

# Versatile vectors to study recoding: conservation of rules between yeast and mammalian cells

Guillaume Stahl\*, Laure Bidou<sup>1</sup>, Jean-Pierre Rousset<sup>1</sup> and Michel Cassan

Institut de Génétique et Microbiologie, URA CNRS 1354, Bâtiment 400, Université Paris-Sud, 91405 Orsay Cedex, France and <sup>1</sup>UFR 927, Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris Cedex 05, France

Received December 16, 1994; Revised and Accepted March 16, 1995

## ABSTRACT

In many viruses and transposons, expression of some genes requires alternative reading of the genetic code, also called recoding. Such events depend on specific mRNA sequences and can lead to read through of an in-frame stop codon or to +1 or –1 frameshifting. Here, we addressed the issue of conservation of recoding rules between the yeast *Saccharomyces cerevisiae* and mammalian cells by establishing a versatile vector that can be used to study recoding in both species. We first assessed this vector by analysing the site of +1 frameshift of the Ty1 transposon. Two sequences from higher organisms were then tested in both yeast and mammalian cells: the *gag-pol* junction of human immunodeficiency virus type 1 (HIV-1) (a site of –1 frameshift), and the stop codon region of the replicase cistron from the tobacco mosaic virus (a site of UAG read through). We show that both sequences direct a high level of recoding in yeast. Furthermore, different mutations of the target sequences have similar effects on recoding in yeast and in mouse cells. Most notably, a strong decrease of frameshifting was observed in the absence of the HIV-1 stem-loop stimulatory signal. Taken together, these data suggest that mechanisms of some recoding events are conserved between lower and higher eukaryotes, thus allowing the use of *S.cerevisiae* as a model system to study recoding on target sequences from higher organisms.

## INTRODUCTION

The mechanism of translation is one of the cell features which is best conserved in evolution. Triplets of nucleotides are sequentially decoded in amino acids which are added to the nascent peptidyl chain until a stop codon is encountered. However, there are exceptions to this general rule; the best documented cases being frameshifting (i.e. slippage of the ribosome either backward or forward), and read through (i.e. natural suppression of a stop codon in-frame between two coding regions). These modifications of the classical reading of triplets depend on specific mRNA sequences and/or structures (1) and have been termed 'recoding' by Gesteland *et al.* (2).

Accurate quantification of recoding efficiencies is fundamental for dissecting *cis* elements involved in these mechanisms. We have previously constructed expression vectors to quantify recoding efficiency in mammalian cells in culture (pRSVL series) (3). Using these vectors, we demonstrated that a palindromic sequence downstream of the frameshift site in HIV-1 acts as a stimulator of frameshifting in different cell types (4). We were also able to show that HIV-1 infection of T lymphoid cells does not significantly increase the frameshifting efficiency (4). However, identification of *trans*-acting factors is limited in mammalian cells because of the lack of genetic approach. We thus decided to explore the yeast *Saccharomyces cerevisiae* as an alternative eukaryotic model.

Here we present a versatile expression cloning vector based on a simple and highly sensitive assay that can be used to analyse *cis*- and *trans*-acting factors involved in recoding efficiency on heterologous targets, either in yeast or in mammalian cells.

The following specific requirements are met with this system: (i) expression in yeast and mammalian cells, (ii) an internal reference for translation initiation efficiency, and (iii) oriented cloning sites to insert recoding target sequences.

