days 3, 4, 6 and 8)

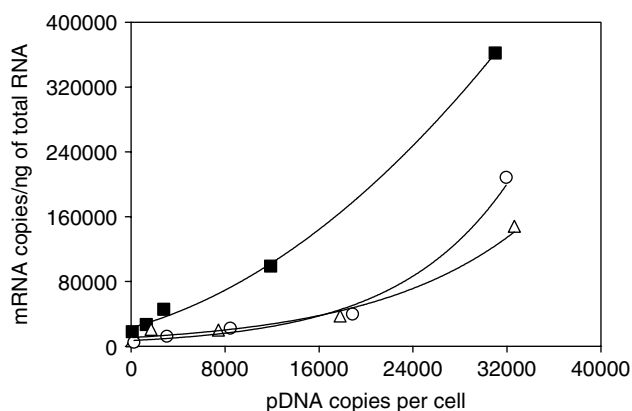


Figure 7. Correlation between plasmid copies per cell versus eGFP mRNA level in transfected CHO cells using pVAX1GFP-BGH (■), pVAX1GFP-SV40 (△), and pVAX1GFP-Synt (○). The chart was plotted using the average data presented in Figures 4 and 6

variants. Figure 7 plots the total pDNA per CHO cell versus the eGFP mRNA content. If for the sake of argument we assume that the number of pDNA molecules that reaches the nucleus is proportional to the total number of plasmids in the cells, the results indicate that the BGH poly A sequence has an efficiency up to 4-fold higher than the SV40 and Synt poly As, a value obtained by comparing the ratios mRNA/pDNA at 48 h post-transfection.

Clearly, the differences in transfection levels found between the vectors are not only the result of the different numbers of pDNA molecules that reach the nucleus, but should also be attributed to differences in the steady-state levels of mRNA. The fact that the plasmid containing the BGH poly A is more efficient at the post-transcriptional level may result in cells with higher fluorescence levels (above the threshold of the flow cytometer), even with a smaller number of pDNA molecules in the cell nucleus. In fact, CHO cells transfected with pVAX1GFP-BGH presented almost twice the fluorescence intensity

than cells transfected with the pVAX1GFP-SV40 and pVAX1GFP-Synt (Figure 8). A good correlation (r^2 above 0.95) was found when plotting the average intensity of cell fluorescence as a function of the mRNA level.

Transfection experiments using mouse hybridoma and HeLa cells were also performed using the different pDNA vectors. Despite the much lower transfection level, experiments with hybridoma cells led to results similar to those obtained using CHO cells (Figure 8). However, interesting differences were found when transfecting HeLa cells. As obtained with CHO and hybridoma cells, transfection with pVAX1GFP-BGH resulted in higher values of eGFP mRNA and average intensity of cellular fluorescence (Figure 8). On the other hand, the transfection levels obtained with pVAX1GFP-SV40 and pVAX1GFP-Synt were approximately the same or even higher (for pVAX1GFP-Synt) than those found with pVAX1GFP-BGH. The highest number of plasmid copies per HeLa cell (around 15 090) was also obtained with pVAX1GFP-Synt. These results indicate that not only the efficiency of the poly A sequence at the post-transcriptional level but also pDNA resistance plays a major role in the detected transfection level of the studied plasmids.

