

**THESE de DOCTORAT de l'UNIVERSITE PARIS 6**

Spécialité :

Génétique Cellulaire et Moléculaire

**présentée par Guillaume Stahl**

pour obtenir le grade de DOCTEUR de l'UNIVERSITE PARIS 6

Sujet de la thèse

**Recherche de facteurs modulant le recodage chez  
*Saccharomyces cerevisiae***

soutenue le 9 juillet 1998

devant le jury composé de

Pr. Marguerite PICARD (présidente)

Dr. Olivier JEAN-JEAN (rapporteur)

Dr. Mathias SPRINGER (rapporteur)

Pr. François LACROUTE (examineur)

Pr. Jean-Pierre ROUSSET (examineur)

## Remerciements à

ADDOC, Alain, Angelos, Chantou, Christophe, le CIES de Versailles, Enzo, Eric, Evelyne, François, Hongmei, Isabelle, Jean-Luc, Jean-Pierre, Laure, Loul, Marco, Marguerite, Mathias, Mich', Michel, Michèle, Monique, Olivier, Phil, Philippe, Pierre, Sam, Sophie, Véronique,...

Il faut reprendre à César tout ce qui ne lui appartient pas  
André Breton & Paul Éluard

---

à Anne et Riquet

# Table des matières

|   |    |
|---|----|
| Remerciements à.....  | 3  |
| Table des matières.....   | 4  |
| Abbréviations .....   | 8  |
| Nombre de références bibliographiques .....                     | 9  |
| Liste des figures.....  | 9  |
| Conventions.....  | 9  |
| Introduction .....  | 11 |
| A. Les erreurs de traduction.....                               | 11 |
| B. Le recodage.....   | 15 |
| B.1. Décalage de phase en -1 .....                              | 19 |
| B.1.a. L'heptanucléotide, séquence glissante .....              | 19 |
| B.1.b. La structure secondaire .....                            | 24 |
| B.1.c. La distance heptanucléotide/structure<br>secondaire..... | 27 |
| B.1.d. Un codon stop en 3' de la séquence<br>glissante.....     | 27 |
| B.1.e. La butée du ribosome .....                               | 29 |
| B.1.f. Modèles temporels .....                                  | 31 |
| B.1.g. Modèles mécanistiques .....                              | 33 |
| B.2. Décalage de phase en +1 .....                              | 41 |
| B.2.a. Famille Ty1,2 et 4.....                                  | 41 |
| B.2.b. Ty3.....   | 42 |
| B.2.c. RF2.....   | 44 |
| Le codon glissant.....  | 44 |
| Le contexte du codon Stop .....                                 | 46 |
| Une séquence Shine-Dalgarno.....                                | 46 |
| B.2.d. Antizyme.....  | 46 |
| B.2.e. EST3.....  | 48 |
| B.3. Translecture d'un codon stop .....                         | 50 |
| B.3.a. Contexte nucléotidique et terminaison .....              | 50 |
| L'environnement naturel des codons stop.                        | 51 |
| L'influence de ce contexte sur la<br>terminaison.....           | 51 |
| B.3.b. TMV.....   | 52 |
| B.3.c. Autres virus de plantes .....                            | 54 |
| B.3.d. Sindbis .....  | 54 |
| B.3.e. MoMuLV .....   | 55 |
| B.3.f. Gènes cellulaires.....                                   | 56 |

|  |    |
|--|----|
| E. coli : gènes nécessaires à l'assemblage<br>des pili .....   | 56 |
| Kelch.....   | 57 |
| Headcase .....   | 57 |
| Sélénocystéine.....  | 57 |
| B.3.g. Les codons ochre .....  | 58 |
| B.4. Saut de ribosome.....   | 60 |
| Chapitre I. Construction d'un outil de mesure du recodage.....   | 63 |
| I.A. État des lieux en 1993.....   | 63 |
| I.B. Le système de mesure .....  | 64 |
| I.C. Validation du système rapporteur .....  | 64 |
| I.D. Étude de cibles de recodage hétérologues chez <i>S. cerevisiæ</i> .....   | 65 |
| I.E. ARTICLE I .....   | 66 |
| I.F. Décalage de phase en +1 chez <i>S. cerevisiæ</i> .....  | 67 |
| I.G. Levures et cibles issues d'eucaryotes supérieurs.....   | 68 |
| I.G.1. TMV.....  | 68 |
| I.G.2. VIH-1 .....   | 69 |
| I.H. Le recodage chez les procaryotes.....   | 69 |
| Chapitre II. Étude de l'influence de la stabilité de la structure<br>secondaire du VIH-1 sur le taux de décalage de phase en -1..... | 73 |
| II.A. État des lieux en 1995.....  | 73 |
| II.B. ARTICLE II .....   | 75 |
| II.C. Existence de la tige-boucle en aval du site de glissement<br>de VIH-1.....   | 76 |
| II.D. Mutant plus stable que le sauvage.....   | 77 |
| II.E. Comparaison entre levure et cellules de mammifère,<br>facteurs trans.....  | 79 |
| Chapitre III. Facteurs Trans modulant le recodage .....  | 81 |
| III.A. Facteurs trans identifiés du recodage .....   | 81 |
| III.A.1. Sélénocystéine .....  | 82 |
| III.A.2. Gènes cellulaires.....  | 82 |
| III.A.2.a. protéines ribosomales .....   | 82 |
| III.A.2.b. Facteurs de la traduction .....   | 83 |
| Initiation .....   | 83 |
| Élongation.....  | 84 |
| Terminaison.....   | 85 |
| Les facteurs NMD .....   | 85 |
| III.A.2.c. ARNr.....   | 86 |
| 23S.....   | 86 |

|  |     |
|--|-----|
| 16S.....   | 87  |
| 5S.....  | 87  |
| III.A.2.d. ARNt .....  | 88  |
| Décalage de phase +1.....  | 88  |
| Décalage de phase -1.....  | 89  |
| III.A.3. Facteurs de l'ARNm agissant en trans .....                      | 90  |
| III.A.4. Peptide en croissance.....                                      | 91  |
| III.A.5. Polyamines .....  | 91  |
| III.A.6. Substances antibiotiques .....                                  | 92  |
| III.A.7. Conditions de croissance .....                                  | 93  |
| III.A.7.a. Phases de croissance .....                                    | 94  |
| III.A.7.b. Carence en acides aminés .....                                | 95  |
| III.A.7.c. Source de carbone.....  | 95  |
| III.A.7.d. Sénescence.....   | 96  |
| III.A.8. Ferritine.....  | 96  |
| III.B. Crible de facteurs modulant la translecture du TMV ...            | 98  |
| III.B.1. Principe.....   | 99  |
| III.B.2. Gène rapporteur .....   | 99  |
| III.B.3. Construction de la souche de mutagenèse .....                   | 100 |
| III.B.4. Mutagenèse .....  | 102 |
| III.B.5. Crible primaire.....  | 102 |
| III.B.6. Crible secondaire .....   | 102 |
| III.B.7. Phénotypes secondaires .....                                    | 103 |
| III.B.8. Résultats .....   | 104 |
| III.B.9. Faisabilité du crible.....                                      | 108 |
| Conclusion.....  | 111 |
| Système de mesure, les éléments cis .....                                | 111 |
| Les facteurs trans, l'avenir ?.....                                      | 114 |
| (i) Recodage, infidélité ou fidélité ?.....                              | 115 |
| (ii) Facteurs trans comme révélateurs des mécanismes<br>du recodage..... | 116 |
| (iii) Le recodage comme modèle d'étude de la<br>traduction.....          | 117 |
| Annexes .....  | 121 |
| Matériels et Méthodes.....   | 121 |
| Escherichia coli.....  | 121 |
| Saccharomyces cerevisiae .....   | 121 |
| Techniques.....  | 121 |
| Tests b-galactosidase sur boîte.....                                     | 121 |
| sur filtre .....   | 121 |

|  |     |
|--|-----|
| en Top-Agarose .....                                   | 121 |
| Souches .....  | 122 |
| Tableau récapitulatif des résultats de mutagenèse..... | 123 |
| Code génétique .....                                   | 124 |
| Références bibliographiques.....                       | 125 |
| Résumé en anglais - Abstract .....                     | 147 |
| Résumé.....  | 149 |

# Abbréviations

|                      |  |
|----------------------|--|
| aa                   | : acide aminé  |
| ADN                  | : Acide désoxy-ribonucléotidique                       |
| ADNr                 | : ADN codant pour l'ARN ribosomique                    |
| ARN                  | : Acide ribonucléotidique                              |
| ARNm                 | : ARN messenger  |
| ARNr                 | : ARN ribosomique                                      |
| ARNt                 | : ARN de transfert                                     |
| ATP                  | : adénosine tri-phosphate                              |
| BLV                  | : Bovine Leukaemia Virus                               |
| BYDV                 | : Barley Yellow Dwarf Virus                            |
| Da                   | : dalton   |
| <i>E. coli</i>       | : <i>Escherichia coli</i>                              |
| eRF                  | : facteur de terminaison eucaryote                     |
| FIV                  | : Feline Immunodeficiency Virus                        |
| HIV                  | : Human Immunodeficiency Virus (VIH)                   |
| HTLV-2               | : Human T-cell Leukemia Virus type II                  |
| IBV                  | : Infectious Bronchitis Virus                          |
| kDa                  | : kilo-dalton  |
| kb                   | : kilo-base  |
| kpb                  | : kilo-paire de base                                   |
| MMTV                 | : Mouse Mammary Tumor Virus                            |
| MoMuLV               | : Virus Moloney de la leucémie murine                  |
| NMD                  | : Non-sense Mediated mRNA decay                        |
| nt                   | : nucléotide   |
| ODCase               | : Ornithine Décarboxylase                              |
| ORF                  | : Open Reading Frame - phase ouverte de lecture        |
| PAP                  | : Pokeweed Antiviral protein                           |
| pb                   | : paire de base  |
| R                    | : purine (A ou G)                                      |
| RF                   | : Release Factor (facteur de terminaison)              |
| RMN                  | : Résonance Magnétique Nucléaire                       |
| RSV                  | : Rous Sarcoma Virus                                   |
| RRL                  | : Rabbit Reticulocyte Lysate                           |
| <i>S. cerevisiae</i> | : <i>Sacharomyces cerevisiae</i>                       |
| ScV                  | : <i>S. cerevisiae</i> virus (L-A, ou L-BC)            |
| SD                   | : Shine-Dalgarno                                       |
| SRV-1                | : Simian Retrovirus 1                                  |
| ts                   | : thermo-sensible                                      |
| TMV                  | : Tobacco Mosaic Virus (virus de la mosaïque du tabac) |
| UV                   | : ultra-violets  |
| VIH                  | : Virus de l'Immunodéficience Humaine (HIV)            |
| Y                    | : pyrimidine (C, T ou U)                               |
| X-Gal                | : 5-bromo-4-chloro-3-indolyl- -D-galactopyranoside     |

# Nombre de références bibliographiques

220

## Liste des figures

01. page 10 : Représentation schématique du Proofreading
02. page 20 : Génome de l'IBV
03. page 22 : Décalage simultané des ARNt sur un heptamère glissant
04. page 22 : Une seule mutation dans un heptamère glissant produit deux changements d'acides aminés au cours du recodage
05. page 25 : Différentes représentations schématiques équivalentes d'un pseudo-nœud de type H
06. page 32 : Décalage PRÉ-peptidyl transfert
07. page 32 : Décalage POST-peptidyl transfert
08. page 34 : Essais de translocation provoquant un décalage de phase en -1
09. page 36 : Déroulement/ré-enroulement de la structure secondaire provoquant un décalage de phase en -1
10. page 37 : Torsion Resistance Model
11. page 40 : Décalage de phase en +1 du Ty1 et du Ty3
12. page 45 : Autorégulation de RF-2 par décalage de phase en +1
13. page 47 : Boucle de régulation négative de l'antizyme par décalage de phase en +1
14. page 101 : Construction des souches contenant le gène rapporteur *lacZ* derrière une cible de recodage
15. page 101 : Souche YTX8b et disruption du locus ASN1 par recombinaison homologue
16. page 105 : Crible de facteurs trans sur la translecture du TMV
17. page 110 : Récapitulatif des mécanismes de recodage

## Conventions

La séquence des anticodons sera écrite, comme indiqué, dans le sens 3' vers 5' pour faciliter la lecture des appariements avec les codons.

Les phases de lecture seront indiquées avec des espaces entre les codons. Sans indication, c'est la phase initiale (phase 0) qui sera prise en compte.

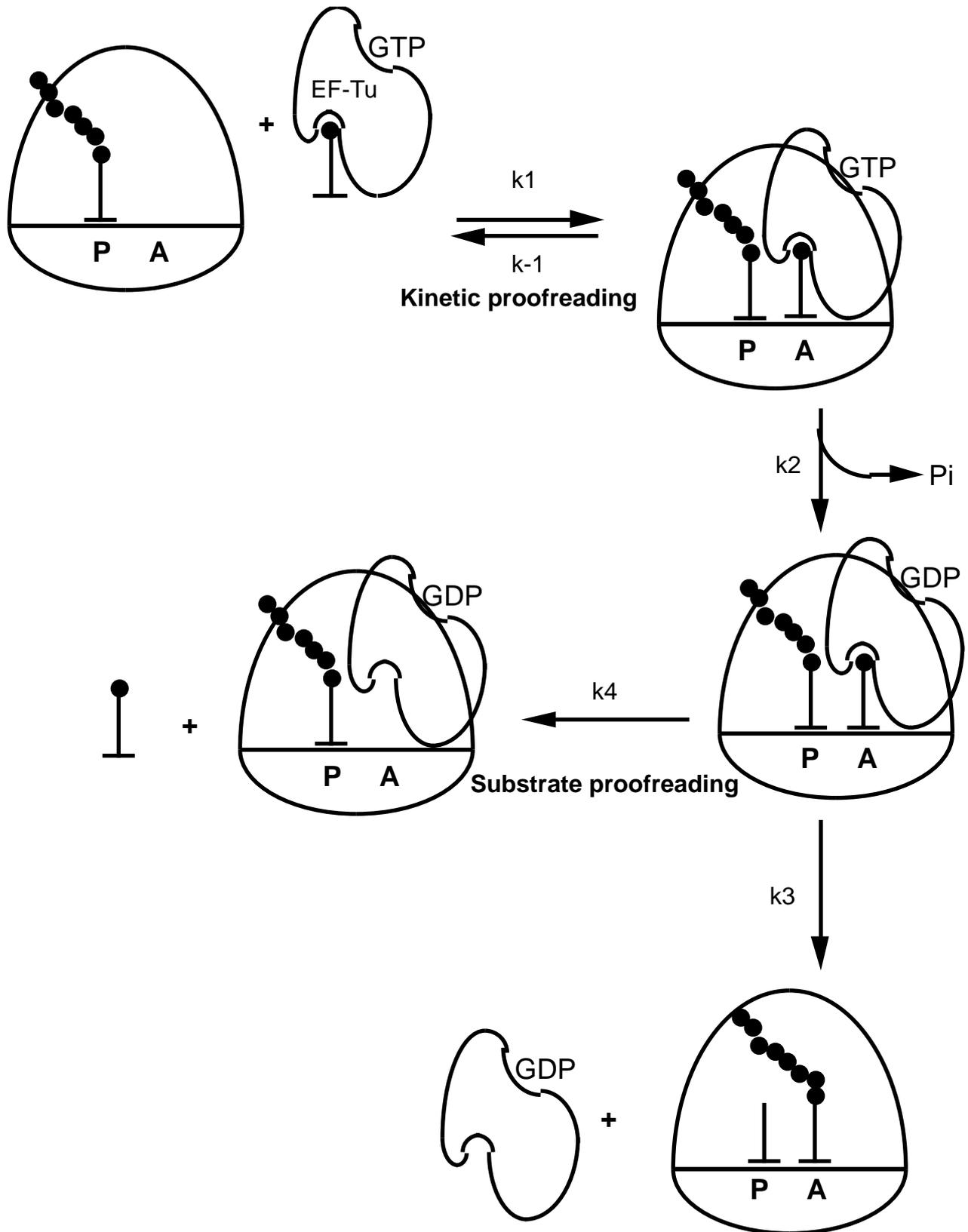


Figure 1 : **Représentation schématique du Proofreading**

$k_1$  ARNt cognat <  $k_1$  ARNt non cognat : le complexe ternaire contenant un ARNt non cognat a tendance à se dissocier avant hydrolyse du GTP ( $k_2$ )  
 $k_4$  ARNt cognat <  $k_4$  ARNt non cognat : l'ARNt non cognat à tendance à se dissocier avant que ne se forme la liaison peptidique ( $k_3$ )  
 Les constantes de vitesse  $k_1$ ,  $k_2$  et  $k_3$  sont identiques entre un ARNt cognat et non cognat

# Introduction

## A. Les erreurs de traduction

Comme le suggèrent les théories actuelles, le monde vivant aurait basculé d'une composition essentiellement ARN à une composition mixte où le vivant dépend de la combinaison acides nucléiques/protéines. Cette étape a nécessité l'apparition d'une machine moléculaire, le ribosome, permettant le transfert, par traduction, de la plupart des capacités catalytiques des ARN vers les protéines, laissant aux premiers les capacités de support génétique. L'intégrité des protéines, et donc des fonctions cellulaires, nécessite une étape de traduction du code beaucoup plus fiable que ne le permettent les simples réactions physico-chimiques en jeu, notamment la reconnaissance codon-anticodon (Kurland 1992). Il a été suggéré que la fidélité du ribosome lui est conférée par un système de relecture-correction cinétique ("kinetic proofreading" : Ninio 1974). La présence d'un ARNt cognat (ARNt le plus apte à décoder un codon donné) est favorisée par des interactions non spécifiques du ribosome avec tout ARNt : la vitesse à laquelle se fait cette stabilisation de l'interaction au site A favorisera un ARNt restant plus longtemps (ARNt cognat) que les autres (ARNt non cognats, cf. la constante cinétique  $k_1$  dans la Fig. 1). Ce système est optimisé par ce que l'on a appelé le "substrate proofreading" (Hopfield 1974; Ninio 1975), qui constitue un délai supplémentaire permettant le même type de sélection non spécifique que le "kinetic proofreading". Ce délai serait fourni par une réaction indispensable à la poursuite de la traduction : l'hydrolyse du GTP par le facteur d'élongation EF-Tu/EF-1A, qui a amené l'ARNt au ribosome sous la forme d'un complexe ternaire. Cette étape supplémentaire permet de favoriser un ARNt cognat en augmentant le délai entre l'arrivée de l'ARNt au site de décodage A (accepteur), et la suite de l'élongation (transfert du peptidyl), et donc le temps où les ARNt non cognats peuvent se dissocier à cause d'une constante cinétique  $k_4$  défavorable (Fig. 1, et pour revue Thompson 1988; Kurland 1992). Ces deux étapes permettent d'amplifier la fidélité de lecture (ARNt cognat *versus* ARNt non cognat) aux dépens de la vitesse de traduction. Toute modification, même mineure, des caractéristiques cinétiques de l'appareil traductionnel ont des conséquences dramatiques sur la fidélité de traduction.

