Efficient GFP mutations profoundly affect mRNA transcription and translation rates

Andrea Sacchetti, Tarek El Sewedy[1], Ashraf F. Nasr, Saverio Alberti*

Laboratory of Experimental Oncology and Biotech Group, Department of Cell Biology and Oncology, Instituto di Ricerche Farmacologiche Mario Negri, Consorzio Mario Negri Sud, 66030 Santa Maria Imbaro (Chieti), Italy

Received 5 January 2001; accepted 6 February 2001

First published online 23 February 2001

Edited by Julio Celis

Abstract Green fluorescent protein (GFP) variants with higher expression efficiencies have been generated by mutagenesis. Favorable mutations often improve the folding of GFP. However, an effect on protein folding fails to explain the efficiency of several other GFP mutations. In this work, we demonstrate that mutations of the GFP open reading frame and untranslated regions profoundly affect mRNA transcription and translation efficiencies. The removal of the GFP 5′ untranslated region halves the transcription rate of the GFP gene, but hugely improves its translation rate. Mutations of the GFP open reading frame or the addition of peptide sequences differentially reduce the GFP mRNA transcription rate, translation efficiency and protein stability. These previously unrecognized effects are demonstrated to be critical to the efficiency of GFP mutants. These findings indicate the feasibility of generating more efficient GFP variants, with optimized mRNA transcription and translation in eukaryotic cells.

Key words: Green fluorescent protein; RNA transcription rate; RNA translation efficiency; Protein stability; Protein folding; Gene mutation

1. Introduction

Green fluorescent protein (GFP) is a spontaneously fluorescent protein that is widely used as a recombinant protein tag in living cells [1–3]. However, the expression of GFP is suboptimal in several experimental systems [4], often because of a low folding efficiency at high temperatures [4,5]. Several GFP mutants have thus been produced that demonstrate better fluorescence and generate higher protein levels than the wild-type (wt) [6–11]. Quite a few of these also show an improved protein folding [6,8,10,11]. However, it was open to question if all of the mutations with major effects on GFP expression levels affected protein folding or protein stability, or if they acted through other, as yet unrecognized, mechanisms [4–6,8].

Little is known about the effects of structural alterations of the GFP gene on its transcription and translation abilities [12]. In most genes, the latter are influenced by the 5′ and 3′ untranslated regions (UTs) of the mRNA [13–18], and by sequences surrounding the translation start site, such as the Kozak motif [19–23]. Sequence elements in the coding region also affect the efficiency of transcription, the mRNA stability and the mRNA translation rate [24–29]. Thus, we investigated if alterations in the GFP UT and open reading frame (ORF) (deletions, point mutations and added peptide tags, as test sequences for fusion to heterologous proteins [30,31]) could significantly modify the GFP mRNA transcription and translation rates, or mRNA stability. Our findings demonstrate that mutations of the GFP ORF and UT have major consequences on mRNA transcription and translation. These previously unrecognized effects are demonstrated to be critical to the efficiency of GFP mutants. These results indicate the feasibility of more efficient GFP expression constructs, and of specific mutagenesis and screening programs aimed at optimizing mRNA transcription and translation in eukaryotic cells.

2. Materials and methods

2.1. Vector construction

The full-length wtGFP cDNA (p10.1 plasmid) [12] was utilized to generate the wtGFP expression constructs with or without UTs. The wtGFP, and the S65T [6], Bex1 and Vex1 [7,32] mutants were amplified by PCR and subcloned in the pRK-5 expression vector [33] (Fig. 1). N-terminal (wtGFP-myc) and C-terminal (wtGFP-SB) tagged wtGFP were also generated. All constructs were sequenced to ascertain the absence of PCR-induced mutations.

2.2. Transfection of the GFP expression constructs

293T cells were cultured in DMEM with 10% fetal calf serum (Gibco-BRL, Paisley, UK) and 200 μg/ml geneticin (Sigma Chemical Co., St. Louis, MO, USA). Cells were seeded in 10 cm diameter dishes (Nunclon, Nunc, Denmark), and transfected by calcium phosphate co-precipitation, as previously described [34].