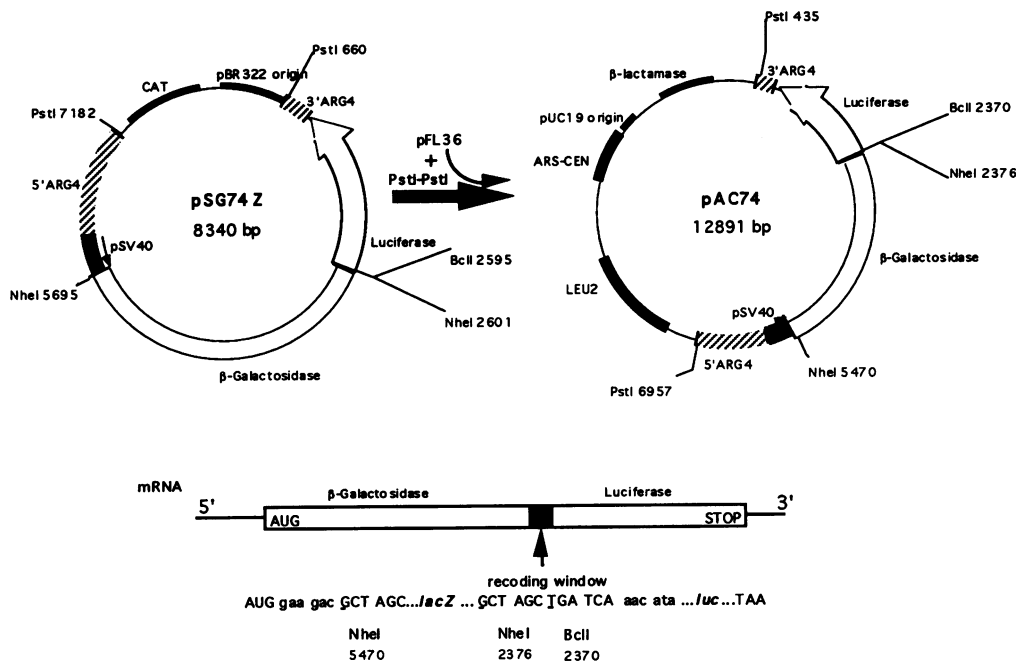
Using two heterologous targets: the –1 frameshift site of the *gag-pro/pol* junction of HIV-1 and the UAG read through of the TMV replicase cistron, we show that both are sites of highly efficient recoding in yeast. Furthermore, mutations known to decrease recoding in the natural host cells have the same effect in yeast. These results establish that yeast is a suitable host to study recoding target sequences from higher eukaryotes *in vivo* and suggest that mechanisms involved in at least some –1 frameshifting and read through events are conserved between lower and higher eukaryotes.

## MATERIALS AND METHODS

### Recoding target sequences

After *NheI* and *BclI* digest which destroys the *lacZ* ORF in pAC74 (see Fig. 1), pairs of complementary oligonucleotides containing the target sequence were inserted (white/blue selection). The oligonucleotides are flanked by a *NheI* site and a *BclI* compatible cohesive end without the first T, in order to destroy the TGA stop codon (see sequence in Fig. 1). The *lacZ* fragment (from the CmR pSG74 Z plasmid) is then re-inserted in the new plasmid at the *NheI* site (blue/white selection). Cm resistance and Amp sensitivity of pSG74Z allow a parental selection for the pAC derivatives: only

\* To whom correspondence should be addressed



**Figure 1.** Description of the cloning vectors. pSG74 Z derives from pSG74, which was obtained by insertion of a *PvuII* fragment from pCYM1L, containing *lac* under control of SV40 early promoter, in pMY237 *HpaI* site. The pCYM1L is a pCYM01 derivative (5), and pMY237 is a pUC derivative containing the *CAT* gene, 5' (1 kb) and 3' (150 bp) parts of *ARG4* *PstI-PstI* fragment from *S.cerevisiae* (M. Cassan, unpublished results). A *NheI-NheI* fragment containing the *E.coli lacZ* ORF with neither initiation nor stop codons was constructed from pUR278 (6) by PCR amplification using the oligonucleotides: W1 5'-GGAAACAGCTAGCACCATGAT-TACG-3' and C2 5'-CTAGAGTCGAGCTAGCGGGATCCCC-3'. The fragment was then inserted at the beginning of the *lac* coding sequence in the *NheI* site of pSG74, in-frame with the *lac* AUG, giving rise to pSG74 Z. The largest *PstI* fragment of the pSG Z family constructs was introduced in the unique *PstI* site of pFL36, giving rise to the pAC family. pFL36 is a shuttle vector, containing a replicative origin in *E.coli* and a  $\beta$ -lactamase gene (ampicillin resistance); it is also replicative (ARS/CEN) and selectable (*LEU2*) in *S.cerevisiae* (12).

blue clones on ampicillin have the right structure. Each region [150 nucleotides (nt)] surrounding the recoding site has been sequenced. All recoding target sequences studied in this report are shown in Figure 2.