Comme on l'a vu, le complexe ternaire peut se dissocier après hydrolyse du GTP, avant transfert du peptidyl, et ceci est plus fréquent pour des

ARNt non cognats. Lors d'erreurs à répétition, il y a donc création de cycles futiles de consommation d'énergie sous forme de GTP. Ces systèmes de relecture-correction, et l'importance quantitative du processus (36% du poids sec de *E. coli* est dédié à la traduction, Goodsell 1991) font de la traduction le système cellulaire le plus coûteux en énergie. La machinerie traductionnelle a évolué en tendant vers un équilibre satisfaisant pour la vie cellulaire dans un milieu donné, entre les deux grandeurs contradictoires de vitesse et de fidélité de traduction. Des isolats naturels ont d'ailleurs des temps de génération, ainsi que des fidélités de traduction très variables (Kurland 1992). La résolution de ce conflit nous donne aujourd'hui à observer une traduction qui semble relativement fidèle, puisqu'on estime à environ  $5 \times 10^{-4}$  les erreurs faux sens (Parker 1989; ces chiffres, et tous les suivants sont donnés pour des souches sauvages de laboratoire de *E. coli*). Malgré ce taux d'erreur relativement faible, on en a déduit que la traduction d'une phase ouverte de lecture de 500 acides aminés produit 25% de protéines erronées (mais pas forcément inactives), ou encore qu'une protéine de grande taille telle la  $\beta$ -galactosidase, ou l'ADN polymérase, contient statistiquement au moins un acide aminé erroné. Pas un seul ribosome (2,5 MDa) de la cellule n'est strictement identique à un autre (Kurland 1992). Ces exemples montrent l'importance du contrôle de la fidélité de la traduction, et ses conséquences potentielles en cascade, dont l'issue fatale a été appelée la catastrophe d'erreurs ("error catastroph"), a été suspectée d'être une des causes de la sénescence (Orgel 1970).

Les erreurs faux-sens, dont l'impact peut être assez limité vu la tolérance des protéines aux changements d'acides aminés sur la majeure partie de leur longueur, constituent environ les deux tiers des erreurs observées lors de la traduction (Kurland 1992). À celles-ci s'ajoutent les erreurs de processivité, dont on ne connaît pratiquement pas le contrôle, et qui provoquent une dissociation précoce du peptidyl-ARNt du ribosome. Cette dissociation peut être provoquée par différents événements :

- un mécanisme de "ribosome editing", qui provoque la dissociation spontanée du peptidyl-ARNt du ribosome ("drop-off"), plus fréquente s'il y a modification du taux des erreurs faux-sens; ceci augmente le contrôle dû au mécanisme de proofreading,
- la terminaison précoce au niveau d'un codon sens. Cette erreur est très rare, de l'ordre de  $10^{-5}$  par codon (Jorgensen *et al.* 1993),

- des erreurs dans le maintien du cadre de lecture; en effet, le décalage dans une des deux autres phase de lecture mène fatalement à la rencontre d'un codon stop au bout de quelques dizaines de codons. Cette erreur a été estimée à un maximum de  $3 \times 10^{-5}$  par codon (Kurland 1992). Les conséquences beaucoup plus néfastes de ce type d'erreur (terminaison précoce) par rapport aux erreurs faux-sens ont sans doute contribué à la sélection de cette fidélité de processivité. Peu d'approches méthodologiques à ce jour ont permis d'aborder l'étude des mécanismes de contrôle de cette fidélité. L'une de ces approches serait d'étudier plus spécifiquement des contextes extrêmement défavorables à la processivité normale du ribosome, comme nous le verrons plus loin. Une autre erreur de traduction dont nous parlerons peut également être considérée comme une erreur de processivité. Il s'agit de la mauvaise reconnaissance d'un codon de terminaison de la traduction, et donc de la poursuite de l'élongation au delà d'un codon stop. Ce type d'erreur peut être sans conséquences (par exemple dans le cas où des codons stops multiples se suivent à la fin du messager), ou délétères pour la cellule, dans le cas d'extensions C-terminales ayant un effet dominant négatif. On ne sait pas actuellement si le système de proofreading des erreurs faux-sens médié par EF-1A contrôle également les erreurs de processivité; mais le décalage du cadre de lecture est augmenté chez des mutants EF-Tu (Hughes *et al.* 1987) ou EF-1A (Sandbaken *et al.* 1988). On verra plus loin que les trois processus dont nous avons parlé (faux-sens, processivité, terminaison) semblent intimement liés.

Malgré toutes ces erreurs, et ce sans doute car les erreurs processives restent minoritaires, les conséquences des erreurs de traduction sont limitées. Mais on a pu mettre en évidence au début des années soixante-dix des phénomènes traductionnels extraordinaires qui impliquent des erreurs de traduction à un haut niveau et de manière reproductible : les erreurs se produisent toujours au même endroit sur un messager donné, et le type d'erreur est identique à chaque fois. Ainsi la traduction peut être modifiée au cas par cas pour certains ARNm : cadre de lecture modifié, codons stop non reconnus, ou même déplacement du messager sur une grande distance par rapport au ribosome. Ces événements semblent impliquer des règles différentes pour ces ARNm, et donc un codage spécifique. On a alors parlé de "recoding" en anglais, que l'on pourrait traduire par recodage. Au long de cette introduction, je vais faire un état des lieux de ce domaine expérimental. J'aborderai les questions qu'il suscite, et auxquelles il répond. Enfin j'exposerai les questions que j'ai moi-même abordé au cours de ma thèse.

Je définirai tout d'abord les notions utilisées telles que le recodage, et les différents types d'infidélités traductionnelles que ce mot englobe; en essayant de faire ressortir les principes et les règles qui commencent à émerger. Ces rappels vont nous permettre d'aborder les différentes étapes de ma thèse que sont :

- Chapitre I : la construction d'un système de mesure des erreurs traductionnelles chez la levure *Saccharomyces cerevisiæ*, et son application à des cibles provenant d'eucaryotes supérieurs, pour déterminer si cet organisme peut être utilisé comme modèle d'étude du recodage,

- Chapitre II : l'étude de l'influence d'une structure secondaire de l'ARNm sur un décalage de phase issu du génome du Virus de l'Immunodéficience Humaine (VIH-1) en parallèle dans des cellules de mammifères en culture et chez *S. cerevisiæ*,

- Chapitre III : le projet que je développe pour identifier des facteurs cellulaires modulant différents types d'erreurs traductionnelles.

## B. Le recodage

Les erreurs de traduction programmées, ou infidélités de traduction, ont été baptisées "recoding" en anglais (Gesteland *et al.* 1992). Le recodage peut être défini de manière simple comme tout ce qui n'est pas "codage" classique de l'information génétique : un codon (3 nucléotides), qui est traduit en un acide aminé précis dans la chaîne peptidique en élongation; le prochain triplet nucléotidique code le prochain acide aminé, et ce jusqu'à la rencontre d'un signal de fin de traduction (codon stop: UAA, UAG, UGA). En effet, certains messagers codent non seulement des acides aminés de cette façon classique, mais possèdent en plus des signaux spécifiques en *cis* (séquence et/ou structures), qui modifient le codage normal de la traduction en provoquant une "erreur programmée", un recodage. Comment le messenger oblige-t-il l'appareil traductionnel à augmenter l'efficacité et l'étendue de tels événements ? Existe-t-il un interprète particulier de ces programmes ? Quels que soient les mécanismes envisagés, la compréhension de ces systèmes alternatifs de la traduction permettra une meilleure compréhension du contrôle des erreurs traductionnelles stochastiques, naturelles, et même une meilleure compréhension de la traduction en général.

Non seulement, la compréhension de ce nouveau code est d'importance pour la compréhension des mécanismes moléculaires *per se*, mais elle est aussi d'actualité pour la modélisation informatique de nos connaissances. L'accumulation de séquences (plus de 70 génomes en cours de séquençage), pose le problème du traitement automatisé de la signification de ces séquences. Si on ne sait pas définir une séquence codante, ou si le code que l'on y définit est incomplet, on néglige ou on biaise une partie des informations. En effet, toute séquence soumise au décalage de phase ou à la lecture d'un codon stop apparaîtra dans les banques de données avec une "traduction" erronée. Les gènes nécessitant un événement de recodage ont un message génétique qui, par définition, n'est pas en phase. En fait ces séquences seront répertoriées le plus souvent comme séquences non-codantes ou pseudogènes, au vu de leurs tailles apparentes, comme cela a été le cas du gène *EST3* (Ever Shorter Telomeres), identifié génétiquement. D'après les critères du "Yeast Genome Project" (Dujon *et al.* 1994), ce gène ne contenant pas de phase ouverte de lecture de taille supérieure à 100 codons, ne coderait pas pour une protéine. Vu l'importance d'une matrice ARN lors de la télomérisation, il a été suggéré que ce gène codait pour un ARN.

Récemment, un travail précis a montré qu'il s'agissait en fait d'une protéine exprimée à partir de deux phases de lecture ouvertes de moins de 100 codons, séparées par un décalage programmé de phase de lecture en +1 (Morris *et al.* 1997). Les gènes soumis à recodage sont ignorés actuellement lors des analyses informatiques du génome définissant les phases potentiellement codantes (ORF pour "Open Reading Frame", phase ouverte de lecture), et donc les séquences qui vont être analysées lors de l'analyse fonctionnelle du génome (disruption en masse des ORFs, microplaques contenant des oligonucléotides de référence pour toutes les ORFs de levure,...). Chez *S. cerevisiae*, sans parler de tous les éléments Ty (voir plus loin), des travaux en cours au laboratoire (Olivier Namy, communication personnelle) ont permis d'évaluer que de nombreuses phases de lecture interrompues ont été ignorées de la sorte lors du traitement des données du séquençage. Des expériences sont en cours pour déterminer si ces phases de lecture interrompues sont exprimées, s'il y a recodage lors de leur traduction, et à quel taux. S'il y a recodage, cette information n'a pas été référencée automatiquement, et les données disponibles pour ces séquences dans les banques de données sont incomplètes ou fausses.

Une autre conséquence de la méconnaissance du recodage provient de l'épuration des erreurs de séquençage contenues dans les banques de données. Afin de détecter des erreurs de séquençage avec insertion ou délétion d'un nucléotide, des programmes informatiques détectent des décalages de phases entre deux parties codantes d'un même gène (grâce à une analyse de type linguistique par exemple Fichant *et al.* 1995). Ce programme corrige alors automatiquement ces supposées erreurs. Or le recodage (processif) et le résultat d'une insertion/délétion due à une erreur de séquençage sont des objets informatiquement identiques. Comme on le voit, la génomique ne peut être envisagée comme approche globale sûre que si tous les mécanismes d'expression génétique sont connus, spécialement ceux de la traduction, utilisés dans les toutes premières étapes de la modélisation du génome : découpage des séquences en phases codantes, traduction automatique des phases codantes en séquences protéiques, recherche de fonction par comparaison de séquences protéiques, ... C'est aussi dans ce cadre que nous nous efforçons, par la compréhension des informations contenues sur les messagers, et par l'identification des acteurs de la traduction, de comprendre les mécanismes et le rôle du recodage.

Il existe une troisième raison, plus appliquée, d'étudier le recodage : il s'agit d'une cible anti-virale potentielle. En effet, à ce jour, la majorité des événements de recodage caractérisés l'ont été chez des virus, ou des éléments cellulaires apparentés, comme les transposons. Cette prépondérance du recodage chez les virus peut être dû à une meilleure connaissance de leur mode d'expression, mais peut également être dû à certaines spécificités des virus :

- la protéine de fusion gag-pol est constituée d'une partie structurale de la nucléocapside (gag) intégrée automatiquement au virus, qui entraîne la partie enzymatique (pol), qui est amarrée au virion de cette manière (Ribas *et al.* 1998),

- les virus à réplication à ARN sont en général interdits d'épissage, puisque leurs ARNs servent en même temps de génome et de message; ceci ne concerne qu'une partie des virus, puisque d'autres stratégies permettent d'éviter l'encapsidation de molécules épissées,

- ce sont des génomes compacts; or le recodage permet de bénéficier de deux phases de lecture sur un seul messenger, et ce en contrôlant le rapport précis de la quantité relative des deux produits synthétisés. Les virus semblent être très sensibles aux variations du taux de recodage, une modification d'un facteur deux peut entraîner un défaut dans la propagation du virus (Felsenstein *et al.* 1988; Park *et al.* 1991; Dinman *et al.* 1992; Kawakami *et al.* 1993; Farabaugh 1995; Hung *et al.* 1998).

Le recodage est donc une cible anti-virale, si l'on en trouve des modulateurs spécifiques (Dinman *et al.* 1998). Il faut pouvoir estimer l'effet de ces modulateurs sur les cellules eucaryotes pour une utilisation thérapeutique :

- si les cellules eucaryotes ne dépendent pas de l'utilisation du recodage pour l'expression de leurs fonctions vitales, on peut bloquer la réplication virale en faisant varier spécifiquement le taux de recodage, sans affecter la cellule hôte,

- à l'inverse, si le recodage est également répandu chez les eucaryotes, et certains exemples récents sembleraient confirmer cette idée (Xue *et al.* 1993; Matsufuji *et al.* 1995; Morris *et al.* 1997; Steneberg *et al.* 1998), il faut pouvoir mesurer l'impact précis des "anti-recodage" sur la cellule, et éventuellement identifier des facteurs spécifiques des systèmes viraux.

Ces considérations montrent l'importance des études du recodage dans plusieurs champs de la biologie : la génétique moléculaire, la modélisation informatique et la génomique, la lutte contre les pathogènes.

### **Définition du recodage**

Le recodage a été défini par John Atkins et Raymond Gesteland (Gesteland *et al.* 1992; Gesteland *et al.* 1996). Le recodage est un phénomène :

- (i) dynamique (ou encore temporaire), c'est-à-dire qu'il ne concerne qu'un seul ribosome à la fois,
- (ii) en compétition avec le codage; les deux phénomènes sont exclusifs, et si l'un devient moins efficace, l'autre voit sa fréquence augmenter,
- (iii) dépendant d'un site d'action spécifique sur le messager,
- (iv) dont l'amplitude est modulée par un stimulateur, qui est généralement une structure secondaire de l'ARNm.

Nous proposons d'ajouter à cette définition, pour éviter un débat sémantique sur la notion de reprogrammation qui peut parfois se confondre avec une erreur à un taux élevé de l'appareil traductionnel, qu'il n'y a recodage que si le produit issu d'un tel événement traductionnel possède une fonction biologique.

Dans cette introduction, je présenterai les facteurs présents sur l'ARNm qui induisent le recodage (pour revue, voir Farabaugh 1996). J'évoquerais plus loin, au troisième chapitre, les facteurs *trans* connus du recodage.

## **B.1. Décalage de phase en -1**

Le décalage programmé du ribosome d'un nucléotide en amont est apparemment le mécanisme de recodage le plus conservé et le plus répandu dans la nature. Ce décalage en -1 se retrouve en général à la jonction entre les gènes *gag* et *pol*, ou leurs équivalents, chez la plupart des rétrovirus, dans des virus de plantes, de levure, des éléments transposables bactériens ou eucaryotes, (pour revue Brierley 1995). Le gène *pol* code les fonctions catalytiques du virus, dont la polymérase nécessaire au cycle du virion. Son expression est dépendante de la phase de lecture de *gag*, bien que *pol* soit codé dans la phase -1 par rapport à *gag*. Certains ribosomes vont changer de cadre de lecture à la jonction entre les deux phases, et donc traduire sous forme de protéine de fusion le polypeptide Gag-pol (Fig. 2). Ce phénomène a été décrit pour la première fois pour le RSV (Jacks *et al.* 1985). Je présenterai ici l'exemple le plus étudié à ce jour, et le plus canonique : le coronavirus aviaire de la bronchite infectieuse (IBV).

Le génome de l'IBV est un ARN simple-brin de 27,6 kb, de polarité plus, et est donc directement traduit dans les cellules infectées. En 5' de ce messenger se trouvent deux très grandes ORFs *1a* et *1b*, capables de coder deux protéines de 440 kDa et 300 kDa respectivement. *1b*, qui code une ARN polymérase ARN-dépendante, est chevauchante par rapport à *1a* sur 42 nucléotides, dans la phase en -1. Cet arrangement est très semblable à celui identifié chez le RSV, et il a donc été suspecté une régulation traductionnelle de type décalage de phase en -1 auparavant seulement identifiée chez les rétrovirus ou les rétrotransposons. En effet, la jonction 1a/1b clonée dans un gène rapporteur, en respectant les deux phases décalées, provoque *in vitro* un décalage de phase avec une efficacité d'au moins 30% (Brierley *et al.* 1987).