Effect of polyadenylation efficiency on transgene expression

In order to demonstrate the effect of polyadenylation efficiency on the transfection level and transgene expression, a plasmid containing a mutant BGH poly A was constructed (Figure 1). The mutant (pVAX1GFP-BGH-M1) has a deletion of 61 bp downstream of the AATAAA signal. The deleted sequence included part of the 3' efficiency elements of the BGH poly A and also part of a putative hot spot sequence for nuclease action, as described previously [9]. No significant changes were detected in the *in vitro* resistance of pVAX1GFP-BGH-M1 to nuclease action. A secondary structure mapping with nuclease S1 indicated that the putative hot spot was still present, probably as a result of the high frequency of purine-rich zones in the BGH poly A sequence (results not shown) [9]. Figure 9 presents the average values of pDNA (copies/cell), mRNA (copies/ng total RNA), transfection level (%) and average intensity of cellular fluorescence, 48 h post-transfection of CHO and HeLa cells with the mutant pVAX1GFP-BGH-M1. A comparison of the data obtained for pVAX1GFP-BGH and pVAX1GFP-BGH-M1 shows that the deletion in the BGH poly A sequence led to: (i) a decrease in the transfection level (2 times), (ii) a decrease in the mean fluorescence (3 times), (iii) a similar number of plasmid copies per cell, and (iv) a decrease in eGFP mRNA steady-state level (8 times). Clearly, the mutation in the BGH poly A signal did not result in an increase in pDNA resistance to nucleases, but severely impaired transgene expression. So far, it is difficult to evaluate if the lower levels of mRNA obtained when using the Synt, SV40 and BGH-M1 poly As in comparison with