### Strains

Caesium chloride purified plasmids were used to transform CM5 $\alpha$  *Escherichia coli* strain (5,6).

Haploid *S.cerevisiae* strain Fy1679-18B $\alpha$  (7) was transformed using the lithium acetate method (8). At least three independent transformants were grown in 2 ml rich media (1% yeast extract, 2% bactopectone, 2% glucose) to early stationary phase. In these conditions, there is <4% plasmid loss per generation (data not shown). Culture aliquots were then diluted 1:20 in 4 ml of rich medium, grown to reach OD<sub>600</sub> = 1.5 and centrifuged. Crude cellular extracts were obtained by vortexing cell pellets for 30 min at 4°C in 150  $\mu$ l of luciferase assay buffer (9), including an equal volume of glass beads. The crude extracts were cleared by centrifugation at maximum speed for 5 min at 4°C in a Eppendorf centrifuge. There were no significant clonal variation between independent yeast transformants.

The same constructs were transfected in NIH3T3 cells, using the calcium phosphate precipitation method (10). Crude extracts were obtained as described before (3).

### Enzyme assays

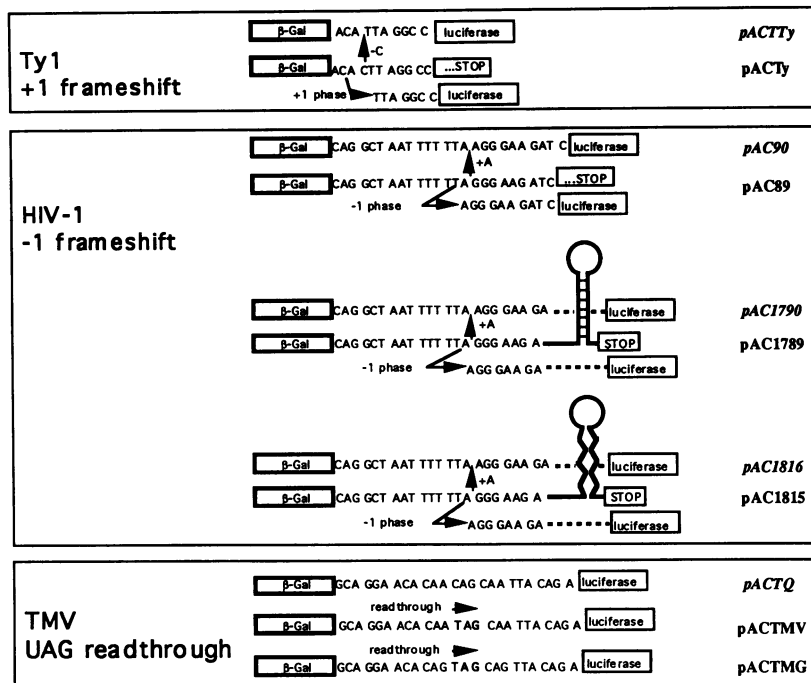
Luciferase (9) and  $\beta$ -galactosidase activities (11) were measured from the same crude extract as described elsewhere (4).

## RESULTS AND DISCUSSION

### General strategy

A hybrid *lacZ-luc* gene under control of the SV40 early promoter, was cloned on a shuttle expression vector (pFL36) replicative in *E.coli* and *S.cerevisiae* (12), giving rise to the pAC series (see Fig. 1). As the SV40 promoter is functional both in yeast and higher organisms (5), the pAC vectors can be used directly to transform yeast cells and to transfect mammalian cells in culture, allowing direct and precise comparison of recoding efficiencies. Target sequences were inserted at the junction of the *lacZ-luc* hybrid gene in such a way that each ribosome initiating translation will give rise to  $\beta$ -galactosidase activity but an active luciferase will be synthesized only if a recoding event takes place during translation of the hybrid mRNA (see Fig. 1). The  $\beta$ -galactosidase is used as an internal reference to estimate the overall expression level of each construct thus controlling for vector stability, transfection efficiency, transcriptional and translational rates.