### **B.1.a. L'heptanucléotide, séquence glissante**

La jonction 1a/1b contient un motif décrit comme étant une "séquence glissante" chez les rétrovirus : un heptanucléotide constitué de deux triplets homopolymères de type ..X XXY YYZ. Les deux ARNt présents aux sites P et A en train de décoder les codons XXY et YYZ "dérapent" le long de l'ARNm dans certaines conditions. La propriété de "wobble" des ARNt

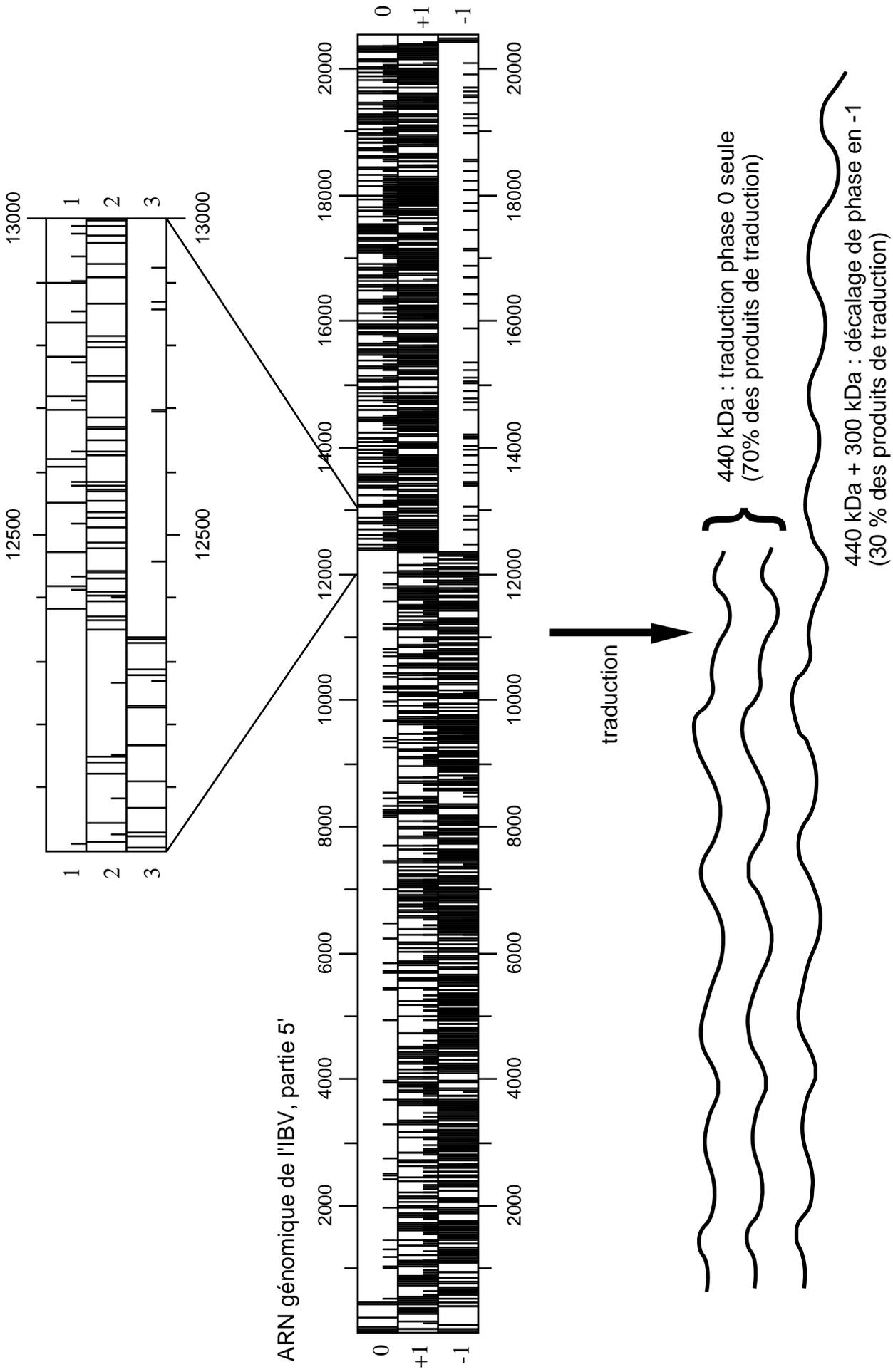


Figure 2 : **Génome de l'IBV**  
 Les produits de traduction obtenus sont schématisés en dessous

permet au ribosome de se recalculer sur les séquences homopolymériques, mais en ayant effectué un décalage d'un nucléotide en 5' : XXX YYY Z..; le ribosome est alors capable de traduire la phase *1b* (Fig. 3). Cette propriété de décalage en tandem des ARNt lors de la traduction d'une "séquence glissante" a pour l'instant toujours été confirmée, par des expériences de mutagenèse dirigée sur ces séquences, ou par séquençage des protéines de fusion. Ce modèle "*donne la meilleure explication de l'événement de décalage de phase jusqu'à présent*" (Hatfield *et al.* 1990). La preuve que l'ARNt décodant la partie 3' de l'heptamère subit un décalage (à l'exclusion d'un mécanisme de type décodage de deux nucléotides) est qu'une mutation appropriée de la dernière base de l'heptamère (par exemple acA AAU UUA ua -> acA AAU UUU ua) modifie la séquence protéique du dernier acide aminé avant décalage Leu (UUA) -> Phe (UUU), et du premier acide aminé après décalage de phase Ile (Aua) -> Leu (Uua) (Fig. 4). Ceci prouve bien que ce nucléotide est décodé deux fois, par deux ARNt différents (Jacks *et al.* 1988a). Ces expériences délimitent la zone 3' de la séquence glissante. Par ailleurs, une mutation du premier nucléotide de l'heptamère diminue le taux de décalage de phase en -1 presque aussi fortement que les mutations sur le deuxième codon de la séquence glissante. Ceci délimite la région 5' de cette séquence et suggère que les ARNt présents aux sites P et A du ribosome se décalent simultanément (tandem shift en anglais). Par exemple en ce qui concerne l'IBV, la séquence X XXY YYZ a été mutée en différentes séquences, avec les conséquences suivantes sur le taux de décalage de phase (Brierley *et al.* 1989):

| heptamère glissant | ARNt au site P     | ARNt au site A     | décalage de phase (%) |
|--------------------|--------------------|--------------------|-----------------------|
| U UUA AAC (*)      | Leu <sub>AAU</sub> | Asn <sub>UUG</sub> | 25                    |
| (- absente -)      | Leu <sub>AAU</sub> | Asn <sub>UUG</sub> | 0                     |
| <b>C</b> UUA AAC   | Leu <sub>AAU</sub> | Asn <sub>UUG</sub> | 2,5                   |
| U UUA <b>U</b> AC  | Leu <sub>AAU</sub> | Tyr <sub>AUG</sub> | 0,5                   |
| U UUA <b>C</b> AC  | Leu <sub>AAU</sub> | His <sub>GUG</sub> | 0                     |

(\*) séquence sauvage.

Les anticodons sont donnés dans le sens 3' vers 5'

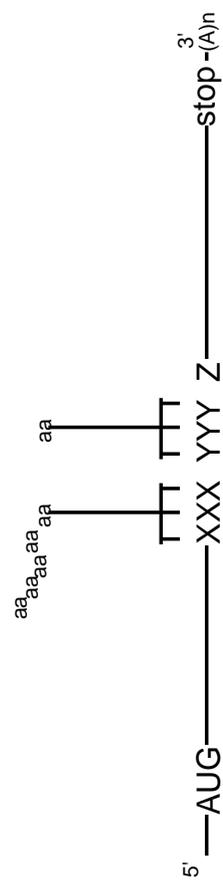
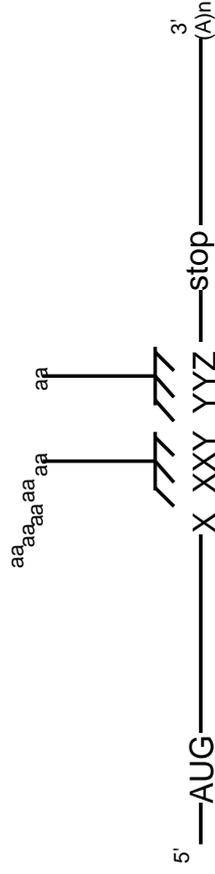
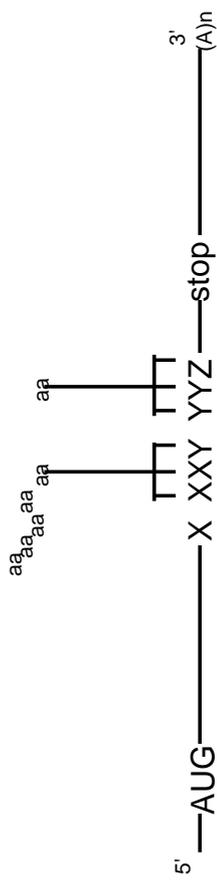
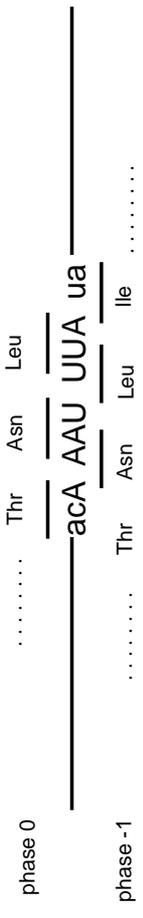


Figure 3 : Décalage simultané des ARNt sur un heptamère glissant

**Séquence sauvage**



**Séquence mutante AAUUUUA -> AAUUUUU**

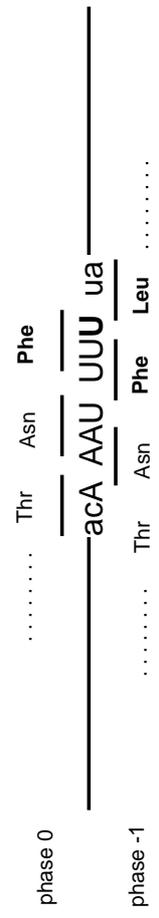


Figure 4 : Une seule mutation dans un heptamère glissant produit deux changements d'acides aminés au cours du recodage

En ce qui concerne le mutant de la première base de l'heptanucléotide U->C, l'ARNt Leu<sub>3'</sub>AAU<sub>5'</sub> ne pourrait que difficilement se réappairier du codon ..C UUA AAC vers le codon situé 1 nucléotide en amont CUU AAA C.. Ne subsisterait alors que la possibilité d'un décalage du seul ARNt situé au site A, d'où sans doute ce faible taux de décalage de phase en -1 (2,5 %). Pour les deux autres mutants ponctuels, le décalage de phase est pratiquement aboli, ce que l'on pourrait expliquer par l'absence totale de complémentarité de l'anticodon (3'AUG<sub>5'</sub> ou 3'GUG<sub>5'</sub>) avec le codon chevauchant en phase -1 (AUA C.. et ACA C.. respectivement). Mais d'une manière générale, les mutants touchant le codon au site A réduisent plus fortement le taux de décalage de cadre que ceux touchant le site P (voir par exemple Jacks *et al.* 1988a). Des mutants plus en 5' n'ont pas d'effet sur le taux de décalage de phase, ce qui suggère que l'ARNt présent au site E du ribosome n'est pas impliqué dans le processus (site E pour "exit", dans lequel se trouverait l'ARNt dé-acylé avant de quitter le ribosome, ce qui rend le site A de nouveau accessible).

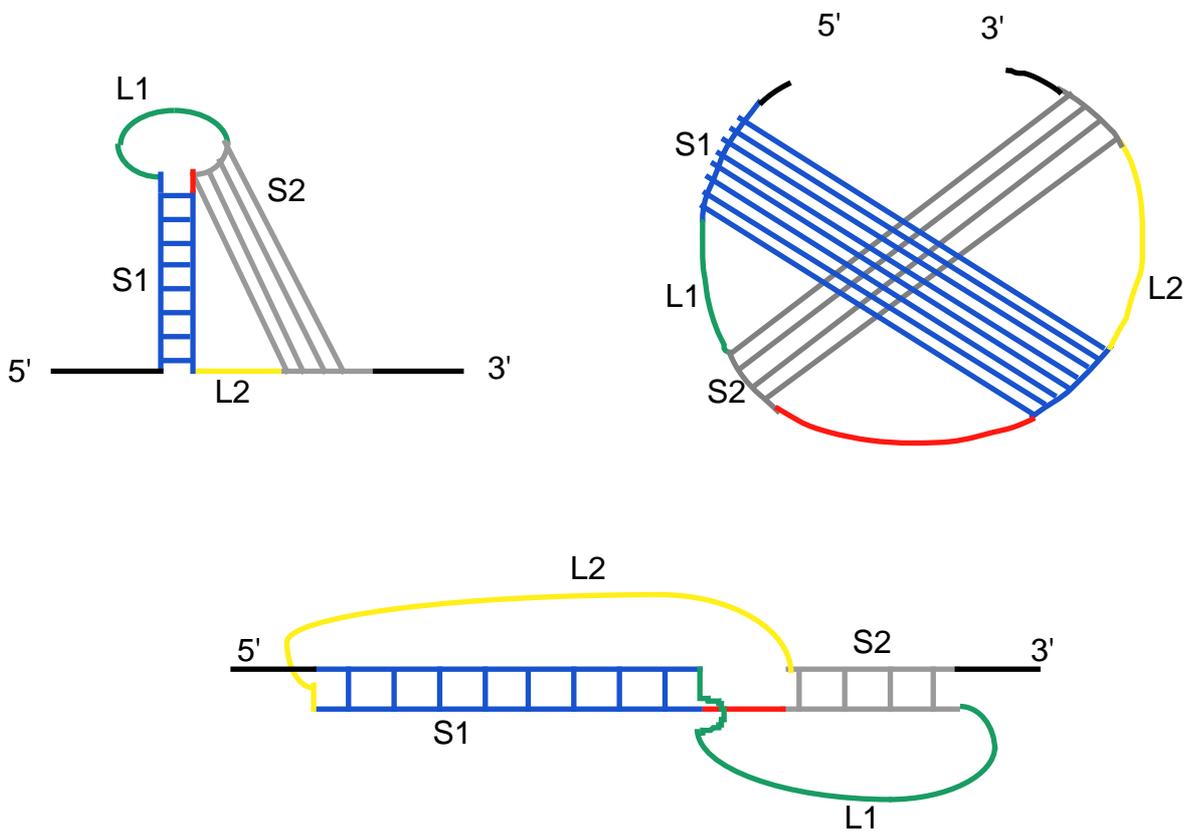
Tous les décalages de phase en -1 étudiés à ce jour chez les eucaryotes correspondent au modèle canonique décrit ci-dessus (Brierley 1995). En ce qui concerne les procaryotes, je donnerais quelques détails à la fin du chapitre I. Parmi 38 sites de décalage trouvés dans la nature, 11 séquences heptamériques différentes ont été identifiées (ten Dam *et al.* 1990); presque tous ces heptamères ont été testés dans le contexte de l'IBV : leur taux de décalage de phase en -1 varie de 16 à 40% (Brierley *et al.* 1992). Des études menées chez la levure ou en lysats de réticulocytes de lapin (RRL) tendent à montrer que tous les nucléotides ne sont pas interchangeable dans les séquences glissantes (Dinman *et al.* 1991; Brierley *et al.* 1992). Au niveau du premier homopolymère, les pyrimidines provoquent le meilleur décalage de phase, suivies par les purines, dans l'ordre UUU>CCC>AAA>GGG. Au niveau du deuxième homopolymère, seules les bases formant deux liaisons hydrogènes sont utilisables pour obtenir des taux significatifs de décalage de phase : AAA ou UUU. La septième base peut être n'importe laquelle, excepté G (Dinman *et al.* 1991; Brierley *et al.* 1992; Dinman *et al.* 1992). Certaines séries, bien que respectant ces règles ne semblent pas permettre un niveau élevé de décalage de cadre (par exemple A AAA au site A chez les mammifères Jacks *et al.* 1988a). Parmi les sites identifiés chez divers virus, les codons présents au site A avant décalage de phase sont AAC, AAU, UUA, UUC, UUU chez les eucaryotes (AAA, AAG chez les procaryotes) (Farabaugh 1996). Les contraintes en bases de ces séquences reflètent certainement la possibilité pour les ARNt de se

désappairier efficacement du premier codon pour se réappairier avec le codon en phase -1, mais on ne peut pas exclure qu'il y ait des ARNt qui soient particulièrement enclins à provoquer le décalage de phase (Jacks *et al.* 1988a), comme c'est le cas pour le décalage de phase en +1 (Curran 1993; Vimaladithan *et al.* 1994). Ceci pourrait expliquer les différences de niveau de décalage de phase entre eucaryotes et procaryotes sur certaines cibles (voir par exemple Tsuchihashi 1991; Brierley *et al.* 1997). Il existe néanmoins des exceptions au paradigme du triplet homopolymérique, comme dans le cas du Mouse Mammary Tumor Virus (MMTV), à la jonction entre pro et pol, le site de glissement est GGA UUU A, qui provoque un décalage de phase avec un taux de 25% (Jacks *et al.* 1987), ou l'heptamère G UUA AAC de l'Équine Artris Virus, qui provoque un décalage de phase presque au même taux que la séquence de l'IBV, ainsi que d'autres constructions artificielles contenant un G à la première position de l'heptamère (Brierley *et al.* 1992).

### **B.1.b. La structure secondaire**

La séquence glissante est nécessaire mais non suffisante pour rendre compte des taux élevés de décalage de phase en -1 mesurés. Dans tous les cas cités ci-dessus, il existe une séquence nucléotidique palindromique, qui pourrait adopter une structure secondaire de type tige-boucle en aval de l'heptanucléotide. Cette structure secondaire est invoquée pour expliquer ces hauts taux de décalage de phase. L'influence de cette structure secondaire a pu être vérifiée par des mutations ponctuelles et des délétions. Dans le cas de l'IBV, une délétion de cette séquence palindromique abolit le décalage de phase. Mais une délétion, indépendante, de la séquence en aval de cette tige-boucle produit un effet presque aussi drastique (2 à 3 % de décalage de cadre de lecture, par rapport aux 25 % du sauvage) (Brierley *et al.* 1989; Brierley *et al.* 1991). Cet ARNm est en effet capable de prendre une conformation tertiaire plus complexe, où la boucle de la première tige-boucle est appariée avec une séquence située en aval, constituant ce que l'on appelle un pseudo-nœud de type H (H pour "hairpin-like") : deux tiges coaxiales "empilées" reliées par une boucle (ten Dam *et al.* 1992) (Fig. 5).

En revanche, il existe des systèmes qui naturellement n'utilisent qu'une tige-boucle pour stimuler le décalage de phase (Parkin *et al.* 1992). Mais un pseudo-nœud stimule le décalage de phase plus qu'une tige-boucle, à



- S1 : Tige 1
- S2 : Tige 2
- L1 : Boucle 1
- L2 : Boucle 2
- Adénine intercalée provoquant la courbure du pseudo-nœud (voir texte)

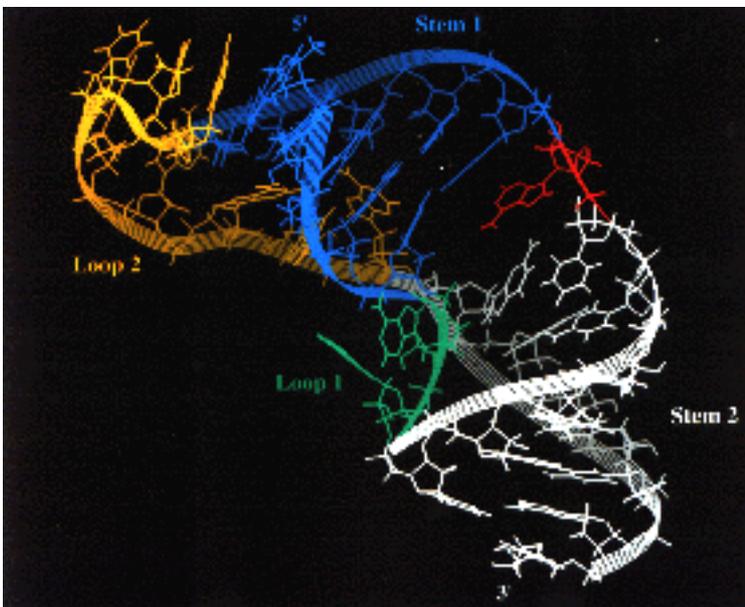


Figure 5 : **Différentes représentations schématiques équivalentes d'un pseudo-nœud de type H.**  
 Sur fond noir, structure obtenue par RMN, de Shen *et al.* 1995

stabilités thermodynamiques théoriques égales (Brierley *et al.* 1991; Chamoro *et al.* 1992). Certains décalages de cadre de lecture semblent se faire sans structure secondaire évidente (ten Dam *et al.* 1990), et à l'inverse, quelques structures secondaires caractérisées s'avèrent beaucoup plus complexes qu'un simple pseudo-nœud, telles des "oreilles de lapin", des double tige-boucles, des "tiges boucles s'embrassant" (revues dans Miller *et al.* 1995; Farabaugh 1996).