2.3. Northern blot analysis

RNA for Northern hybridization was extracted from transfected 293T cells. The GFP mRNA stability was quantified as described previously [33]. Briefly, transfected cells were treated with 10 μg/ml actinomycin D to stop RNA polymerase II transcription. Total RNA was extracted at different times after the treatment and analyzed by Northern blotting. The levels of the GFP transcripts were measured by densitometry, and analyzed with NIH Image 1.62, using a Kodak gray scale for internal standardization (http://www.kodak.com/country/US/en/motion/postProduction/tools/taf.shtml). mRNA levels were plotted against time (Fig. 2), and used to determine the mRNA half-life (Table 1). The ratio between steady-state mRNA levels and RNA stability was used to quantify the mRNA transcription rate. Since the mRNA stability of the different constructs was found to be essentially identical, the levels of mRNA were used as direct indicators of mRNA transcription rates (Table 1).

*Corresponding author. Fax: (39)-0872-570 412.
E-mail: alberti@cmns.mnegri.it

[1] Present address: King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia.

PII: S0014-5793(01)02246-3

2.4. Western blot analysis

Pre-cleared cell lysates were run on 12% SDS–PAGE and transferred to nitrocellulose filters. The filters were treated with rabbit anti-GFP (Clontech) and peroxidase-conjugated anti-rabbit antisera (Calbiochem, La Jolla, CA, USA) essentially as described previously [35]. Antibody-binding was revealed by chemiluminescence (ECL, Amersham, Buckinghamshire, UK) (Fig. 3). The GFP protein stability was quantified after blocking protein synthesis with cycloheximide (25 μg/ml). Cells were lysed at different times after the treatment, and analyzed by Western blotting (Fig. 3). The GFP levels (*P*) at the different time points were measured by densitometry, analyzed with NIH Image 1.62, and used to determine the protein half-life (Table 1).

The mRNA translation rates of the different constructs (how many molecules of GFP are synthesized per mRNA molecule) were calculated as the ratio between corresponding protein and mRNA levels. To correct for the different stabilities of the encoded proteins, the protein levels were normalized versus the protein half-lives ($P/P\ t_{1/2}$). As for the mRNA transcription rates, there was no need to normalize for mRNA stability, since this was constant among the different constructs (Fig. 2).

2.5. Fluorescence analysis

The GFP fluorescence analysis was performed by flow cytometry [34] (FACStar and Vantage, Becton-Dickinson, Synnyvale, CA, USA) (Fig. 3), and by spectrofluorimetry (CM1T11I Spex spectrofluorimeter, Spex Industries, Edison, NJ, USA). Blue-excited GFPs are better chromophores [36] than the wtGFP by 1.7-fold. However, the wtGFP is excited about 3.5-fold less at 488 nm than at 396 nm [37,38]. Thus, blue-excited GFPs emit around 6-fold more light than wtGFP when excited at 488 nm in flow cytometry. Fluorescence half-lives essentially coincided with those of the corresponding peptides reported in Table 1.

3. Results and discussion

The wtGFP, and first (S65T) [6] and second (Bex1 and Vex1) [7] generation GFP mutants were expressed in mammalian cells and analyzed. These constructs were designed to include either or both the 5′ and 3′ UTs of the GFP gene, N- or C-terminal GFP amino acidic tags, different Kozak initiation of translation sequences [21], and different mutations of the GFP ORF (Fig. 1 and Table 1). The S65T construct bears this single mutation [6], whereas Bex1 and Vex1 bear multiple ORF mutations (S65T, V163A for Bex1; V163A, S202F, T203Y for Vex1) [7,32]. All constructs were transiently transfected into mammalian cells. The GFP mRNA transcription and translation rates and stabilities were measured, together with the GFP protein and fluorescence levels and stabilities (Figs. 2 and 3). The ratio between steady-state mRNA levels and RNA stability was used to quantify the GFP mRNA transcription rates. Since the mRNA stability was found constant for the different GFP variants, the mRNA levels were used as direct indicators of the efficiency of mRNA transcription (Table 1). The ratio between steady-state protein and mRNA levels was used to quantify the mRNA translation rates. This ratio was normalized versus the protein half-life, to correct for the differential effects of each mutation on protein stability. Table 1 presents the results obtained in transfected 293T cells. Analogous results were obtained in transfected COS-7 and 293 cells, supporting the widespread validity of these findings.