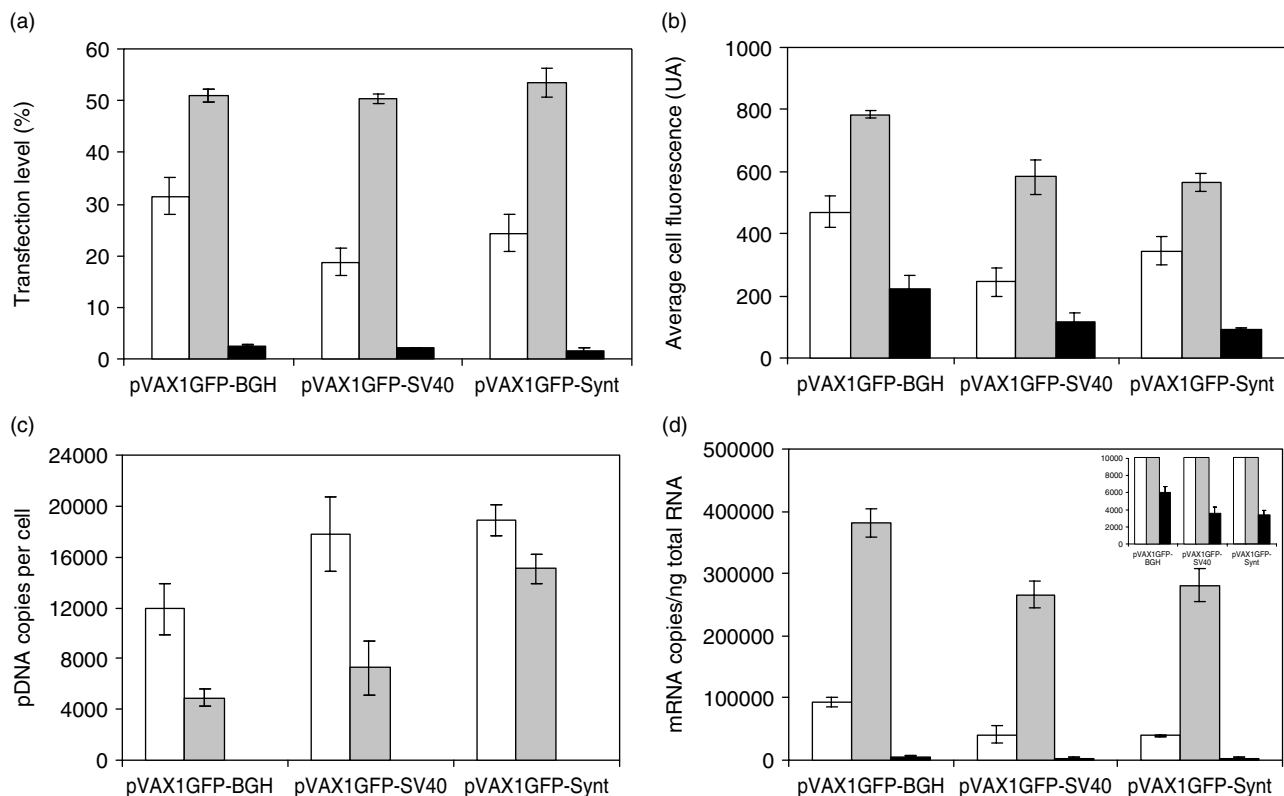


Figure 8. Analysis of transfected CHO (white bars), HeLa (grey bars) and hybridoma (black bars) cells using the pDNA vectors with different polyadenylation sequences. Cells were transfected with Lipofectamine[™] for 12 h (CHO cells) or 6 h (hybridoma and HeLa cells), collected after 48 h, and analyzed for transfection level (a), average cellular fluorescence (b), pDNA copies per cell (c), and eGFP mRNA levels (d) as described in the 'Materials and methods' section. Values of pDNA copies per hybridoma cell were below the detection limit of the method (<350 copies per cell). Insert in (d) was introduced for a better visualization of the hybridoma cells results. Error bars indicate standard deviation between duplicates of two independent transfection experiments, each one performed in duplicate

those obtained with BGH poly A, resulted from differences in the mRNA maturation process or stability (turnover), since both directly influence steady-state mRNA levels [18].

Discussion

In a previous report we have shown a direct correlation between the number of homopurine-rich sequences in the polyadenylation signals BGH, SV40 and Synt and the susceptibility of plasmids harbouring each sequence (pVAX1LacZ-BGH, pVAX1LacZ-SV40, pVAX1LacZ-Synt) to digestion with the single-stranded specific nuclease S1 [9]. The BGH poly A sequence has the highest frequency of purine-rich zones (six), followed by the SV40 (three) and Synt (none). The results presented in this work confirm the importance of these regions, which are prone to form single-stranded structures, in plasmid stability and resistance to mammalian nuclease attack. The mapping of secondary structures also indicated that the replacement of the plasmid reporter gene (*lacZ* by *eGFP*) in each of the three reference plasmids did not alter the presence of single-stranded secondary structures (at least not at detectable levels). The *in vitro* incubation of these new

eGFP plasmids with mammalian nucleases confirmed that the suppression of the hot spot in the poly A signal results in significant improvements in plasmid resistance (Figure 2). This effect was more evident for mice plasma and CHO endosome/lysosome nucleases. Furthermore, the plasmids containing the SV40 and the synthetic poly A presented remarkable stability upon storage, when compared to the plasmid with BGH poly A (Figure 3).

Although not so evident as in the *in vitro* assays, the analysis of pDNA decay in transfected CHO cells also indicates that the pVAX1GFP-SV40 and pVAX1GFP-Synt variants are more resistant to intracellular degradation (Figure 4). However, despite this increased resistance to nucleases, the pVAX1GFP-SV40 and pVAX1GFP-Synt plasmids presented significantly lower transfection efficiency, as quantified by the number of CHO cells expressing the reporter protein eGFP at levels detectable by the flow cytometric analysis (Figures 5 and 8). Since the plasmids differed only in their poly A signals, the role and function of these sequences at the post-transcriptional level should be responsible for the differences in transgene expression in CHO cells.

It is known that the BGH poly A pre-mRNA forms an extensive hairpin loop secondary structure at the 3'-untranslated region (UTR) [19]. Although less extensively, SV40 poly A also forms a malleable stem-loop