Since introduction of an additional sequence at the amino terminal end of the luciferase protein may alter its specific activity, we included control constructs in which no recoding is required to obtain a  $\beta$ -galactosidase-luciferase fusion protein. Each fusion protein synthesized from the control constructs will be exactly the same as that obtained after recoding on the test constructs. To quantify recoding, we estimated the ratio between the luciferase activity of the various test constructs to that of the corresponding control constructs, after standardization for expression efficiency using  $\beta$ -galactosidase activity.



**Figure 2.** Synopsis of recoding systems and sequence of the recoding sites. Name of the test vectors are in plain letters; those of control vectors are in italics. Sequence of oligonucleotides are listed in references (3) and (4).

### Ty1 +1 frameshifting

In order to assess whether the *lacZ-luc* hybrid gene is functional in yeast, we studied the slippery site of Ty1 transposon in its natural host (13,14). This frameshift event is mediated by a 7 nt sequence CUU AGG C where the rare AGG arginine codon acts as a stimulatory signal (14). When tRNA<sup>ARG</sup> is in limiting concentration, the tRNA<sup>LEU</sup> slips one base forward, from CUU to UUA, thus achieving a +1 frameshift. In the control construct, a C was omitted in the first position of the slippery sequence CTT AGG C (14; and see Fig. 2). Figure 3a shows that, in yeast, the +1 frameshifting efficiency of the Ty1 slippery site carried by the pACTy is 32% ± 2.5. Previous reports, using a *lacZ* reporter system, have shown that the Ty1 +1 frameshifting efficiency varies between 25 and 40% in yeast (13,14). Our results are therefore in good agreement with those already described, which indicated that the *lacZ-luc* reporter hybrid gene is suitable for quantification of recoding events in yeast.

By contrast, only a very low recoding efficiency ( $7 \times 10^{-4}$ ) was observed with this construct in mammalian cells (Fig. 3a). This very low level may be due to minor but specific differences between the translational machineries of yeast and mammals. For example, the involved tRNA<sup>ARG</sup> may be more abundant in mammalian cells, thus inhibiting frameshift initiation. This interpretation is consistent with the observation that the corresponding AGG codon is much more frequent in mouse than in yeast (15).

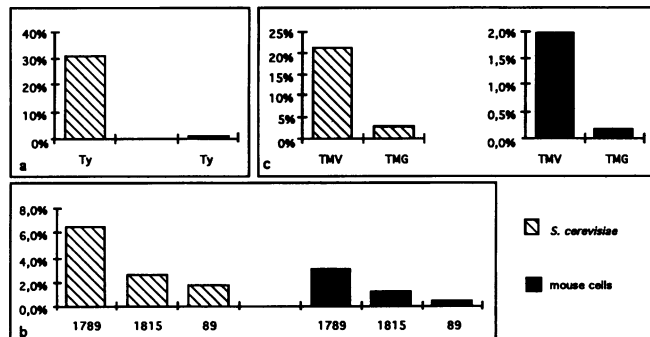
### HIV-1 frameshifting

This frameshift takes place at a 7 nt slippery site (action site): TTT TTA. In mammalian cells, frameshifting efficiency is increased by a secondary structure (stimulatory signal) precisely

positioned downstream from the action site. The stimulatory signal is a stem-loop and possibly acts through slowing down the ribosome. By contrast, Wilson and co-workers (16) did not detect any effect of the palindromic sequence in a study where the HIV-1 *gag-pol* frameshift has been analysed in yeast cells. This might be due to difficulties to accurately measure frameshifting in the detection system used or alternatively, to fundamental differences between yeast and mammalian cells regarding the mechanism of frameshifting. We have investigated the effect of the potential secondary structure on frameshifting efficiency, both in yeast and mammalian cells. We used pAC vectors to compare the frameshifting efficiency obtained with the HIV-1 minimal sequence (pAC89) and that obtained in presence of the natural HIV-1 downstream stimulatory signal (pAC1789). We also included a mutated sequence (pAC1815), where most of the base pairing of the stem has been destroyed and which has been shown to result in a 2-fold decrease in frameshifting efficiency in mammalian cells (4). Each control construct was designed by addition of an A after the TTA codon from the slippery sequence (17; and see Fig. 2).