3.1. UTs

The deletion of the 5′ UT from wtGFP caused a 50% reduction in the GFP mRNA levels (Table 1). Since the stability of the Δ5′ UT mRNA was not affected (Table 1 and Fig. 2), this indicated that the removal of the 5′ region of the gene reduced the GFP mRNA transcription rate by 50%. On the other hand, the deletion of the 5′ UT led to a 200-fold increase in the GFP protein levels (Table 1 and Fig. 3; wtGFP-Δ5′ UT versus wtGFP) [12]. The wtGFP-Δ5′ UT and wtGFP constructs expressed identical, native GFP proteins, and these were confirmed to have an identical half-life (Table 1). Thus, wtGFP-Δ5′ UT mRNA possesses a 200-fold higher translation rate than wtGFP mRNA. The wtGFP-Δ5′ UT and wtGFP share identical Kozak initiation of translation sequences (Table 1 and see below) [21]. Thus, potent translation regulatory elements other than the Kozak motif exist within the 5′ UT of GFP. These may impair ribosome scanning [39]

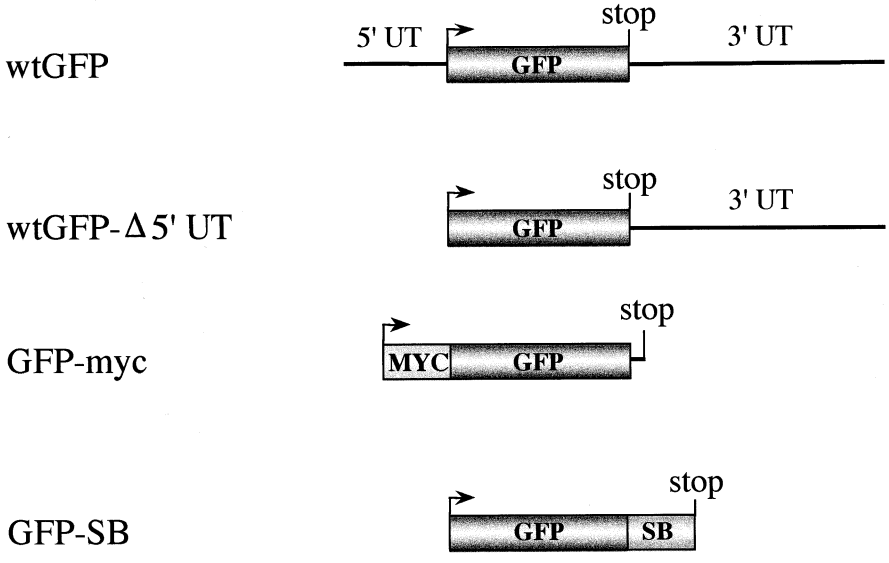

Fig. 1. Engineered GFP variants. wtGFP: full-length wtGFP cDNA in germline configuration. This includes the 5′ and 3′ UTs. wtGFP-Δ5′ UT: wtGFP that includes the 3′ UT, but is devoid of the 5′ UT; GFP-myc: N-terminal myc-tagged GFPs (wt, S65T, Vex1 or Bex1), devoid of the 5′ and 3′ UTs; GFP-SB: C-terminal streptavidin-binding tagged wtGFP, devoid of the 5′ and 3′ UTs. UT: untranslated region; bent arrow: ATG translation start codon; stop: stop codon; MYC: myc tag; SB: streptavidin-binding tag.