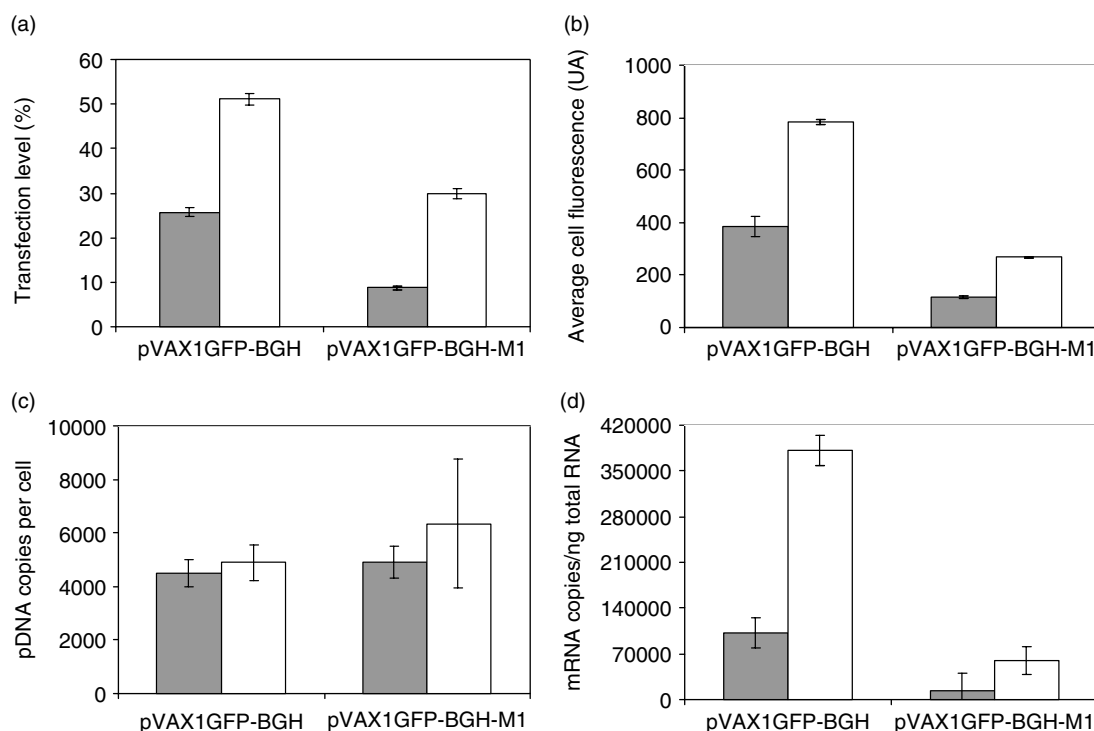


Figure 9. Comparison between transfections performed with plasmid pVAX1GFP-BGH and the poly A modified plasmid pVAX1GFP-BGH-M1 using CHO (grey bars) and HeLa (white bars) cells. Transfections were done as described in the 'Materials and methods' section using Lipofectamine™. Cells were collected 48 h post-transfection and analyzed for transfection level (a), average cellular fluorescence (b), pDNA copies per cell (c), and eGFP mRNA levels (d) as described in the 'Materials and methods' section. Error bars indicate standard deviation between duplicates of two independent transfection experiments, each one performed in duplicate

due to the presence of GU- and U-rich elements upstream and downstream of the AAUAAA signal sequence [10,11]. However, the much smaller Synt poly A signal has no GT- or T-rich sequences upstream of the AATAAA sequence, even though our *in silico* analysis indicates that a loop structure downstream of that signal does exist in solution (result not shown). All together, the longer and more malleable structure of the BGH poly A signal may result in a more efficient polyadenylation sequence.

In the case of HeLa cells, however, a different picture emerged when comparing plasmids pVAX1GFP-BGH and pVAX1GFP-Synt (Figure 8). In fact, and although levels of mRNA and average cell fluorescence were almost 40% higher with the BGH poly A signal, transfection efficiency was very similar to those obtained with plasmids pVAX1GFP-SV40 and pVAX1GFP-Synt. This may be directly attributed to the higher number of plasmid copies per cell detected, especially in the case of plasmid pVAX1GFP-Synt (15 090 vs. 4900), demonstrating that in some cases plasmid resistance is important. The more nuclease-resistant plasmid, presenting a longer half-life of the SC isoform, was probably able to reach the nucleus of a larger number of cells. The results recently reported by Remaut *et al.* [20] showing the much higher transfection efficiency of the SC isoform when compared to the OC and L isoforms corroborates this hypothesis. In the case of pVAX1GFP-BGH and the modified pVAX1GFP-BGH-M1, similar levels of pDNA copies were found in the cells. However, the low efficiency of the modified

poly A sequence led to a significant decrease in the 'apparent' transfection level. Taken together the results indicate that the detected transfection level results from a compromise between plasmid resistance (hence number of plasmids reaching the cell nucleus) and post-transcriptional efficiency of the poly A signal.