L'existence de ces structures secondaires et leur implication dans le phénomène du décalage de phase en -1 ont été vérifiées par des mutations ponctuelles ou compensatoires dans plusieurs systèmes : IBV (Brierley *et al.* 1989; Brierley *et al.* 1991), ScV (Dinman *et al.* 1991), MMTV (Chamoro *et al.* 1992), Feline Immunodeficiency Virus (FIV, Morikawa *et al.* 1992), Simian Retrovirus (SRV1, ten Dam *et al.* 1994 et 1995). Dans la majorité des mutants analysés, on peut corrélérer le taux de décalage de phase avec la présence d'une structure secondaire, la structure primaire n'ayant apparemment pas d'influence propre. Mais des mutants compensatoires ne restaurent pas toujours un taux de décalage de phase identique au sauvage (Jacks *et al.* 1988a; Brierley *et al.* 1991; Chamoro *et al.* 1992; ten Dam *et al.* 1994; ten Dam *et al.* 1995). De plus la corrélation entre la stabilité de la structure secondaire et le taux de décalage de phase n'est pas toujours évidente (Brierley *et al.* 1989; Kollmus *et al.* 1994; Chen *et al.* 1995; Kollmus *et al.* 1996).

D'autres expériences menées pour analyser la structure spatiale des pseudo-nœuds par des sondes chimiques ont également suggéré que la stabilité thermodynamique des structures secondaires n'est pas seule en jeu. Ces études ont porté sur la conformation spatiale du pseudo-nœud du MMTV, et ont été depuis étendues à d'autres structures secondaires nécessaires au décalage de phase en -1. L'étude d'un variant du pseudo-nœud du MMTV, nommé VPK, permettant un fort taux de décalage de phase en -1, a montré une structure coudée. Une adénine à la jonction entre les deux tiges empêche celles-ci de se positionner de façon coaxiale. Ce type de structure est prédit pour être moins stable que deux tiges empilées coaxialement, et pourtant le décalage de phase en -1 est plus stimulé par ce type de pseudo-nœud (Chen *et al.* 1995). Les auteurs ont confirmé ces prédictions en étudiant précisément la structure tridimensionnelle de différents pseudo-nœud par Résonance Magnétique Nucléaire (RMN) (Shen *et al.* 1995). Le VPK est un pseudo-nœud hautement organisé, le nucléotide à l'interface entre les deux tiges crée un angle de 112° entre les deux, et les deux boucles sont très précisément

placées par rapport aux sillons de l'hélice A : la boucle 1, de petite taille (en général 1 à 2 nucléotides) passe dans le petit sillon, alors que la boucle 2 passe dans le grand sillon (cf. Fig. 5). Pratiquement tous les nucléotides de la boucle semblent participer à la formation de la structure, par des contraintes de taille/distance entre les deux tiges. Ce pseudo-nœud est donc très contraint, mais les analyses en RMN indiquent que quatre nucléotides de la structure sont capables de mouvements rapides, donnant une certaine flexibilité à cette structure rigide (Shen *et al.* 1995). Par la suite, on a pu démontrer que dans d'autres pseudo-nœuds il y avait également une adénine non appariée entre les deux tiges (Sung *et al.* 1998).

La stabilité thermodynamique de la structure secondaire ne semble donc pas être le seul facteur nécessaire pour expliquer de forts taux de décalage de phase en -1. Ces considérations seront discutés plus bas, où j'essayerai de donner un modèle expliquant le rôle de tous les acteurs dans l'obtention de forts taux de décalage de phase de lecture en -1.

### **B.1.c. La distance heptanucléotide/structure secondaire**

Des expériences de délétion en amont de la structure secondaire, et en aval de la séquence glissante ont montré que l'on ne peut pas ou peu faire varier la distance entre les deux éléments *cis* sans affecter fortement le décalage de phase en -1. En ce qui concerne l'IBV, l'ajout d'un codon aux 6 nucléotides séparant séquence glissante et pseudo-nœud abolit le décalage de phase; un codon en moins le fait baisser d'un facteur 10 par rapport à la situation normale (Brierley *et al.* 1989). Cette distance semble donc déterminée pour une cible particulière, mais par contre elle est variable en fonction des cibles étudiées : de 3 à 14 nucléotides en fonction des virus (Brierley 1995). Si les différents pseudo-nœuds sont capables de prendre des formes différentes, plus ou moins coudées, cela pourrait expliquer que la distance optimale varie d'une cible à l'autre.

### **B.1.d. Un codon stop en 3' de la séquence glissante**

Un codon stop est retrouvé juste en aval de l'heptanucléotide glissant dans une majorité de virus. Il a été démontré dans le cas du RSV que celui-ci pouvait être remplacé par un autre codon stop, ou un codon sens,

sans variation du taux de décalage de phase (Jacks *et al.* 1988a). Réciproquement, un codon stop ajouté en aval de la séquence glissante ne modifie pas le taux de décalage de cadre du HTLV-2 (Kollmus *et al.* 1994).

Mais toutes les expériences ne présentent pas des résultats aussi contrastés. En remplaçant un codon sens GGG par un stop juste après l'heptamère glissant du VIH-1 le décalage de phase est réduit d'un facteur 10 environ chez *E. coli*, et le rend sensible au taux d'expression des facteurs de terminaison de la traduction (Horsfield *et al.* 1995). La conclusion des auteurs est que le décalage de phase a lieu quand l'heptamère U UUU UUA est placé dans les sites E et P, et le stop au site A. Mais il n'a pas été déterminé si le décalage de phase était dépendant de la première partie de l'heptamère, c'est-à-dire si le décalage de phase se faisait en tandem comme dans les systèmes eucaryotes. Il pourrait s'agir dans ce contexte (chez *E. coli*, et sans structure secondaire) d'un phénomène de glissement du seul deuxième ARN<sub>t</sub><sup>Leu</sup> de U UUA vers UUU A. Le codon GGG juste en aval, rare chez *E. coli* pourrait ralentir le ribosome, et donc promouvoir ce nouveau type de décalage par une pause traductionnelle. Le remplacement de GGG par un codon stop rendrait alors ce décalage de phase sensible au taux de RF présent dans la cellule. D'autant plus que d'autres expériences ont montré par séquençage protéique que le site minimal du VIH-1 pouvait provoquer un décalage de phase du seul premier ARN<sub>t</sub> du site (supposé être au site P) (Jacks *et al.* 1988b; Yelverton *et al.* 1994). Le site de décalage de phase très monotone du VIH-1 pourrait être très glissant, autorisant donc trois types de décalage d'ARN<sub>t</sub> : en tandem, ou chacun des ARN<sub>t</sub> seuls de leur côté. Cette hypothèse paraît plus plausible, puisqu'il ne semble pas y avoir d'interaction codon/anticodon au site E (Farabaugh 1996). Dans des expériences indépendantes, et qui n'ont donné lieu à aucune vérification depuis, il a été montré qu'un codon stop placé deux codons en aval de la séquence glissante dans la phase 0 augmente le décalage de phase de 1,8% (U UUU UUA GGG AAG) à 2,8% (U UUU UUA GCC **UAG**) (Weiss *et al.* 1989). En ce qui concerne le BYDV, un codon stop est nécessaire au décalage de cadre de lecture en -1 (Brault *et al.* 1992).

A ce jour, il est généralement admis que les codons stops n'influent pas sur le décalage de phase, et les expériences sus-citées n'ont donné lieu à aucun approfondissement. La question reste ouverte, mais il est sûr que cet effet, s'il existe, n'est pas généralisable.

### **B.1.e. La butée du ribosome**

Pour expliquer le rôle de la structure secondaire et l'importance de l'espacement entre la séquence glissante et la structure secondaire, on a invoqué un système de barrière physique, énergétique, précisément placée (le pseudo-nœud ou la tige-boucle), où le ribosome viendrait buter, permettant aux ARNt présents aux sites A et P de se recaler un nucléotide en arrière (des modèles plus précis sont discutés à la fin de ce paragraphe). Les premières preuves expérimentales de cette pause du ribosome ont été obtenues par hasard durant l'étude *in vitro* du décalage de phase en -1 nécessaire à l'expression de la sous-unité gamma de l'ADN polymérase III de *E. coli* sur le messager du gène *dnaX*. Lors d'expériences témoin réalisées pour prouver l'absence de décalage de phase en l'absence de séquence glissante, il est apparu des intermédiaires de traduction dont la taille correspond exactement à celle obtenue si la traduction était bloquée juste en amont de la tige-boucle. Ces intermédiaires sont détectables quand la réaction de traduction est interrompue, mais disparaissent si l'on laisse la réaction se poursuivre. L'auteur a interprété ce résultat comme un arrêt transitoire, une pause, du ribosome en cours de traduction au niveau de la tige-boucle (Tsuchihashi 1991). Cette pause est indépendante de l'heptamère glissant, vu que la construction testée en est dépourvue.

Cette "pause" du ribosome a été étudiée de manière approfondie en ce qui concerne l'IBV (Somogyi *et al.* 1993) et le ScV (Tu *et al.* 1992). Les deux approches utilisent une technique différente, et apportent des conclusions complémentaires. Les travaux sur l'IBV ont été réalisés de manière similaire à celle évoquée ci-dessus pour *dnaX* : les auteurs ont réalisé des expériences de traduction *in vitro* sur des ARNm ne comportant pas la séquence glissante. Des intermédiaires de traduction sont visualisables à 26°C, et en présence d'édéine (utilisée pour synchroniser l'initiation de la traduction). L'un des intermédiaires de traduction, de taille conforme à un arrêt de traduction juste en amont du pseudo-nœud, est dépendant de la présence du pseudo-nœud. De plus les auteurs ont fait varier les conditions physico-chimiques afin de modifier la stabilité de la structure secondaire : une hausse de la température de 26 à 35°C diminue le taux de produit intermédiaire, alors qu'une hausse de la concentration en Mg<sup>2+</sup> de 1,5 à 2 mM diminue la vitesse de traduction, mais augmente la quantité de produit de

traduction résultant d'une pause du ribosome. Des facteurs stabilisant (le  $Mg^{2+}$ ) ou déstabilisant (une hausse de température) le pseudo-nœud font donc également varier le taux d'apparition de l'intermédiaire de traduction : il y a une corrélation entre la stabilité pseudo-nœud et fréquence/durée de la pause (Somogyi *et al.* 1993).

Les travaux réalisés sur le ScV consistent à visualiser des fragments d'ARN protégés de la digestion nucléasique micrococcalle par les ribosomes présents sur le messager (Wolin *et al.* 1988). Cette technique permet de cartographier très précisément la zone de protection : *in vitro*, il y a protection préférentielle de la zone située juste en amont du pseudo-nœud, à deux positions séparées de 3 nucléotides. Les auteurs proposent que ces deux positions correspondent à une pause avant et après le décalage de phase, invoquant pour expliquer que ces deux pauses soient séparées par 3 nucléotides et non un seul, que la nucléase agit à des sites préférentiels (Tu *et al.* 1992). Une expérience permettant de déterminer si une de ces deux positions est dépendante de la présence de l'heptamère, et donc du décalage de phase, n'a pas été réalisée.

Pour les travaux évoqués ci-dessus, les structures secondaires étudiées (sauvages, mutants ponctuels, compensatoires, délétions), montrent une corrélation entre pause du ribosome et décalage de phase en -1, même s'il n'y a pas forcément causalité. Mais la corrélation inverse n'est pas absolue : quelques structures secondaires provoquent une pause (parfois de moindre importance), bien que le décalage de phase ne soit plus du tout décelable. Par exemple, une délétion de la seconde partie du pseudo-nœud de l'IBV, laissant seule une tige-boucle, provoque quand même une certaine pause; mais il n'y a plus décalage de phase (Somogyi *et al.* 1993). De même, 2 bases enlevées dans la deuxième tige du pseudo-nœud de ScV ne changent pas l'empreinte observée, alors qu'on ne détecte plus du tout de décalage de phase (Tu *et al.* 1992). Il peut donc y avoir pause sans décalage de phase.

Ceci, et le fait que la stabilité théorique des structures secondaires ne soit pas toujours corrélée au taux de décalage de phase nous amène à proposer des modèles plus compliqués :

(i) la progression du ribosome ne dépend pas seulement de la force des structures secondaires, mais aussi de leur composition et/ou structure tridimensionnelle (il serait par exemple plus difficile de dérouler un pseudo-nœud qu'une tige-boucle, cf. ci-dessous § B.1.g.iii, "Torsion Resistance Model"),

(ii) la structure secondaire ne joue pas qu'un rôle simple de barrière physique.

Avant de préciser quelles sont les hypothèses quant à ce rôle supplémentaire (interaction avec le ribosome, avec un facteur soluble de la traduction,...), je vais préciser quels sont actuellement les modèles temporels de l'événement de décalage de phase en -1.

### **B.1.f. Modèles temporels**

Il existe en théorie plusieurs moments où les deux ARNt présents dans le ribosome sont capables de se décaler d'un nucléotide en 5' sur le messager (Fig. 6 et 7).

(i) Décalage avant le transfert peptidique

Il a été proposé que le glissement du peptidyl-ARNt et de l'aminocyl-ARNt se faisait après qu'ils se soient positionnés dans les sites P et A, mais avant qu'il y ait eu transfert du peptidyl d'un ARNt à l'autre (Fig. 6 et Jacks *et al.* 1988a).

(ii) Décalage après le transfert peptidique

Par la suite, il a été proposé que le glissement se fasse d'une façon post-peptidyl transfert, mais avant la translocation de l'ARNt déacylé et du peptidyl-ARNt vers les sites E et P (Fig. 7 et Weiss *et al.* 1989).

Il n'y a pas actuellement, par séquençage protéique ou par mutagenèse dirigée, de moyens pour discriminer entre ces deux modèles. Mais des arguments permettent de favoriser l'un ou l'autre :

- le fait que le décalage de phase du VIH-1 *in vitro* (Jacks *et al.* 1988b), ou celui d'autres cibles heptamériques en expression hétérologue chez *E. Coli* (Weiss *et al.* 1989; Yelverton *et al.* 1994), se produisent parfois grâce au glissement d'un seul d'ARNt au site P, indique que le décalage pré-peptidyl transfert est possible; mais il ne s'agit pas là de décalage de phase en -1 en tandem, et ce décalage en singleton peut aussi avoir lieu juste après la translocation,

- l'effet d'antibiotiques spécifiques du transfert peptidique (Dinman *et al.* 1997b), de la Pokeweed Antiviral Protein (PAP) inhibitrice de la translocation (Tumer *et al.* 1998), de mutants EF-1A (Dinman *et al.*

1997a) sur le taux de décalage de phase en -1 du ScV *in vivo* suggèrent que le décalage a lieu dans les toutes premières étapes de l'élongation, sans doute juste avant l'hydrolyse du GTP par EF-1A (Dinman *et al.* 1997a),