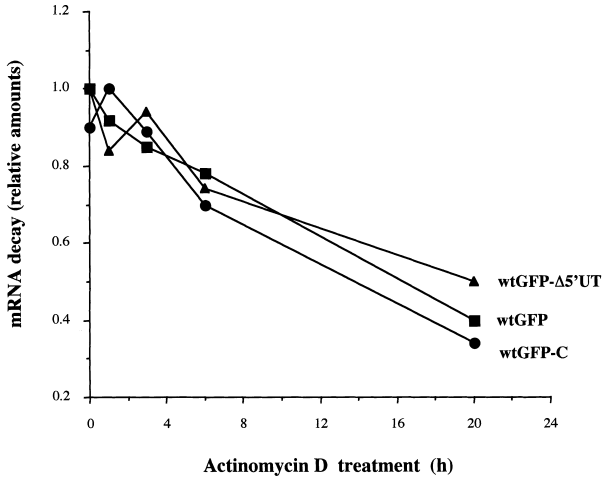

Fig. 2. GFP mRNA stability in transfected 293T cells. mRNA synthesis was blocked with actinomycin D and the relative residual amounts of GFP mRNA were quantified at different time points. wtGFP: squares; wtGFP-Δ5′ UT: triangles; wtGFP-myc: circles.

or bind negative modulators of GFP translation [29,40], and their removal allows the efficient translation of the GFP mRNA. On the other hand, all of the mRNA devoid of a 3′ UT displayed a lower translation rate than the wtGFP-Δ5′ UT, suggesting a translation-enhancing effect of the 3′ UT of GFP [16,18].

3.2. Point mutations

Mutations in the GFP ORF were found to affect the translation rate of the corresponding mRNA. The V163A mutation was shown to cause a reduction in the translation rate of the GFP mRNA by almost one-half (Table 1: compare Bex1-myc with S65T-myc). This result was rather unexpected, since Bex1-myc is an efficient V163A mutant [7]. However, a similarly low translation rate was demonstrated by Vex1-myc (Table 1). The finding that a single codon change in the GFP ORF is sufficient to profoundly affect the GFP translation rate is novel. It would be of interest to determine if this single codon mutation directly affects translation, e.g. by translation pausing due to RNA secondary structures [28], or if it induces the binding of translation-inhibitory factors [29].

3.3. GFP tags

GFP is routinely used as a fusion partner to heterologous proteins, and these have been demonstrated to heavily affect GFP function [5,30,41]. GFP tags known to affect the folding of GFP [41] were used in this work as examples of heterologous fusion partners. Remarkably, the SB tag was found to affect both mRNA transcription and translation rates, confirming the role of coding sequences in regulating the mRNA transcription efficiencies [24] and translation rates [28,29].

The wtGFP-SB and wtGFP-myc bear Kozak initiation of translation sequences of different strength [21] (Table 1). However, different Kozak sequences are unlikely to account for major differences in GFP translation efficiencies. Indeed, wtGFP-Δ5′ UT demonstrates the highest translation rate among the constructs tested in this work. In particular, it is twice as efficient as wtGFP-myc, in spite of the largely suboptimal Kozak sequence of the former versus the 'perfect' Kozak motif in the latter [21,42]. The 200-fold better translation of wtGFP-Δ5′ UT compared to that of the wtGFP, in spite of an identical germline Kozak motif, is also consistent with an overall minor role in GFP translation (Table 1, Fig. 3 and see above). Kozak motifs were first identified by comparison of numerous eukaryotic mRNA sequences that surround the AUG start codon [43]. However, the relative strength of different Kozak sequences was experimentally assessed largely in expression systems for preproinsulin or chloramphenicol acetyl transferase [21]. Thus the strength of Kozak sequences may vary depending on their sequence context [44]. In other words, mRNA elements other than the Kozak motif contribute to the determination of their overall translation competence, and the relative importance of these different elements probably varies among different mRNAs.

wtGFP-SB, Bex1-myc and Vex1-myc demonstrate the highest folding levels (up to 80% of the GFP molecules) [41] and the lowest translation rates among the GFP variants analyzed (Table 1). This raised the possibility that high folding levels were mechanistically linked to a 'slow' translation [45,46]. However, a comparison of wtGFP and wtGFP-Δ5′ UT, which are identical wtGFP molecules with widely different mRNA translation rates, demonstrates essentially identical protein folding abilities (25% of the GFP molecules [41]). Thus, the GFP folding efficiency appears to be a fundamental property of specific peptide sequences, and is unlikely to be causally related to translational 'speed'.