So far, few studies have demonstrated the importance of incorporating post-transcriptional regulatory elements in gene therapy vectors in order to improve gene expression via an increase in the level of mRNA and the efficiency of translation [21–24]. In those studies, the approach used involved the insertion of the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) into the 3'-UTRs of adeno-associated virus vectors. Despite the benefits in gene expression, the excessive size of the WPRE sequence proved to be a significant limitation [25]. To the best of our knowledge, this type of approach has not been applied to non-viral vectors.

In conclusion, the efficiency of a pDNA vector will ultimately result from its efficiency in overcoming physical and metabolic barriers during trafficking to the cell nucleus, and also from the efficiency of the regulatory sequences (e.g. promoter, poly A sequence, etc.) responsible for the transgene expression. As recently demonstrated for the influence of pDNA size in transfection efficiency [26], the pDNA resistance should be taken into consideration during pDNA design and evaluation because it may significantly interfere in the

final transfection results. Our studies strongly suggest that the polyadenylation sequences play an important role in both, resistance and post-transcriptional efficiency. This information may be relevant not only for the optimization of vectors used in non-viral gene therapy studies, but also for the optimization of vectors employed in the production of recombinant proteins by transient transfection of mammalian cells.

Conclusions

The hypothesis that the modification of nuclease-sensitive sequences (hot spots) such as those present in BGH polyadenylation signals may significantly increase the resistance of plasmid vectors towards mammalian nucleases and their stability during storage has been confirmed. The elimination of these hot spots in poly As, by replacing for a synthetic poly A sequence, or the attempt to eliminate the hot spot by deletion of nucleotide sequences, typically lead to low levels of transgene expression by interfering in mRNA maturation or turnover. In one case, however, there are strong indications that the use of a synthetic poly A leads to higher transfection levels as a direct result of increased plasmid resistance. Further studies on the role of secondary structures located upstream and downstream of the AATAAA poly A signals are still necessary to increase the current knowledge of their effects on plasmid DNA stability and pre-mRNA maturation. The rational design of poly A signals could, in the near future, allow the construction of plasmid vectors with improved resistance to nucleases and specifically optimized at the post-transcriptional level.

Acknowledgements

The authors thank the Portuguese Fundação para a Ciência e a Tecnologia for financial support (POCI/BIO/55799/2004, and post-doctoral grant BPD/14879/2003 to A. R. Azzoni).