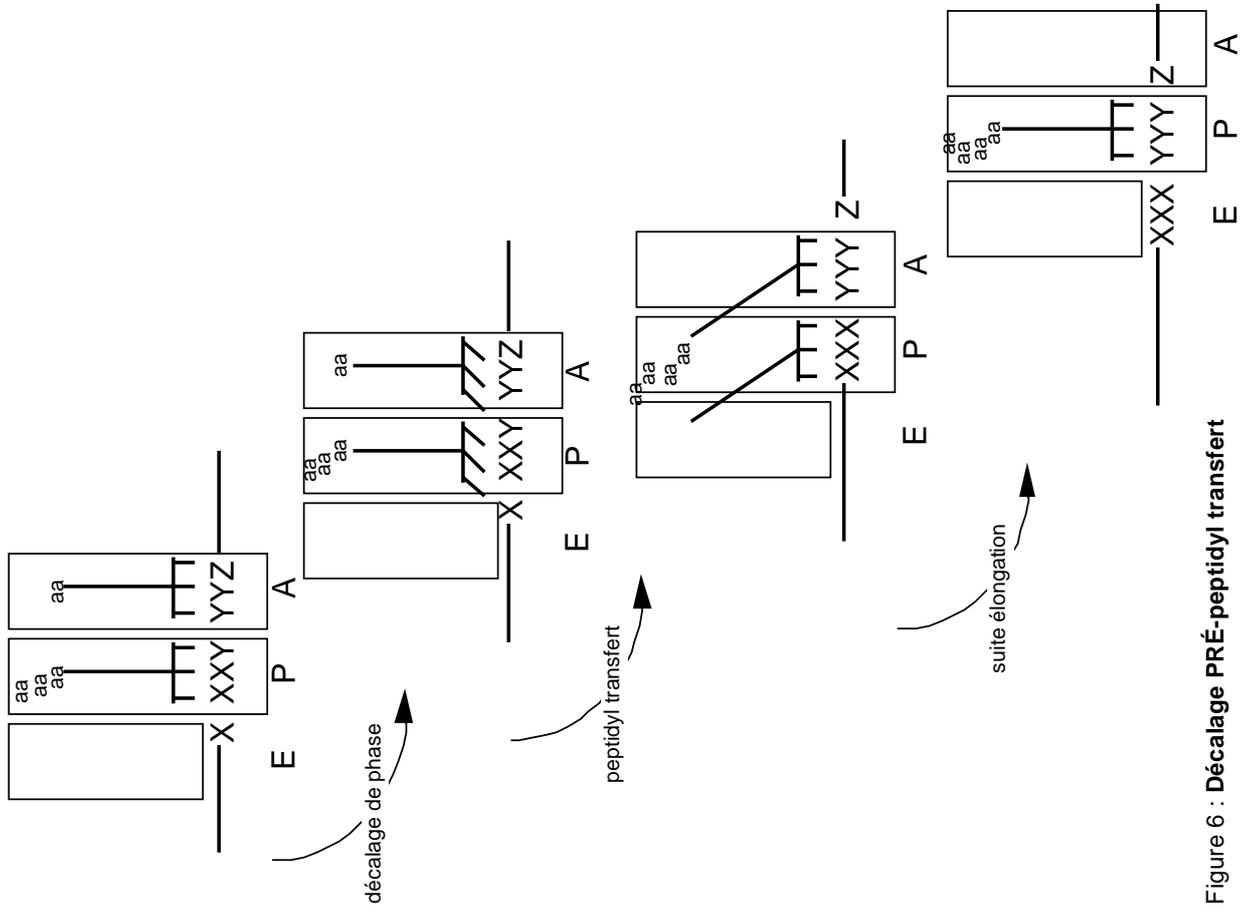


Figure 6 : Décalage PRÉ-peptidyl transfert

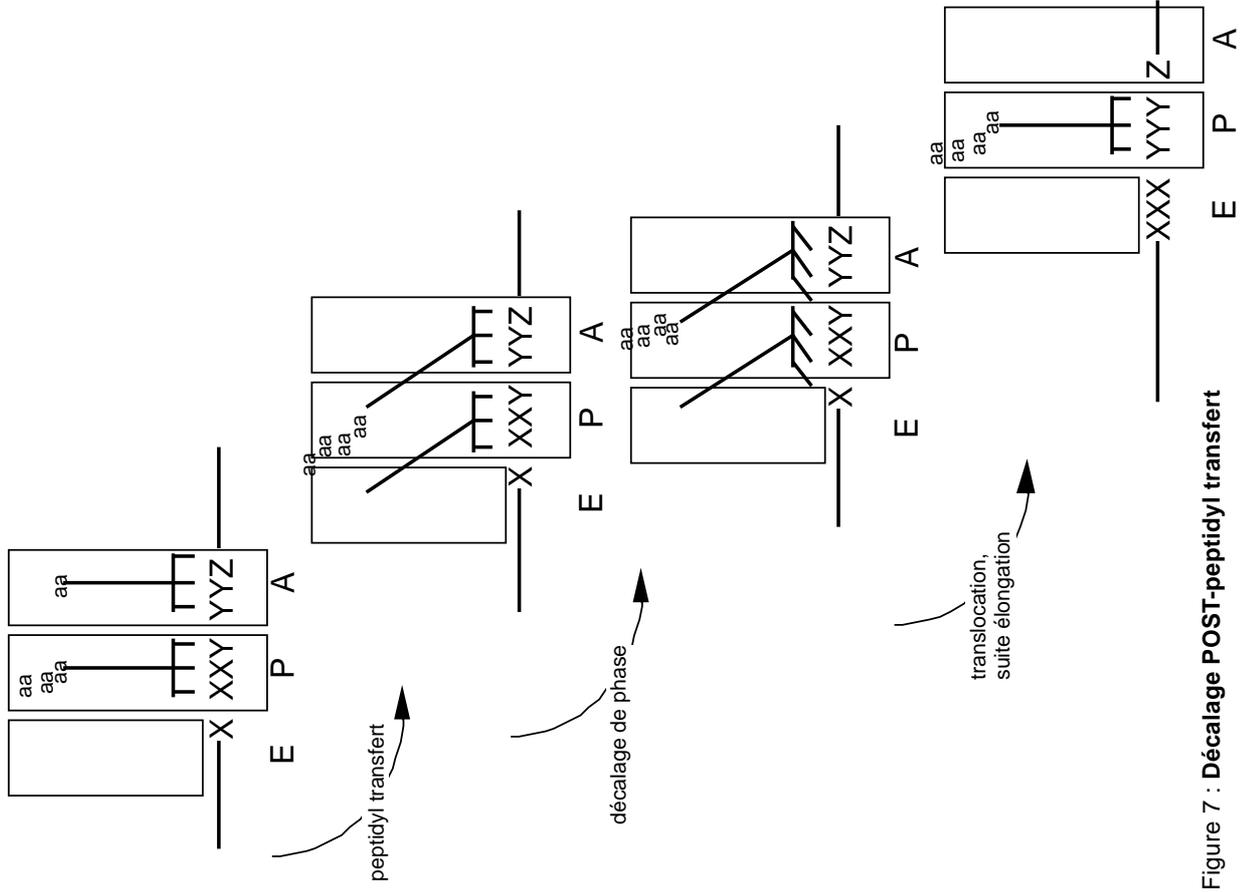


Figure 7 : Décalage POST-peptidyl transfert

- d'un autre côté, des arguments cinétiques favorisent le deuxième modèle : le transfert du peptidyl serait extrêmement rapide une fois le site A occupé (en fait la durée de cette étape ne peut être mesurée, et elle est considérée comme étant instantanée, Farabaugh 1996). Le décalage des ARNt se ferait donc entre le transfert du peptidyl et la translocation.

### **B.1.g. Modèles mécanistiques**

#### (i) Modèle stochastique simple

On peut imaginer que la simple pause de l'élongation au niveau d'un site glissant permet de favoriser cinétiquement le décalage par rapport à l'élongation normale. Mais ce modèle n'explique pas tous les résultats contradictoires de corrélation entre durée de pause (stabilité de la structure secondaire) et le taux de décalage de phase. Plusieurs autres modèles tentent de prendre en compte ces exceptions et donnent en plus une cause énergétique au décalage de phase, notion absente du modèle stochastique.

#### (ii) Décalage au cours de tentatives abortives de translocation

Si la structure secondaire bloque le ribosome, cela doit interférer avec la translocation, au moment où le ribosome tente de faire coulisser le complexe ARNm/ARNt de 3 nucléotides. Ce mouvement de translocation est initié au niveau du peptidyl-ARNt (site A) qui, par un changement de conformation stabilisé par EF2, est transféré dans le site P pendant que l'ARNt déacylé passe du site P vers le site E. L'ARNm étant bloqué à l'entrée du ribosome par la structure secondaire, il ne peut pas coulisser de façon simultanée avec les ARNt, ce qui provoque alors un décalage entre ARNt/ARNm, favorisé par la présence de l'heptamère glissant (Fig. 8). Si ce modèle est vrai, on peut imaginer que la séquence UU UUU UUA pourrait provoquer un décalage de phase en -2 pendant les essais de translocation successifs, mais ceci n'a jamais été testé spécifiquement. Le décalage de phase du ScV n'est pas affecté par un inhibiteur de la translocation, la PAP (Tumer *et al.* 1998), ce qui laisse supposer que la fréquence et/ou l'efficacité des essais de translocation pendant la pause du ribosome n'est pas cruciale dans ce modèle.

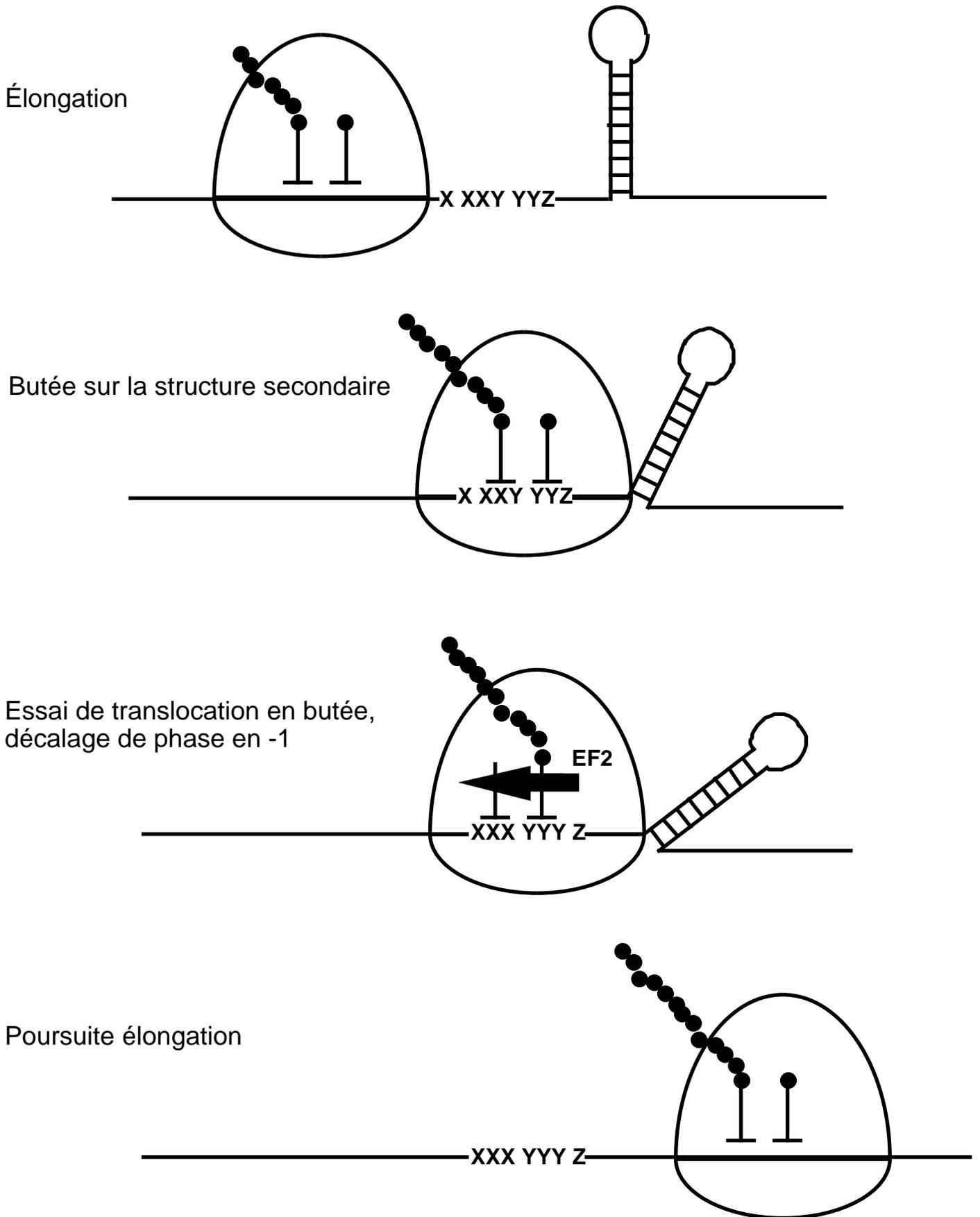


Figure 8 : Essais de translocation provoquant un décalage de phase en -1

(iii) Décalage grâce à un déroulement/ré-enroulement partiel de la structure secondaire

Ce deuxième modèle suppose qu'une activité hélicase associée au ribosome est responsable du déroulement des structures secondaires sur le messager. Au moment où le ribosome est bloqué en amont d'une structure secondaire, cette activité hélicase permet d'ouvrir la double hélice ARN pour poursuivre l'élongation. Si la structure secondaire est particulièrement stable, elle peut avoir tendance à se reformer, et donc à tirer l'ARNm par rapport au ribosome et aux ARNt (Fig. 9). Ce modèle, en plus d'une explication énergétique (la force du ré-appariement de la structure secondaire donne l'énergie nécessaire au décalage de phase), expliquerait des résultats jusqu'alors difficilement interprétables :

- dans les expériences réalisées pour mettre en évidence les pauses du ribosome en amont d'un pseudo-nœud, les résultats de cartographie précise montrent deux endroits de pause séparés par 3 nucléotides (Tu *et al.* 1992). Deux pauses distinctes sont difficilement explicables s'il y a une seule structure secondaire précisément positionnée par rapport au ribosome. On pourrait envisager que le ribosome, bloqué par le pseudo-nœud à sa base (première pause), puisse ensuite avancer encore de 3 nucléotides en aval (deuxième pause) lors d'un déroulement partiel de la structure secondaire (Fig. 9). Après cette seconde pause, le pseudo-nœud se reformerait complètement, et le complexe ribosome/ARNt serait contraint de reculer en 5' par rapport à l'ARNm;

- si on élimine le premier appariement du pseudo-nœud de l'IBV, le taux de décalage de phase ne varie pratiquement pas (Brierley *et al.* 1991). Si effectivement le pseudo-nœud est partiellement défait avant décalage des ARNt, le rôle des premiers appariements peut être moins critique,

- si on fait varier la distance intercalante ("spacer") heptamère-structure secondaire du FIV (cf. aussi supra) par addition ou délétion de 3 nucléotides, le décalage de phase diminue de 6 à 15 fois. Si l'on mute les 3 derniers nucléotides de la région intercalante afin de leur permettre de s'apparier dans la première tige, la structure secondaire est de fait rapprochée de 3 nucléotides de la séquence glissante; mais dans ce cas on n'observe aucune modification du taux de décalage de phase (Morikawa *et al.* 1992). Si ces appariements supplémentaires sont défaites par le ribosome, jusqu'à atteindre la distance optimale pour le décalage (3 nucléotides plus loin), ces résultats deviennent interprétables simplement.

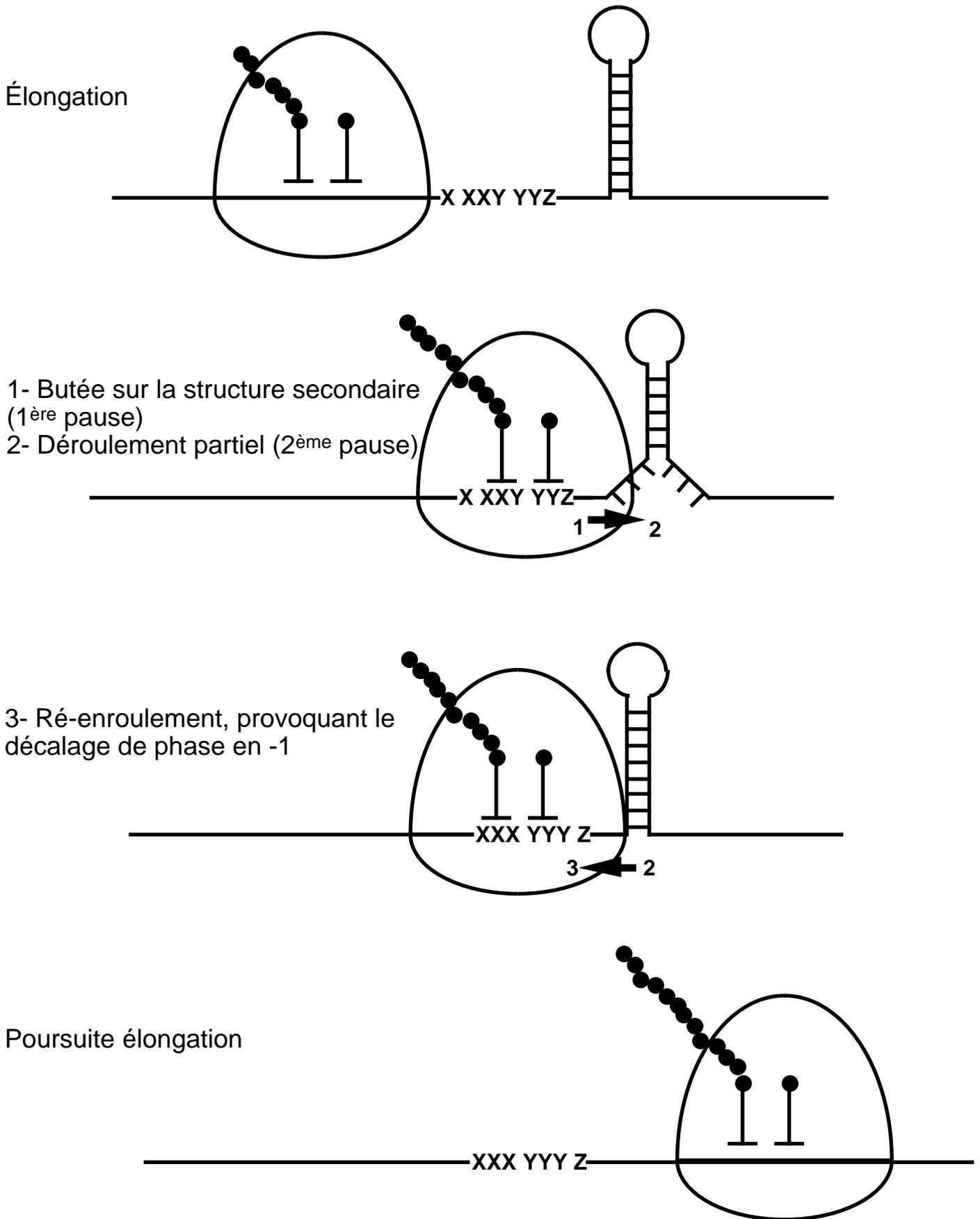
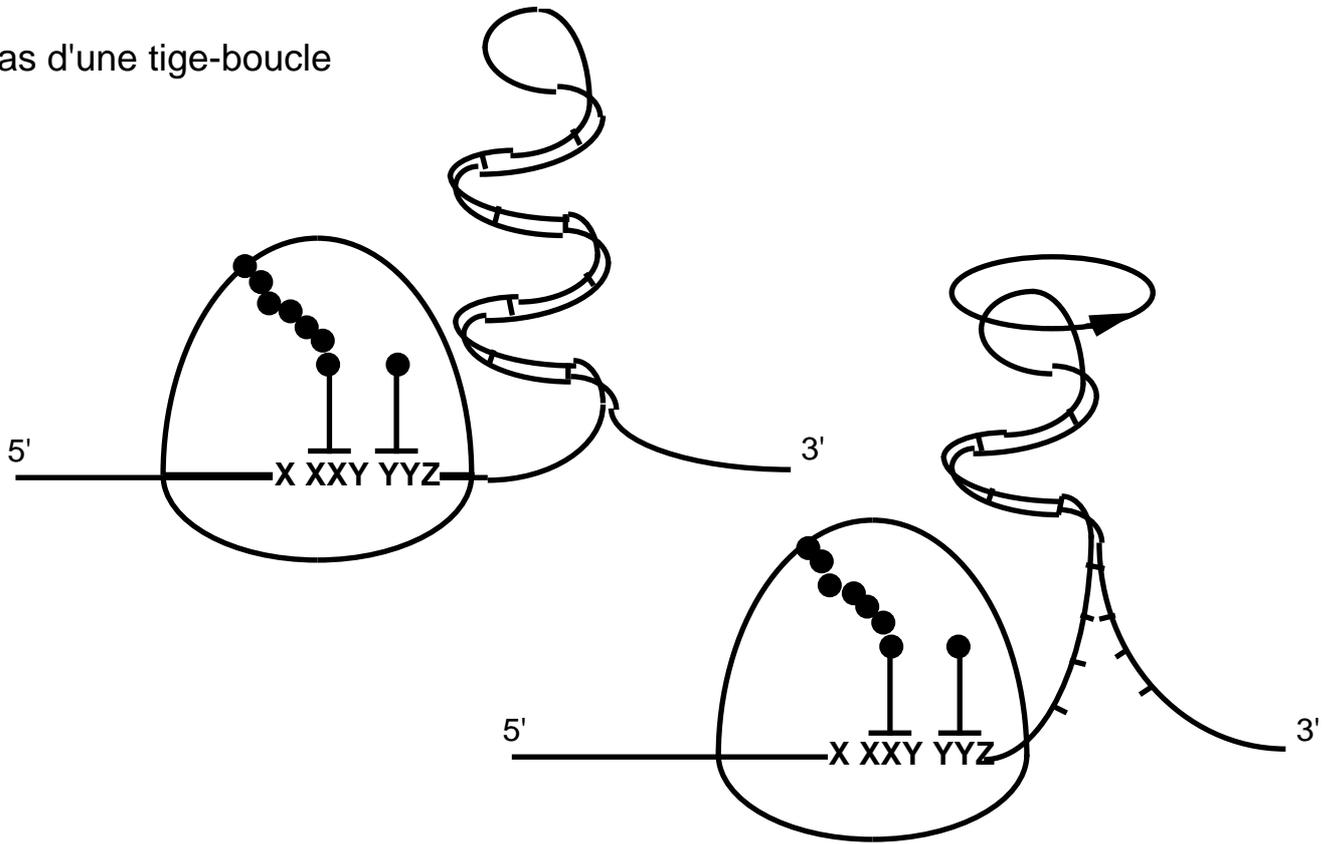