The results obtained in NIH3T3 mouse fibroblasts are shown in Figure 3b. The frameshifting efficiency is 0.6% ± 0.4 with the minimal sequence (pAC89) and is increased five times (3% ± 1.4) when the stimulatory signal is present (pAC1789). The pAC1815 construct with the disrupted palindrome gave an intermediate frameshifting efficiency (1.3% ± 0.1). These recoding efficiencies are very similar to (i) those previously obtained using the pRSVL series vectors, with no *lacZ* internal control (4), (ii) other constructs in various mammalian cells (3,4,18), and (iii) another reporter system in avian or simian cells (19).

In heterologous yeast cells, a 3.3-fold increase of frameshifting efficiency was obtained in presence of the palindromic sequence: 2% ± 0.2 for the pAC89 minimal construct, 6.6% ± 1.5 for the pAC1789 construct containing the wild-type palindrome. The



**Figure 3.** Recoding efficiency in yeast and mouse cells. (a) Ty1 +1 frameshifting (7 nt minimal sequence). (b) gag-pol -1 frameshifting: HIV-1 slippery sequence with a wild-type (1789), mutated (1815) or without (89) palindromic downstream sequence. (c) TMV replicase read through: wild-type (TMV) or mutant (TMG) sequence. Results are expressed as the ratio of luciferase activity obtained with the test vectors compared to their respective control, after  $\beta$ -galactosidase normalisation. Each value represents the mean of at least five experiments in yeast and three in NIH3T3. Since identical results were obtained with pRSVL and pAC derivative vectors for the HIV-1 constructs, only pRSVL derivative were used for TMV and TMG measurements in NIH3T3. In NIH3T3, the construct with the smallest efficiency (pACTy) gave a luciferase activity between  $3 \times 10^2$  and  $6 \times 10^3$  relative light units (RLUs) (mock transfected cells gave a luciferase activity of  $10^2$  RLU). For the control constructs, the luciferase activity was between  $5 \times 10^3$  and  $4 \times 10^5$  RLU. In yeast, values obtained with pAC89 were between  $8 \times 10^2$  and  $1.2 \times 10^3$  RLU while those of the control constructs were between  $2.5 \times 10^3$  RLU and  $2 \times 10^5$  RLU.

frameshift level is intermediate for the pAC1815 construct containing the disrupted palindrome ( $2.8\% \pm 0.5$ ) (Fig. 3b). Therefore, the stem-loop acts also on -1 frameshifting in yeast.

These results suggest that the translational machineries of yeast and mammalian cells use similar mechanisms in recognizing the frameshift site and the stimulatory signal.

### TMV read through

Finally, we examined whether yeast cells could achieve programmed translational read through, using the region containing the UAG leaky stop codon from TMV. We used a construct (pACTMV) containing the wild-type minimal region UAG CAA UUA (20) and a modified construct where the two CAA codons surrounding the UAG stop codon are replaced by two CAG codons (pACTMG; see Fig. 2). These positions are important for a high level read through in plant cells (20) as well as in mouse cells (21; Rousset and Cassan, unpublished). For the control vector, the UAG codon was replaced by a sense CAG codon (22). As shown in Figure 3c, the efficiency of read through at the TMV site is very high in yeast cells ( $21\% \pm 4.8$ ) and is reduced 7-fold by mutations at the two critical sites (pACTMG:  $3.1\% \pm 0.3$ ). In mouse NIH3T3 cells, the read through efficiency was  $2\% \pm 0.7$  for pACTMV and  $0.2\% \pm 0.08$  for pACTMG. Therefore, the mutation also influences the recoding rate to the same extent in yeast and mouse cells. These results are also in agreement with those obtained in plant protoplasts (20).