Table 1
GFP mRNA stabilities and transcription/translation rates

GFP form	mRNA levels[a]	mRNA $t_{1/2}^{b}$	P^{c}	$P\ t_{1/2}^{d}$	Translation rate[e]	Kozak sequence[f]
wtGFP	100	7	1	54	1	caaagATGa
wtGFP-Δ5′ UT	50	7	170	54	210	caaagATGa
wtGFP-SB	80		30	33	38	cgcatATGa
wtGFP-myc	100	7	100	33	100	agaccATGg
S65T-myc	100		80	25	103	agaccATGg
Bex1-myc	100		60	32	62	agaccATGg
Vex1-myc	100		40	25	55	agaccATGg

The results are average of three independent experiments. All values, except for half-lives, are normalized versus wtGFP-myc and expressed as percentage.
[a]GFP mRNA levels.
[b]GFP mRNA half-life expressed in hours.
[c]GFP protein levels (*P*).
[d]GFP protein half-life expressed in hours.
[e]mRNA translation rates.
[f]Kozak ribosomal binding sequence [21]; the ATG start of translation is in capital letters.

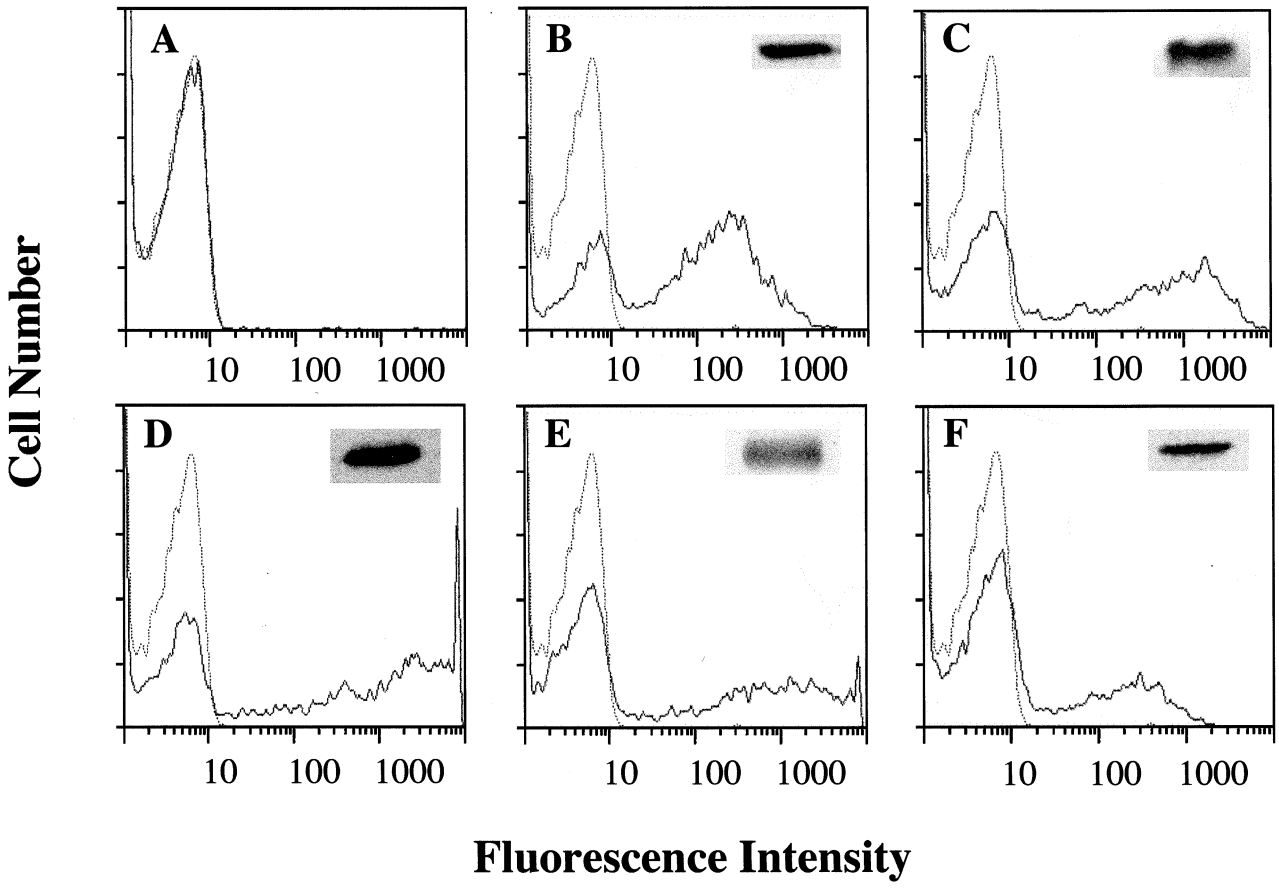

Fig. 3. Expression levels of the transfected GFP constructs. 293T cells were transfected with (A) wtGFP; (B) wtGFP-myc; (C) Bex1-myc; (D) wtGFP-Δ5′ UT; (E) wtGFP-SB; (F) S65T-myc. Panels show log fluorescence profiles of 5000 cells. Solid line: transfected cells. Dotted line: vector-alone transfected cells. Insets: GFP protein levels, as determined by Western blot analysis. The GFP protein levels in (A) were too low to be shown on scale with the other samples.

4. Conclusions

In the past, structural and functional analyses of GFP mutants identified in screening programs have repeatedly raised questions as to their actual mechanisms of action [4,6,8]. Our results demonstrate that several commonly engineered alterations in GFP structure have remarkable consequences on mRNA expression parameters. The major affected mechanism is the efficiency of mRNA translation, that was shown to span a 200-fold range. Significant effects were also demonstrated on protein stability and mRNA transcription rates. These previously unrecognized effects were found to be critical to the overall efficiency of GFP mutants (Table 1; compare the