Figure 9 : **Déroulement / ré-enroulement de la structure secondaire provoquant un décalage de phase en -1**

a- Cas d'une tige-boucle



b- Cas d'un pseudo-nœud

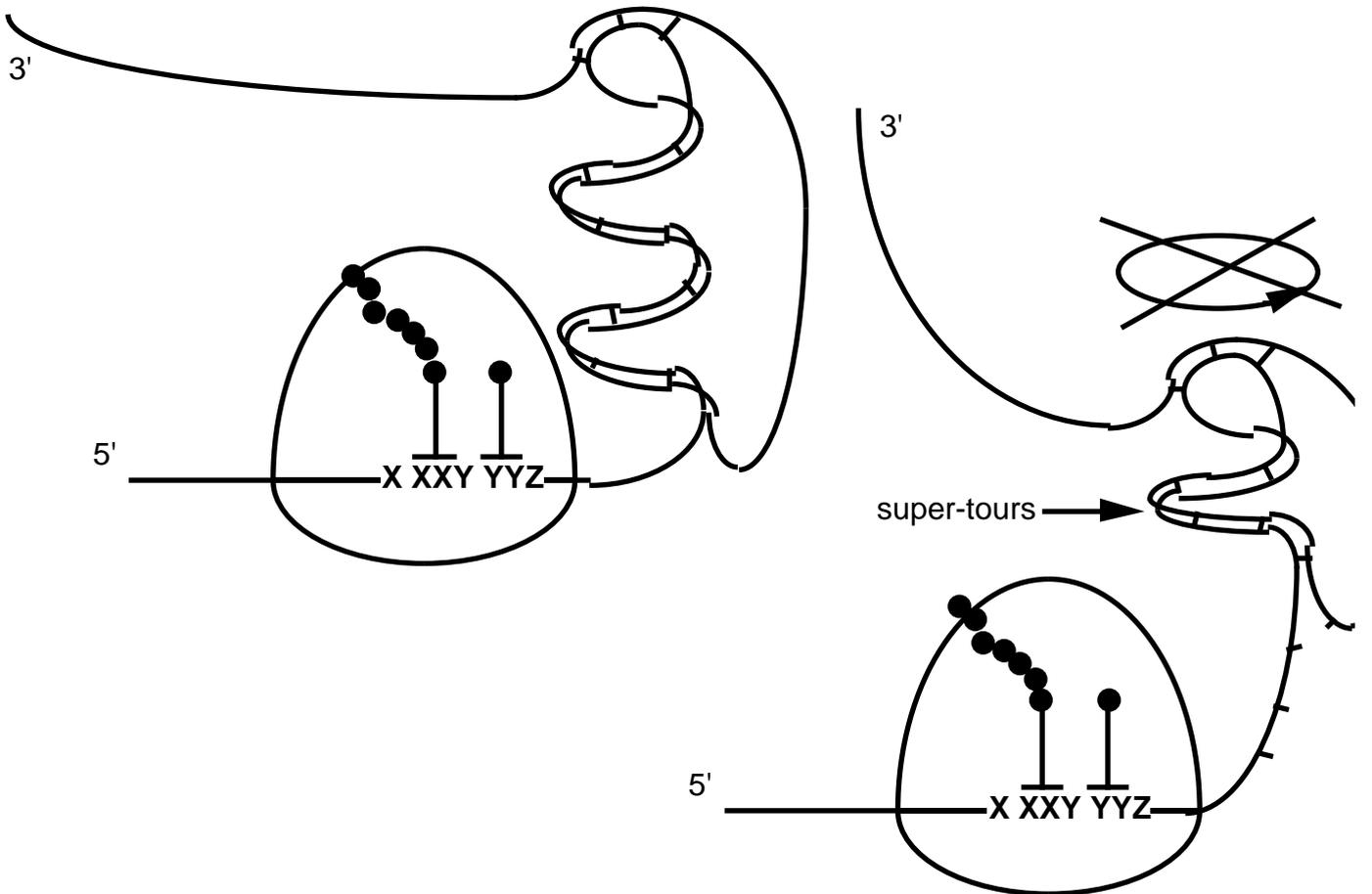


Figure 10 : Torsion Resistance Model

Ce modèle peut également expliquer pourquoi des structures secondaires de type pseudo-nœud sont plus efficaces à induire le décalage de phase que des tige-boucles. Cette hypothèse, évoquée sous une autre forme, et appelée “Torsional Resistance Model” (Dinman 1995) suggère qu’un pseudo-nœud étant maintenu à sa base par le ribosome, et à l’autre extrémité par sa boucle engagée dans l’interaction avec la deuxième tige, toute tentative de déroulement provoquerait la formation de super-tours, qui opposeraient une forte résistance à l’avancement du ribosome, et augmenteraient la quantité d’énergie disponible pour faire reculer le ribosome (Fig. 10b). Ceci n’est pas applicable aux tige-boucles, dont l’extrémité est libre (Fig. 10a).

#### (iv) Facteurs interagissant avec la structure secondaire

Finalement, on ne peut pas exclure des facteurs liant la structure secondaire. Il a été prouvé qu’une protéine se liant à une structure secondaire peut influencer sur le décalage de phase au niveau du site de décalage de cadre du VIH-1 (Kollmus *et al.* 1996). Cette étude montre qu’un facteur *trans* de ce type peut être envisagé, et non qu’il existe. Ce facteur, s’il existe, est difficile à mettre en évidence :

- un excès d’oligonucléotide présentant la structure secondaire du SRV1 ou du MMTV ne modifie pas *in vitro* le taux de décalage de phase en -1 (ten Dam *et al.* 1994; Chen *et al.* 1995),
- des expériences de retard sur gel ne révèlent pas de liaison d’un facteur protéique sur le pseudo-nœud du ScV (Dinman 1995),
- des études génétiques ont été entreprises pour isoler des facteurs faisant varier le taux de décalage de phase du ScV, mais aucun des neuf mutants isolés n’est spécifique de la structure secondaire (Dinman *et al.* 1994; Dinman *et al.* 1995; Cui *et al.* 1996; Cui *et al.* 1998).

Le pseudo-nœud pourrait aussi interagir directement avec le ribosome, et y modifier une activité biochimique. Quelques pistes peuvent être évoquées :

- interaction avec un ARNr,
- interaction avec une protéine du ribosome : S15 chez *E. coli* autorégule sa propre traduction par fixation en 5' de son messenger sur un pseudo-

nœud possédant une adénine intercalée très semblable au pseudo-nœud utilisé par les rétrovirus (Benard *et al.* 1998) et S4, impliquée dans la fidélité de la traduction utilise aussi ce type de régulation (Tang *et al.* 1989),

- interférence avec EF-1A : il a été proposé que les structures observés par RMN pour certains pseudo-nœuds ressemblent fortement à une structure codon/anticodon appariés (Shen *et al.* 1995).

Nous avons donc vu qu'en ce qui concerne le décalage de phase en -1 certains éléments semblaient constants pour toutes les cibles étudiées : heptamère glissant, structure secondaire, pause traductionnelle. La situation semble plus hétérogène pour le décalage de phase en +1, même si des similitudes avec le décalage en -1 existent, notamment une pause en cours de traduction.

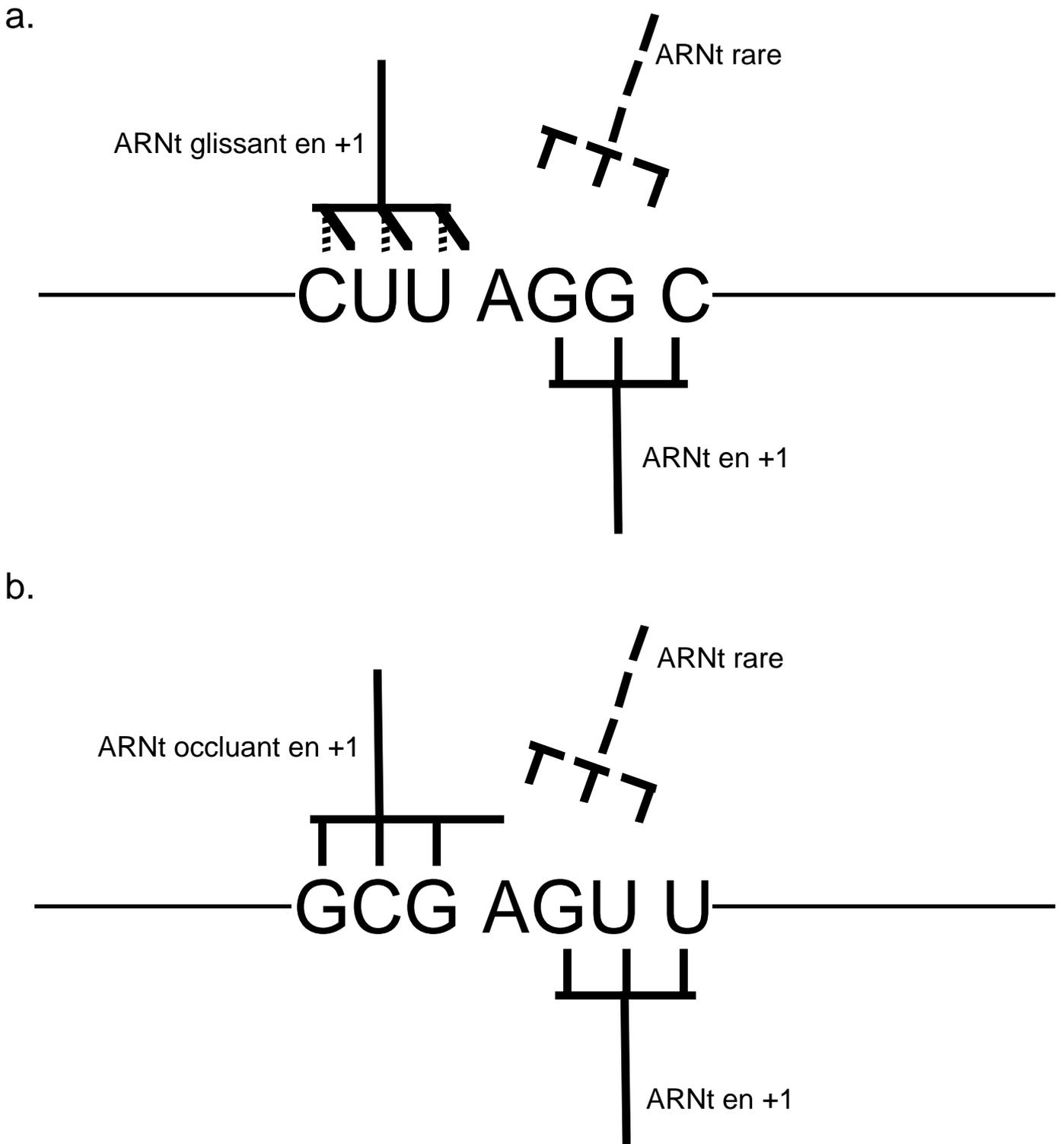


Figure 11 : Décalage de phase en +1  
 a. ARNt glissant du Ty1  
 b. ARNt occluant du Ty3

## **B.2. Décalage de phase en +1**

Pour tous les cas connus où le complexe ribosome/ARNt glisse d'un nucléotide en aval, un ARNt particulier est impliqué au site P. Il peut agir en se ré-appariant vers la base suivante, ou l'occluant, ce qui la rend inaccessible au prochain ARNt dans cette phase, et favorise l'arrivée d'un ARNt en phase +1 (Fig. 11). Il y a donc compétition entre l'ARNt censé décoder le codon en phase 0 au site A et deux ARNt : celui au site P et celui qui va décoder le premier codon de la phase +1 en A. Le codon présent au site A juste avant le décalage de phase correspond à un ARNt cognat rare, ou est un codon Stop; on a appelé ces codons des codons "affamés" ("hungry"). Le site A est de ce fait inoccupé, et permet à l'ARNt présent au site P de favoriser l'arrivée d'un ARNt sur le codon non exactement adjacent, en masquant la première base du codon au site A.

Parmi les gènes connus, dont l'expression nécessite un décalage de phase en +1, la famille des éléments Ty constitue un bon modèle d'étude. Il existe 5 éléments Ty rétrotransposables chez *S. cerevisiae* (Ty1 à 5). Ils ont une capacité codante d'environ 5 kb et, mis à part Ty5, possèdent le même type d'organisation que les virus utilisant le décalage de phase en -1 : une ORF en 5' codant pour une protéine de type *gag* (*TYA*, ou *GAG3* pour le Ty3) suivie, dans la phase +1, d'une ORF (*TYB*, ou *POL3*) codant pour une rétrotransposase, une RNase H, une intégrase et une protéase. Cette deuxième phase de lecture ne possédant pas d'entrée de ribosome indépendante, le décalage de phase en +1 entre *TYA* et *TYB* est indispensable à la propagation du transposon. Ce décalage de phase se passe aussi sur un heptamère (moins clairement défini que dans le cas du décalage de phase -1), et la stimulation semble ici aussi provenir d'une pause du ribosome. Les deux familles de Ty exploitent deux types de mécanismes similaires.

### **B.2.a. Famille Ty1,2 et 4**

C'est le Ty1 qui nous servira de modèle pour décrire cette famille. Pour ces éléments transposables, le décalage de phase se produit *in vivo* avec une fréquence élevée de l'ordre de 20-40%, selon les systèmes de mesure. Cependant une mesure immunologique du taux de décalage de phase du

Ty1 endogène indiquerait qu'il serait de l'ordre de 3% (Kawakami *et al.* 1993), ce qui est confirmé par certains systèmes de mesure enzymatiques (Balasundaram *et al.* 1994). Ces différences de mesure de taux de décalage ne sont pas comprises à ce jour (P. J. Farabaugh, communication personnelle).

Il a été démontré que le décalage de phase en +1 du Ty1 se produisait au site heptamérique CUU AGG C (Belcourt *et al.* 1990). Les mécanismes de décalage en -1 et en +1 ne semblent pas faire appel aux mêmes signaux pour stimuler le recodage. La pause du ribosome permettant un décalage de phase +1 n'est pas initiée par une structure secondaire, mais par la très faible abondance de l'ARNt<sub>Arg</sub> qui vient décoder le deuxième codon de la phase 0 : AGG. La surexpression de cet ARNt provoque une abolition du décalage de phase +1 (Belcourt *et al.* 1990; Xu *et al.* 1990). Les mécanismes ne sont pas complètement compris, puisqu'un autre codon rare ne peut pas forcément prendre la place de celui-là (Belcourt *et al.* 1990, D. Burke, thèse de doctorat de l'Université de Washington à Saint-Louis, USA, cité par Farabaugh 1996).

Le ribosome étant arrêté à un endroit précis, il va devoir encore initier le décalage de phase. Pour le Ty1, l'ARNt<sub>Leu</sub> situé au site P serait capable de se ré-apparier sur le tétranucléotide CUU A en phase +1 grâce à son anticodon 3'GAU5', qui décode les codons Leucine CUU en phase 0 et UUA en phase +1, grâce à sa base U non modifiée (Belcourt *et al.* 1990). Il est notable que CUU est moins bien reconnu que UUA par cet ARNt. L'élongation peut alors se poursuivre grâce au décodage de GGC en phase +1 par un ARNt<sub>Gly</sub> abondant (Fig. 11). Le mécanisme pourrait aussi être de type arrivée hors de phase d'un ARNt au site A, sans que l'ARNt au site P ne se décale en +1 (voir ci-après le mécanisme d'occlusion, ou Pande *et al.* 1995).

### **B.2.b. Ty3**

Le décalage de phase du transposon Ty3 se produit à une fréquence d'environ 5% (Kawakami *et al.* 1993), au niveau de la séquence GCG AGU U. Comme pour le Ty1, le décalage de phase en +1 est favorisé par la rareté de l'ARNt<sub>Ser</sub> qui doit décoder le deuxième codon en phase 0 (AGU) (Farabaugh *et al.* 1993). Ici, le processus de décalage du Ty3 n'est pas initié par le glissement d'un nucléotide en 3' de l'ARNt présent au site P, mais par la propension de cet ARNt à "occlure" la base située en 3',

obligeant le décodage de GUU en phase +1. En effet, l'ARNt (non identifié à ce jour) capable de décoder GCG (anticodon 3'CGU5') serait incapable de décoder le codon en phase +1 CGA (au mieux un seul appariement correct en troisième position). Une expérience de compétition avec l'ARNt<sub>Ala</sub><sup>CGC</sup> isoaccepteur, muté dans son anticodon pour reconnaître le codon GCG, réduit fortement le taux de décalage de phase en +1 dépendant de l'ARNt sauvage. Ceci prouve que l'initiation du décalage de phase en +1 n'est pas dépendant seulement de l'anti-codon d'un ARNt spécifique, mais aussi de l'ARNt lui-même (Vimaladithan *et al.* 1994). En remplaçant le premier codon GCG dans la phase 0 par tous les autres codons possibles il a été montré que cette faculté d'occlure la base juste en 3' est une propriété particulière : seuls 11 ARNt sont capables de promouvoir une arrivée d'un prochain ARNt hors de phase (Vimaladithan *et al.* 1994). Ceci ne dépend pas de la nature de la séquence décodée, puisque des ARNt supposés être glissants, tels l'ARNt<sub>GAA</sub> décodant UUU, l'ARNt<sub>UUU</sub> décodant AAA, l'ARNt<sub>UUC</sub> décodant GAA, ou l'ARNt<sub>UUG</sub> décodant CAA, ne provoquent pas de décalage de phase en +1, alors que des codons que l'on conçoit difficilement glissants le permettent : GCG, CCG, AAU. Il s'agirait d'une caractéristique spécifique à ces quelques ARNt. Contrairement à ce qui se passe chez *E. coli* (Curran 1993, cf. ci-dessous § B.2.c), il n'y a pas de corrélation entre la capacité à glisser et la capacité à induire le décalage de phase en +1 chez *S. cerevisiae*.