References

- Roth CM, Sundaram S. Engineering synthetic vectors for improved DNA delivery: Insights from intracellular pathways. *Annu Rev Biomed Eng* 2004; **6**: 397–426.
- Lechardeur D, Verkman AS, Lukacs GL. Intracellular routing of plasmid DNA during non-viral gene transfer. *Adv Drug Delivery Rev* 2005; **57**: 755–767.
- Pollard H, Toumaniantz G, Amos JL, *et al.* Ca²⁺-sensitive cytosolic nucleases prevent efficient delivery to the nucleus of injected plasmids. *J Gene Med* 2001; **3**: 153–164.
- Lew D, Parker SE, Latimer T, *et al.* Cancer gene therapy using plasmid DNA: pharmacokinetic study of DNA following injection in mice. *Hum Gene Ther* 1995; **6**: 553–564.
- Oh YK, Kim JP, Yoon H, *et al.* Prolonged organ retention and safety of plasmid DNA administered in polyethylenimine complex. *Gene Ther* 2001; **8**: 1587–1592.
- Crook K, McLachlan G, Stevenson BJ, *et al.* Plasmid DNA molecules complexed with cationic liposomes are protected from degradation by nucleases and shearing by aerosolisation. *Gene Ther* 1996; **3**: 834–839.
- Zabner J, Fasbender AJ, Moninger T, Poellinger KA, Welsh MJ. Cellular and molecular barriers to gene transfer by a cationic lipid. *J Biol Chem* 1995; **270**: 18997–19007.
- Gasiorowski JZ, Dean DA. Postmitotic nuclear retention of episomal plasmids is altered by DNA labeling and detection methods. *Mol Ther* 2005; **12**: 460–467.
- Ribeiro SC, Monteiro GA, Prazeres DMF. The role of polyadenylation signal secondary structures on the resistance of plasmid vectors to nucleases. *J Gene Med* 2003; **6**: 565–573.
- Hans H, Alwine JC. Functionally significant secondary structure of the simian virus 40 late polyadenylation signal. *Mol Cell Biol* 2000; **20**: 2926–2932.
- Wu C, Alwine JC. Secondary structure as a functional feature in the downstream region of mammalian polyadenylation signals. *Mol Cell Biol* 2004; **24**: 2789–2796.
- Clark DP. Processing of RNA. In *Molecular Biology: Understanding the Genetic Revolution*, Elsevier Academic Press: San Diego, 2005; 308.
- Rolland A. Gene medicines: the end of the beginning? *Adv Drug Deliv Rev* 2005; **57**: 669–673.
- Yew NS. Controlling the kinetics of transgene expression by plasmid design. *Adv Drug Deliv Rev* 2005; **57**: 769–780.
- Carapuça E, Azzoni AR, Prazeres DMF, Monteiro GA, Mergulhão FJM. Time-course determination of plasmid content in eukaryotic and prokaryotic cells using real-time PCR. *Mol Biotechnol* 2007; accepted for publication.
- Levitt N, Briggs D, Gil A, Proudfoot NJ. Definition of an efficient synthetic poly(A) site. *Genes Dev* 1989; **3**: 1019–1025.
- Lechardeur D, Lukacs GL. Intracellular barriers to non-viral gene transfer. *Curr Gene Ther* 2002; **2**: 1–12.
- Cheadle C, Fan J, Cho-Chung YS, *et al.* Stability regulation of mRNA and the control of gene expression. *Ann N Y Acad Sci* 2005; **1058**: 196–204.
- Gimmi ER, Reff ME, Deckman IC. Alterations in the pre-mRNA topology of the bovine growth hormone polyadenylation region decrease poly(A) site efficiency. *Nucleic Acids Res* 1989; **17**: 6983–6998.
- Remaut K, Sanders NN, Fayazpour F, Demeester J, De Smedt SC. Influence of plasmid DNA topology on the transfection properties of DOTAP/DOPE lipoplexes. *J Control Release* 2006; **115**: 335–343.
- Lipshutz GS, Titre D, Brindle M, Bisconte AR, Contag CH, Gaensler KM. Comparison of gene expression after intraperitoneal delivery of AAV2 or AAV5 in utero. *Mol Ther* 2003; **8**: 90–98.
- Garg S, Oran AE, Hon H, Jacob J. The hybrid cytomegalovirus enhancer/chicken beta-actin promoter along with woodchuck hepatitis virus posttranscriptional regulatory element enhances the protective efficacy of DNA vaccines. *J Immunol* 2004; **173**: 550–558.
- Virella-Lowell I, Zusman B, Foust K, *et al.* Enhancing rAAV vector expression in the lung. *J Gene Med* 2005; **7**: 842–850.
- Zufferey R, Donello JE, Trono D, Hope TJ. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J Virol* 1999; **73**: 2886.
- Le Bec C, Douar AM. Gene therapy progress and prospects – vectorology: design and production of expression cassettes in AAV vectors. *Gene Ther* 2006; **13**: 805–813.
- Yin W, Xiang P, Li Q. Investigations of the effect of DNA size in transient transfection assay using dual luciferase system. *Anal Biochem* 2005; **346**: 289–294.