### Conclusion

The results we obtained in yeast compared to higher eukaryotic cells indicate an overall increased recoding efficiency in yeast

cells, ~2-fold for the HIV-1 constructs and 10-fold for the TMV constructs. Although the reason for such quantitative variations remains unclear, this high level would be useful for isolating down mutants affected in the recoding process (specific 'antisuppressor'), which might lead to the characterisation of new translational factors.

Altogether, the results reported here show that some recoding rules are conserved between yeast and mammalian cells; they strongly suggest conservation of the mechanisms involved. They also clearly indicate that *S.cerevisiae* is a convenient and reliable host to analyse this type of recoding events taking place on target sequences active in higher eukaryotes.

### ACKNOWLEDGEMENTS

We are grateful to Marguerite Picard for constant support throughout the course of this work, and to Mounira Amor who welcomed us in her laboratory for transfection experiments. During the course of this work, LB was supported by the 'Fondation Singer-Polignac'. Part of this work was supported by the 'Association pour la Recherche sur le Cancer' (ARC no. 3065) and by the 'Agence Nationale de Recherches sur le Sida' (ANRS no. 95035).

### REFERENCES

- Atkins, J.F., Weiss, R.B. and Gesteland, R.F. (1990) *Cell*, **62**, 413-423.
- Gesteland, R.F., Weiss, R.B. and Atkins, J.F. (1992) *Science*, **257**, 1640-1643.
- Cassan, M., Berteaux, V., Angrand, P.O. and Rousset, J.P. (1990) *Res. Virol.*, **141**, 597-610.
- Cassan, M., Delaunay, N., Vaquero, C. and Rousset, J.P. (1994) *J. Virol.*, **68**, 1501-1508.
- Camonis, J.H., Cassan, M. and Rousset, J.P. (1990) *Gene*, **86**, 263-268.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. New York.
- Dujon, B. (1994) *Nature*, **369**, 371-377.
- Gietz, D., St. Jean, A., Woods, R.A. and Schiestl, R.H. (1992) *Nucleic Acids Res.*, **20**, 1425.
- Nguyen, V.T., Morange, M. and Bensaude, O. (1988) *Ann. Biochem.*, **171**, 404-408.
- Graham, F.L. and Van der Eb, A.J. (1973) *Virology*, **52**, 456-467.
- Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY.
- Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G., Labouesse, M., Minvielle-Sebastia, L. and Lacroute, F. (1991) *Yeast*, **7**, 609-615.
- Clare, J.J., Belcourt, M. and Farabaugh, P.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6816-6820.
- Belcourt, M.F. and Farabaugh, P.J. (1990) *Cell*, **62**, 339-352.
- Wada, K., Wada, Y., Ishibashi, F., Gojobori, T. and Ikemura, T. (1992) *Nucleic Acids Res.*, **20**, 2111-2118.
- Wilson, W., Braddock, M., Adams, S.E., Rathjen, P.D., Kingsman, S.M. and Kingsman, A.J. (1988) *Cell*, **54**, 1159-1169.
- Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E. (1988) *Nature*, **331**, 280-283.
- Reil, H., Kollmus, H., Weidle, U.H. and Hauser, H. (1993) *J. Virol.*, **67**, 5579-5584.
- Parkin, N.T., Chamorro, M. and Varmus, H.E. (1992) *J. Virol.*, **66**, 5147-5151.
- Skuzeski, J.M., Nichols, L.M., Gesteland, R.F. and Atkins, J.F. (1991) *J. Mol. Biol.*, **218**, 365-373.
- Berteaux, V., Rousset, J.P. and Cassan, M. (1991) *Biochimie*, **73**, 1291-1293.
- Kuchino, Y., Beier, H., Akita, N. and Nishimura, S. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2668-2672.