Finalement, l'abondance de l'ARNt qui va profiter de l'occlusion pour venir décoder la phase +1 influence également le taux de décalage. Sa surexpression augmente le taux de décalage, semblant en quelque sorte "fixer" l'événement initié dans la phase initiale par l'ARNt spécial. Ceci est également vrai si l'ARNt qui provoque le décalage de phase en +1 est de type glissant, comme dans le cas du Ty1. Ceci suggère que le site A n'est pas vide, affamé, mais qu'il est occupé parfois par un ARNt capable de venir décoder au site A un codon de la phase +1 (Pande *et al.* 1995). Les mécanismes permettant cette interdépendance entre peptidyl-ARNt en phase 0 et ARNt amino acylé en phase +1 ne sont pas compris à ce jour.

Il semble qu'en plus de ce système similaire à celui de Ty1, il existerait dans le cas de Ty3 une séquence stimulant 7,5 fois le décalage de phase (Farabaugh 1996). Cette séquence serait capable de s'apparier avec l'ARNr 18S, au niveau des boucles dite "sarcine" et "530" (P.J. Farabaugh, communication personnelle).

### **B.2.c. RF2**

Le séquençage du gène *prfB* de *E. coli*, codant pour RF2, le facteur de terminaison spécifique des codons UGA et UAA, a montré que la phase ouverte de lecture était interrompue au 26<sup>ème</sup> codon par un codon UGA, et que la lecture de la suite du message nécessitait un décalage de phase en +1. Un séquençage peptidique a montré que le ribosome changeait de phase de lecture au codon 25, juste avant le codon stop (Craigén *et al.* 1985), avec une forte efficacité (entre 30 et 50% en fonction des conditions de culture). La séquence incriminée CUU UGA C (Craigén *et al.* 1986) est très similaire à celle du Ty1 de levure CUU AGG C, puisque qu'elle comporte le même codon glissant CUU, et juste après un codon "affamé". Mais ici, le codon "affamé" est le codon stop UGA lui-même, "nourri" par RF2. Avec les mêmes principes que ceux qui gouvernent le décalage de phase en +1 du Ty, on voit qu'il s'agit ici d'une boucle d'autorégulation : plus la concentration en RF2 est élevée, plus le codon UGA est reconnu, moins l'ARNt présent au site P se décale en +1, et donc moins le ribosome produit de RF2 (Fig. 12). À l'inverse, l'absence de RF2 affame le codon UAG, ce qui permet à l'ARNt<sub>Leu</sub> présent au site P de passer du codon CUU UAG C au codon presque cognat CUU UAG C, grâce à l'anti-codon 3'GAG5'. Mais contrairement au Ty1, où seul l'heptamère est responsable du décalage de phase, il faut des éléments *cis* supplémentaires pour provoquer ce décalage de phase en +1 :

#### **Le codon glissant**

CUU est particulièrement sujet au décalage de cadre en +1; on peut le remplacer par des codons glissants comme GUU ou GGG sans perte significative de décalage de phase, mais des codons où le ré-appariement des ARNt se fait en phase +1 avec perte de tout ou partie des interactions codons/anticodon ne permettent plus le décalage de phase en +1 (Weiss *et al.* 1987a). Le système de décalage de phase en +1 a été utilisé pour analyser les propriétés de glissement en +1 des peptidyl-ARNt en remplaçant le codon CUU par 32 codon sens. Chez *E. Coli*, contrairement à ce que nous avons vu chez *S. cerevisiæ* pour le Ty3, le décalage de phase dépend de l'efficacité de l'ARNt du site P à changer de phase (Curran 1993).

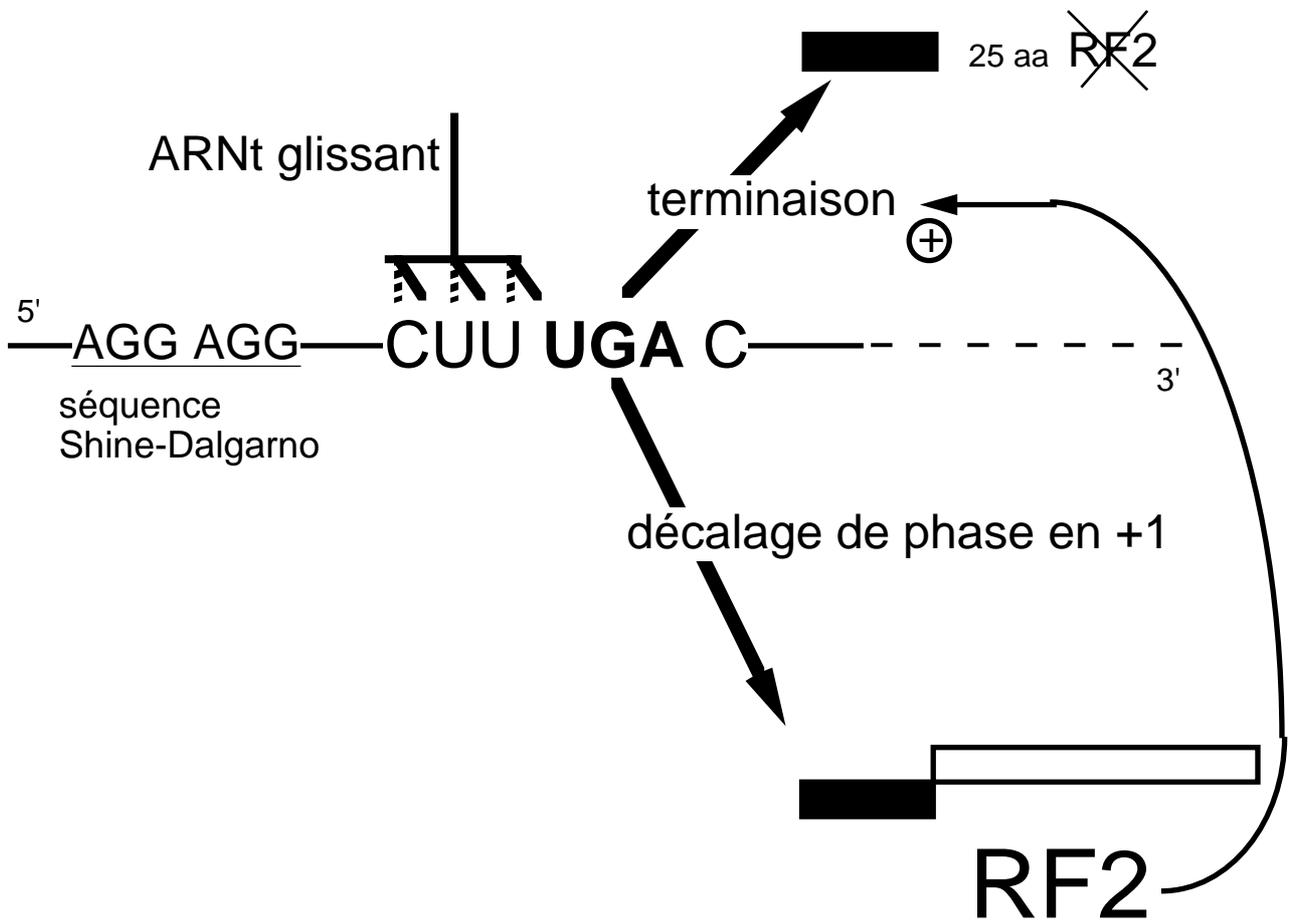


Figure 12 : **Autorégulation de RF2 par décalage de phase en +1**

## **Le contexte du codon Stop**

Le contexte du codon UGA est particulièrement défavorable à la reconnaissance par RF2 (Poole *et al.* 1995; Major *et al.* 1996), et c'est bien l'absence de RF2 qui induirait une pause du ribosome, comme le prouvent des expériences de surexpression de RF2 et de RF1 : seule la surexpression de RF2 inhibe le décalage en +1 de la cible issue de son propre gène (Craigén *et al.* 1986; Donly *et al.* 1990). L'utilisation du site de décalage de RF2 a permis deux types de mesures indirectes en fonction du taux de décalage :

- (i) la mesure de l'efficacité de terminaison en fonction du contexte du stop de RF2 en remplaçant le codon stop (Pedersen *et al.* 1991),
- (ii) la mesure de l'efficacité de décodage relative d'un ARNt en remplaçant le codon stop affamé par des codons sens (Curran *et al.* 1989).

## **Une séquence Shine-Dalgarno**

En plus de ces signaux "classiques", la production de RF2 est dépendante d'une séquence AGGAGG située 3 nucléotides en 5' du codon glissant. Cette séquence riche en purines est très proche du consensus Shine-Dalgarno (SD). L'interaction de cette séquence avec l'ARNr 16S a été prouvée par des mutations et les mutations compensatoires, sur la séquence du gène codant pour RF2 et pour le gène codant pour l'ARNr (Weiss *et al.* 1988). Il a été suggéré que le rôle de cette séquence était de ralentir la progression du ribosome (Weiss *et al.* 1990a).

### **B.2.d. Antizyme**

L'Ornithine Décarboxylase (ODCase) est une enzyme clef de la synthèse des polyamines. La régulation de la quantité de cette enzyme chez les eucaryotes supérieurs se fait par l'antizyme de l'ODCase, qui cible l'ODCase vers le protéasome 26S (Atkins *et al.* 1996). L'expression de l'antizyme est dépendante d'un décalage de phase en +1. Le site de décalage de phase ressemble à celui de RF2, la séquence étant de type codon glissant-codon stop : UCC UGA U. Le codon stop peut être n'importe lequel des trois. Le mécanisme de décalage est de type Ty3,

c'est-à-dire que l'ARNt glissant lit (en quadruplet) ou provoque  
l'occlusion de la base suivant le dernier

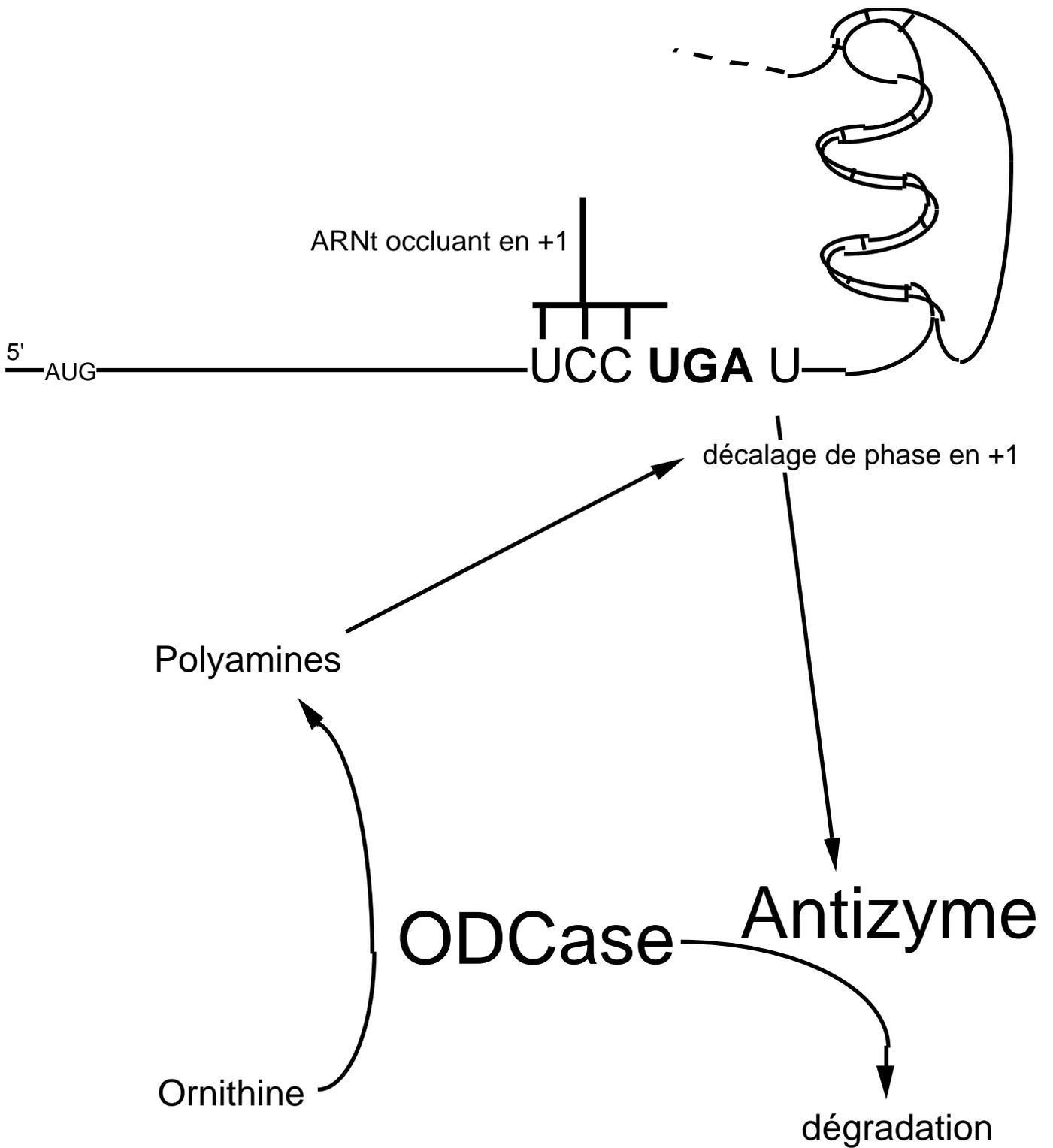


Figure 13 : Boucle de régulation négative de l'antizyme par décalage de phase en +1

codon de la phase 0 (ici, un codon stop). Un pseudo-nœud de type MMTV (avec une base adénine formant un coude) situé 56 nucléotides en aval, joue un rôle dans ce décalage de phase, même si des expériences *in vitro* suggèrent que l'influence de cette structure secondaire est limitée (Matsufuji *et al.* 1995). Ces structures secondaires influent donc sur les trois types de recodage : décalage de phase en -1 (cf. supra), décalage de phase en +1, et translecture de certains codons stops (comme le MoMuLV, cf. infra).

Le taux de décalage est régulé par le taux de polyamines : plus les polyamines sont abondantes dans la cellule, plus le taux de décalage de phase est fort (cf. § III.A.5). Cette régulation par les polyamines est, comme pour RF2, une régulation de type rétrocontrôle négatif : plus il y a d'antizyme, plus l'ODCase est dégradée, moins il y a production de polyamines, d'où diminution du taux de décalage de phase, et baisse de la production de l'antizyme (Fig. 13) (Matsufuji *et al.* 1995).

Le site de glissement de l'antizyme est fonctionnel *in vivo* chez la levure *S. cerevisiæ*. Le séquençage protéique de la protéine obtenue a donné un résultat surprenant : le site de glissement de l'antizyme est recodé avec une bonne efficacité chez *S. cerevisiæ*, mais un décalage de phase en -2 est utilisé au lieu du décalage de phase +1 (Matsufuji *et al.* 1996). Il faut peut-être noter que le décalage de phase en +1 de l'antizyme n'a jamais été prouvé *in vivo*.

### **B.2.e. EST3**

La découverte du décalage de phase en +1 utilisé dans l'expression d'un gène cellulaire (c'est-à-dire autre que de type viral ou transposon) chez la levure est très récente. Le gène *EST3* (Ever Shorter Telomeres), récemment identifié par un crible génétique serait une sous-unité de la télomérase (et aurait de ce fait une possible origine rétrovirale). Le mécanisme de décalage de phase est très semblable aux transposons Ty. Le site de glissement est AUA **CUU AGU UGA** (Morris *et al.* 1997). Il comporte le codon glissant CUU du Ty1 et le codon rare AGU du Ty3. Le taux de décalage mesuré est de 80-90 %, mais les auteurs font remarquer que le système de mesure est exprimé très fortement (promoteur du gène ADH1, plasmide multicopie). Ils font donc l'hypothèse que l'ARNt, déjà rare, le devient plus encore par titration. La rareté accrue de l'ARNt au site A augmente la possibilité qu'a l'ARNt au site P de se décaler dans la

phase +1. Cette explication est à pondérer puisque bien que l'ARNt cognat soit rare (Belcourt *et al.* 1990; Xu *et al.* 1990), le codon lui n'est pas rare (16% des codons sérine dans le génome de *S. cerevisiae* sont des codons AGU) : une surexpression d'un seul ARNm, même importante, peut-elle titrer fortement un ARNt dont le codon est déjà courant ? Indépendamment de ce système de mesure, les mêmes 7 nucléotides provoquent 8% de décalage de phase en +1 dans un système artificiel contenant un élément stimulateur (contexte Ty3) (Vimaladithan *et al.* 1994). On peut suspecter des influences des régions avoisinantes de l'heptamère glissant pour expliquer ce fort taux de décalage de phase dans le gène *EST3* (P.J. Farabaugh, communication personnelle).

On ne connaît pas actuellement le rôle physiologique de ce décalage de phase, puisque si on remplace au locus le gène sauvage par un gène modifié de façon à ne constituer qu'une seule phase, on ne semble pas modifier la croissance du mutant par rapport au sauvage. Cependant l'analyse faite en masse sur une population, au bout de 40 générations, ne permet pas de mesurer des différences fines, au niveau de la longueur des télomères, par exemple. Un contrôle traductionnel pourrait également se faire dans des conditions particulières de croissance (cf. § III.A.7). L'explication du rôle de ce décalage de phase et de sa régulation nécessite des expériences supplémentaires pour confirmer qu'il ne s'agit pas seulement d'une séquence fossile due à l'origine rétrovirale de l'activité ADN polymérase ARN dépendante (Lustig 1998).