translation rates of the different constructs and the corresponding GFP protein levels). Thus, a strategy of optimization of GFP mRNA expression parameters may contribute considerably to the improvement of GFP expression. The engineering of the 5′ and 3′ UTs of GFP, in particular, might permit the reaching of both higher transcription and translation rates. If GFP mutations are found to have similar consequences on GFP mRNA transcription and translation rates in lower eukaryotes, high throughput screening strategies for more efficient GFP mutants could be efficiently designed in commonly used yeast strains [47,48].

Acknowledgements: This work was supported by the Italian Association for Cancer Research (AIRC), the Italian National Research Council (Convenzione CNR, Consorzio Mario Negri Sud), and the Consorzio per la Medicina Tropicale (CMT). A.S. is recipient of a fellowship from Italian Foundation for Cancer Research (FIRC). T.E.S. and A.F.N. were recipients of fellowships from the CMT.

References

[1] Prasher, D.C. (1995) Trends Genet. 11, 320–323.
[2] Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A. and Tsien, R.Y. (1995) Trends Biochem. Sci. 20, 448–455.
[3] Ludin, B. and Matus, A. (1998) Trends Cell Biol. 8, 72–77.
[4] Tsien, R.Y. (1998) Annu. Rev. Biochem. 67, 509–544.
[5] Sacchetti, A. and Alberti, S. (1999) Nat. Biotechnol. 17, 1046.
[6] Heim, R., Cubitt, A.B. and Tsien, R.Y. (1995) Nature 373, 663–664.
[7] Anderson, M.T., Tjioe, I.M., Lorincz, M.C., Parks, D.R., Herzenberg, L.A. and Nolan, G.P. (1996) Proc. Natl. Acad. Sci. USA 93, 8508–8511.
[8] Heim, R. and Tsien, R.Y. (1996) Curr. Biol. 6, 178–182.
[9] Crameri, A., Whitehorn, E.A., Tate, E. and Stemmer, W.P.C. (1996) Nat. Biotechnol. 14, 315–319.
[10] Cormack, B.P., Valdivia, R.H. and Falkow, S. (1996) Gene 173, 33–38.
[11] Siemering, K.R., Golbik, R., Sever, R. and Haseloff, J. (1996) Curr. Biol. 6, 1653–1663.
[12] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C. (1994) Science 263, 802–805.

[13] Pantopoulos, K., Johansson, H.E. and Hentze, M.W. (1994) Prog. Nucleic Acids Res. Mol. Biol. 48, 181–238.
[14] Tanguay, R.L. and Gallie, D.R. (1996) Mol. Cell. Biol. 16, 146–156.
[15] Al-Qahtani, A. and Mensa-Wilmot, K. (1996) Nucleic Acids Res. 24, 1173–1174.
[16] Munroe, D. and Jacobson, A. (1990) Mol. Cell. Biol. 10, 3441–3455.
[17] Hann, L.E., Webb, A.C., Cai, J.-M. and Gehrke, L. (1997) Mol. Cell. Biol. 17, 2005–2013.
[18] Bailey-Serres, J. and Dawe, R.K. (1996) Plant Physiol. 112, 685–695.
[19] Iida, Y. and Masuda, T. (1996) Nucleic Acids Res. 24, 3313–3316.
[20] Le, S.Y. and Maizel, J.V.J. (1997) Nucleic Acids Res. 25, 362–369.
[21] Kozak, M. (1991) J. Biol. Chem. 266, 19867–19870.
[22] Sachs, A.B., Sarnow, P. and Hentze, M.W. (1997) Cell 89, 831–838.
[23] Afshar-Kharghan, V., Li, C.Q., Khoshnevis-Asl, M. and Lopez, J.A. (1999) Blood 94, 186–191.
[24] Xiang, S., Parsons, H.K. and Murray, M. (1998) Gene 209, 123–129.
[25] Bernstein, P.L., Herrick, D.J., Prokipcak, R.D. and Ross, J. (1992) Genes Dev. 6, 642–654.
[26] Veyrune, J.L., Carillo, S., Vie, A. and Blanchard, J.M. (1995) Oncogene 11, 2127–2134.
[27] Ross, J. (1995) Microbiol. Rev. 59, 423–450.
[28] Zama, M. (1999) Nucleic Acids Symp. Ser. 42, 81–82.
[29] Xu, Y.H. and Grabowski, G.A. (1999) Mol. Genet. Metab. 68, 441–454.
[30] Waldo, G.S., Standish, B.M., Berendzen, J. and Terwilliger, T.C. (1999) Nat. Biotechnol. 17, 691–695.
[31] Keiler, K.C., Waller, P.R. and Sauer, R.T. (1996) Science 271, 990–993.
[32] Sacchetti, A., Ciccocioppo, R. and Alberti, S. (2000) Histol. Histopathol. 15, 101–107.
[33] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Cold Spring Harbor Laboratory, New York.
[34] Dell'Arciprete, R., Stella, M., Fornaro, M., Ciccocioppo, R., Capri, M.G., Naglieri, A.M. and Alberti, S. (1996) J. Histochem. Cytochem. 44, 629–640.
[35] El-Sewedy, T., Fornaro, M. and Alberti, S. (1998) Int. J. Cancer 75, 324–331.
[36] Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L. and Lukyanov, S.A. (1999) Nat. Biotechnol. 17, 969–973.
[37] Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R. and Piston, D.W. (1997) Biophys. J. 73, 2782–2790.
[38] Ward, W.W., Prentice, H.J., Roth, A.F., Cody, C.W. and Reeves, S.C. (1982) Photochem. Photobiol. 35, 803–808.
[39] Beyer, D., Skripkin, E., Wadzack, J. and Nierhaus, K.H. (1994) J. Biol. Chem. 269, 30713–30717.
[40] Dubnau, J. and Struhl, G. (1996) Nature 379, 694–699.
[41] Sacchetti, A., Cappetti, V., Marra, P., Dell'Arciprete, R., El Sewedyl, T., Crescenzi, C. and Alberti, S. (2000) J. Cell. Biochem.
[42] Kozak, M. (1986) Cell 44, 283–292.
[43] Kozak, M. (1984) Nucleic Acids Res. 12, 857–872.
[44] Kozak, M. (1989) Mol. Cell. Biol. 9, 5073–5080.
[45] Fedorov, A.N. and Baldwin, T.O. (1997) J. Biol. Chem. 272, 32715–32718.
[46] Harding, H.P., Zhang, Y. and Ron, D. (1999) Nature 397, 271–274.
[47] Cereghino, G.P. and Cregg, J.M. (1999) Curr. Opin. Biotechnol. 10, 422–427.
[48] Sudbery, P.E. (1996) Curr. Opin. Biotechnol. 7, 517–524.