### **B.3. Translecture d'un codon stop**

En plus des décalages de phase, permettant de supprimer des codons stops dans la phase initiale, deux parties d'un gène séparées par un codon stop sont parfois traduites en phase par suppression du stop. Le mécanisme est identique à la suppression génétique (par un suppresseur informationnel, ARNt muté), sauf qu'il se fait dans un contexte sauvage. On parle alors de translecture (readthrough) d'un codon stop. L'efficacité de cette translecture est le résultat de la compétition entre terminaison et poursuite de l'élongation, elle fait donc intervenir un contexte nucléotidique défavorable à la terminaison. La région indispensable à la translecture peut varier d'un nucléotide (Sindbis) à une région structurée de 57 nucléotides (MoMuLV). Un cas particulier de translecture sera évoqué plus loin. Il s'agit là d'une redéfinition d'un codon stop en codon pour le 21ème acide aminé, la sélénocystéine. L'encodage de cet acide aminé fait intervenir tout un système traductionnel parallèle : enzyme de modification, ARNt, facteur d'élongation, ...

Auparavant, je vais résumer ce que l'on sait de l'influence du contexte nucléotidique sur la terminaison, et donc sur la translecture.

#### **B.3.a. Contexte nucléotidique et terminaison**

Le terme de codon non-sens a, dans les premiers temps de l'élaboration du code génétique, servi à définir des codons qui "n'avaient pas de sens", c'est-à-dire ne codant pour aucun acide aminé. On a par la suite pu attribuer aux trois codons non-sens la fonction de terminer la traduction grâce à la reconnaissance par les facteurs de terminaison de la traduction (RF pour Release Factor). Cela a donc été une habitude liée au code génétique de désigner les signaux de terminaison de la traduction sous la forme de triplets UAG, UGA ou UAA. Mais, que les codons stops soient reconnus par un facteur de terminaison ou par l'ARNr 16/18S, la nécessité que le codon stop fasse 3 nucléotides ne lui est pas dicté par une reconnaissance par un ARNt. Des études systématiques pour comprendre ce qui spécifie exactement un stop, ce qui différencie un stop normal d'un stop soumis à translecture ont été menées (pour revue Tate *et al.* 1996). Je ne traiterai pas de l'influence du contexte en 5' des codons stops, mais des travaux récents semblent indiquer un effet des

ARNt ou des acides aminés de la chaîne peptidique en croissance (Mottagui Tabar *et al.* 1997).

### **L'environnement naturel des codons stop**

Une analyse statistique systématique sur les codons stop des gènes séquencés, révèle un fort biais sur le nucléotide +1 (situé juste en 3' du codon stop) chez *E. coli*, *Bacillus subtilis*, *S. cerevisiæ*, *Drosophila melanogaster* et chez les mammifères (Brown *et al.* 1990a; Brown *et al.* 1990b), ce qui confirme qu'un stop ne peut sans doute pas être réduit à un simple triplet. Le biais sur la base en +1 n'est pas le même en fonction des espèces. Par exemple chez *E. coli*, UAAU et UAAG sont les deux codons les plus utilisés, UAAC est utilisé à une fréquence proche de sa répartition théorique, alors que les 9 autres tétranucléotides semblent avoir été contre-sélectionnés. Chez les mammifères, la base +4 est plus souvent une purine (McCaughan *et al.* 1995). Ces résultats doivent être tempérés par le fait que d'une manière générale, les di-, tri- et tétranucléotides ne sont pas répartis équitablement dans le génome, régions non codantes incluses.

Des calculs similaires ont été réalisés sur des sous-ensembles de gènes fortement exprimés (selon leur biais de codon pour *E. coli*, ou selon leur fonction pour les mammifères). L'idée qui sous-tend cette démarche est qu'un gène qui a été contraint d'optimiser son élongation de la traduction, a dû également subir des contraintes quant à l'utilisation d'un codon stop fort. Ces analyses ont permis d'affiner le résultat évoqué ci-dessus. Par exemple chez *E. coli*, 80% des gènes fortement exprimés utilisent comme signal de terminaison de la traduction UAAU ou UAAG, la hiérarchie étant : UAAU>UAAG>UGAU>UAAA UAAC (Tate *et al.* 1992). Les 6 autres sont peu ou pas utilisés. Chez les mammifères, UAAG et UAGG sont sur-représentés, alors que UGAC, UGAU, UAGC et UAGU ne sont pratiquement jamais utilisés (McCaughan *et al.* 1995). Chez les mammifères, le biais s'étend jusqu'à la base +8, ainsi que sur le triplet -1 (McCaughan *et al.* 1995).

### **L'influence de ce contexte sur la terminaison**

Tout comme le biais de nucléotides entourant le codon initiateur s'est révélé biologiquement significatif, la signification du biais évoqué ci-

dessus peut-il être vérifié *in vivo* ? On peut mesurer l'influence du contexte nucléotidique du stop sur la terminaison, en fonction des données obtenues lors des analyses statistiques. C'est le système issu de la cible de recodage de RF2 qui a servi à la mesure de l'efficacité de terminaison, de manière indirecte par la variation de l'efficacité du décalage de phase +1 (cf. supra ou Pedersen *et al.* 1991). La corrélation avec l'analyse statistique est bonne en ce qui concerne UAAN et UGAN, mais pas en ce qui concerne UAGN, qui est un aussi bon terminateur que les deux autres, bien que totalement contre-sélectionné comme codon de terminaison chez *E. coli* (Poole *et al.* 1995). Mais ces tétra-nt UAGN sont également contre-sélectionnés tout au long du génome, ce qui biaise l'analyse statistique. Globalement, un facteur 50 sépare le meilleur codon terminateur UAAU du moins bon UGAC.

Chez les mammifères, l'effet de la base en +1 a été déterminé en mesurant la compétition entre le stop et l'incorporation de sélénocystéine sur un UGA redéfini (celui de l'iodothyronine déiodinase). Ce rapport varie de 1:3 (base +1 = U ou C) à 3:1 (A ou G). Ces mesures sont corrélées à un système *in vitro*, dans lequel l'efficacité de terminaison des autres codons stops UAGN et UAAN a été mesuré. Le même effet de la base en +1 a été observé. Relativement aux autres codons stops, UAA est 10 fois plus affecté par cette base +1 (McCaughan *et al.* 1995). Ces résultats sont encore une fois en accord avec l'analyse statistique.

En résumé : la quatrième base influence la terminaison de la traduction, et ceci influence la fréquence d'utilisation des tétranucléotides dans la terminaison des gènes cellulaires. L'influence est variable en fonction des organismes : U > G > A > C chez *E. coli* (Poole *et al.* 1995), A G >> C U chez les mammifères (McCaughan *et al.* 1995). Pour la levure, des études ont montré que pour UGA et UAA, l'influence de la 4ème position sur l'efficacité de la terminaison est G > U,A > C, alors que pour UAG l'ordre est U,A > C > G (Bonetti *et al.* 1995). On ne sait pas actuellement comment la base en 3' pourrait influencer sur la terminaison : par interaction du complexe de terminaison avec un tétra-nt, ou par influence de la quatrième base sur la structure tridimensionnelle du codon stop dans le site A,...

### **B.3.b. TMV**

Le TMV (virus de la mosaïque du tabac) est un virus de plante (tabac et tomate) à ARN + (servant de messenger), d'environ 6400 nucléotides, dont la réplicase est exprimée sous la forme d'un produit de fusion de 183 kDa résultant de la translecture d'un codon ambre (Pelham 1978). Le remplacement du codon ambre UAG par un codon UAU tyrosine réduit considérablement son pouvoir infectieux; les rares foyers infectieux contiennent un mélange de virus muté (tyrosine) et de virus révertant spontanés (codon tyrosine -> codon ochre UAA) (Ishikawa *et al.* 1986). La délétion du codon stop inhibe totalement le pouvoir infectieux, sans possibilité de réversion (Ishikawa *et al.* 1986).

Il a été montré que la séquence CAA **UAG** CAA UUA, également conservée dans d'autres virus de plantes (cf. ci-dessous), était suffisante pour promouvoir 5% de traduction du codon stop mesuré avec un système rapporteur dans des protoplastes de tabac (Skuzeski *et al.* 1990). Une étude plus fine a montré que la séquence nécessaire et suffisante permettant un taux sauvage de translecture est UAG CAR YYA, où le codon stop peut être remplacé par UGA ou UAA mais avec trois fois moins de translecture (Skuzeski *et al.* 1991). Un ARN<sub>t</sub><sup>Tyr</sup>, avec un anticodon 3'A G5' est capable *in vitro* de supprimer les stops UAG et UAA dans ce contexte (Zerfass *et al.* 1992). La cible de translecture du TMV a également été étudiée dans des systèmes hétérologues, comme des cellules de mammifères en culture, des extraits de réticulocytes de lapin ou de germe de blé. Ces études supplémentaires indiquent un rôle synergique du codon immédiatement en aval du stop : CAA. Dans le cas des cellules de mammifères en culture, le taux de translecture atteint est de 2% (J.-P. Rousset, M. Cassan, manuscrit en préparation); dans les réticulocytes de lapin, 31% en présence d'un ambre, 10% en présence d'un ochre et 69% pour un opale (Valle *et al.* 1992). Au cours de ma thèse j'ai étudié cette cible de translecture dans les cellules de mammifère en culture, dans la levure et chez *E. coli*.

Des virus du même type présentent des sites de translecture similaires (pour revue Miller *et al.* 1995):

- soit CAA **UAG** CAA UYA : pour le Beet Necrotic Yellow Vein Mosaic Virus et Turnip Yellow Mosaic Virus,

- soit AGA **UAG** CAA UUA pour le Tobacco Mild Green Mosaic Virus. A propos de ce dernier, il a été suggéré que l'arginine (AGA) et la glutamine (CAA) étaient deux acides aminés ayant la même composition chimique (Skuzeski *et al.* 1991), et suggère un rôle du dernier acide aminé sur la translecture. Il ne semble pas que ce soit le cas dans les cellules de mammifère en culture, où c'est le contexte nucléotidique qui joue un rôle (J.-P. Rousset & M. Cassan, manuscrit en préparation).

### **B.3.c. Autres virus de plantes**

De nombreuses autres classes de virus utilisent la translecture d'un codon stop, soit pour synthétiser une extension C-terminale de la protéine de surface, soit pour exprimer une activité catalytique de type polymérase. L'environnement des ces stops est très variable en fonction du type viral, mais semble conservé à l'intérieur d'un même groupe (revue dans Miller *et al.* 1995).

En ce qui concerne les lutéovirus exprimant ainsi une extension C-terminale de leur protéine de surface, le site de translecture est extrêmement conservé : CC AAA **UAG** GUA GAC, et il est suivi quelques nucléotides plus loin par des répétitions ininterrompues de type CCX XXX (et donc des répétitions alternées de prolines au niveau protéique). Au moins l'une de ces séquences est de type CCCCA. Les mécanismes à l'œuvre pour promouvoir la translecture au niveau de ce stop en contexte "fort" (**UAGG**) ne sont pas connus, mais ne semblent pas faire intervenir de structures secondaires (Miller *et al.* 1995).

Pour exprimer leur polymérase, les carmovirus utilisent une translecture d'un codon stop : AAA **UAG** G suivi plusieurs dizaines de nucléotides en aval par une séquence CCCCA.

### **B.3.d. Sindbis**

Les alphavirus, virus animaux, dont le Sindbis est le prototype, utilisent un stop en environnement défavorable **UAGC**, qui est suffisant pour permettre 10% de translecture *in vivo* et produire leur polymérase. L'inefficacité de ce stop a été testée *in vitro* dans différents contextes (Li *et al.* 1993). Le sindbis est un virus capable de se reproduire dans deux types d'hôtes : animaux et insectes. Le signal on ne peut plus minimal de

translecture fonctionne également dans les insectes, comme le prouvent des études dans des cellules d'insectes en culture (Jean-Pierre Rousset, résultats non publiés)

D'autres types de virus ont également ce type de signal minimal pour programmer une translecture, comme par exemple le Tobacco Rattle Virus (Pelham 1979).

### **B.3.e. MoMuLV**

Chez certains rétrovirus le site de translecture est suivi d'une structure secondaire stimulatrice. Le virus le mieux étudié de ce genre est le MoMuLV (Virus de Moloney de la leucémie murine). Un codon **UAGG** est décodée par une glutamine (Yoshinaka *et al.* 1985) pour permettre l'expression de la protéine de fusion gag-pol à un taux d'environ 5% *in vitro* (Wills *et al.* 1991; Feng *et al.* 1992). Une région de 57 nucléotides en 3' du stop est nécessaire à cette translecture : la structure secondaire qui suit le stop transluc est un pseudo-nœud, semblable à celui du MMTV (Wills *et al.* 1991; Felsenstein *et al.* 1992); cependant l'un ne peut remplacer l'autre sans perdre 70 à 90 % de son effet stimulateur (Wills *et al.* 1994).

La région intercalante **...UAGGgAGgUCa...** de 8 nucléotides non appariés entre le stop et le pseudo-nœud joue également un rôle pour stimuler la translecture (Feng *et al.* 1992). Des études *in vitro* ont cerné plus précisément les contributions de chacun des nucléotides : je les ai indiqués ci-dessus en capitales. Les nucléotides indiqués en minuscules ont peu ou pas d'effet sur la translecture. Les nucléotides en gras **U** et **C** sont absolument indispensables à la translecture. D'autres nucléotides *a priori* non appariés selon la structure proposée se sont avérés indispensables dans la boucle 2 du pseudo-nœud, et parmi ceux-là, certains sont hypersensibles à la nucléase S1, ce qui suggère qu'ils sont particulièrement exposés dans la structure (Wills *et al.* 1994). Cette situation est différente des pseudo-nœuds des rétrovirus à décalage de phase en -1, dont les nucléotides non appariés sont peu sensibles aux substitutions (Brierley *et al.* 1991); et les nucléotides de la boucle 2 du pseudo-nœud du MMTV qui sont tous également accessibles à la nucléase S1 (Chen *et al.* 1995). En faisant le parallèle avec les rétrovirus utilisant le décalage de phase en -1, qui utiliseraient leur pseudo-nœud pour stopper le ribosome, peut-on ici expliquer la translecture avec le même

type de mécanisme ? On comprend comment une pause avec les sites A et P occupés par des ARNt susceptibles de glisser va favoriser cinétiquement le décalage de phase par rapport à l'élongation normale. En revanche, comment une pause peut-elle favoriser le décodage d'un stop par un ARNt suppresseur par rapport aux facteurs de terminaison ? Le décodage d'un codon stop (site A vide) ou d'un codon affamé, constitue à lui seul déjà une pause (Wolin *et al.* 1988). Augmenter la durée de cette pause peut-elle favoriser un suppresseur par rapport à la terminaison ? Il serait plus tentant, là encore, de proposer une interaction entre le pseudo-nœud et un facteur *trans* (facteur soluble, protéine ou ARN du ribosome) qui altérerait le processus de la terminaison. L'existence de bases indispensables, exposées en surface, et ne faisant pas partie de la structure proprement dite favorise ce type d'hypothèse.

### **B.3.f. Gènes cellulaires**

Très peu de gènes cellulaires faisant appel à la translecture sont actuellement caractérisés. Quelques exemples peu documentés ont été relatés : pour une protéine de la famille des zéïnes chez le maïs (Wandelt *et al.* 1989), l'autre pour une forme légèrement allongée de la  $\beta$ -hémoglobine de lapin apparaissant *in vitro* et *ex vivo* dans des réticulocytes de lapin en culture (Geller *et al.* 1980). Les exemples suivants sont mieux caractérisés, mais là non plus aucune étude n'a confirmé ces résultats, et aucune analyse de contexte et/ou de fonction n'a été entreprise à ce jour. Une exception notable sera détaillée à la fin de ce paragraphe, et concerne l'utilisation d'un codon stop pour permettre l'incorporation du 21<sup>ème</sup> acide aminé : la sélénocystéine, qui se retrouve des procaryotes à l'homme.

#### ***E. coli* : gènes nécessaires à l'assemblage des pili**

Un gène cellulaire dont l'expression est dépendante de la translecture d'un codon stop a été rapporté dans des souches de *E. coli* entérotoxiques (ETEC) (Jalajakumari *et al.* 1989). Cette translecture est nécessaire pour la production d'une protéine de 104 kDa, dont la fonction est inconnue, mais qui participe à la formation des pili nécessaires à l'adhésion des ETEC à la surface intestinale. Ce système est peu caractérisé, mais la souche ayant servi à isoler ce gène contiendrait un suppresseur naturel

(résultats non publiés, cités dans Jalajakumari *et al.* 1989). L'expression de cette protéine est possible dans une souche dépourvue de suppresseurs, ou contenant un suppresseur différent, mais on n'observe alors aucune formation de pili (résultats non publiés, cités dans Jalajakumari *et al.* 1989).

### ***Kelch***

Le gène *kelch* de *Drosophila melanogaster* comporte deux phases ouvertes de lecture ORF1 et ORF2, séparées par un codon UGA. Un produit de fusion (180 kDa) peut être détecté par des anticorps dirigés vers ORF1 (80 kDa), ce qui suggère une translecture du codon stop AUG **UGA** AUG (Xue *et al.* 1993). Les auteurs n'ont pas examiné la possibilité, soit d'une insertion de Sélénocystéine (cf. infra), soit de saut du ribosome de façon à éviter la lecture du stop (cf. ci-dessous, § B.4). Le rapport entre les produits d'ORF1 et le produit de fusion varie en fonction des tissus étudiés. La fonction cellulaire défective dans un mutant peut être restaurée par l'ORF1 seule. Les deux ORF mises en phase ne restaurent que partiellement les phénotypes de transport intercytoplasmique du mutant *kelch* (Robinson *et al.* 1997).

### ***Headcase***

Chez la drosophile, un deuxième gène dont l'expression dépend de la translecture d'un codon stop a été identifié. Le gène *hdc* (headcase) impliqué dans le développement chez la drosophile contient un codon UAA. La translecture de ce codon stop est obligatoire pour permettre de restaurer les fonctions sauvages chez un mutant *hdc*, puisqu'une inactivation de la partie 3' du gène n'y suffit pas (Steneberg *et al.* 1998). L'éventualité d'un saut de ribosome ou de l'insertion de sélénocystéine, comme pour le gène *kelch* ci-dessus, n'a pas été étudiée.

### **Sélénocystéine**

L'encodage du 21ème acide aminé ne se fait pas grâce à un recodage de type faux-sens (c'est-à-dire l'utilisation d'un codon sens pour une nouvelle signification), mais par la redéfinition d'un codon stop UGA.

Cette nouvelle fonction nécessite plus qu'une simple translecture par un ARNt portant une sélénocystéine; il faut en plus un facteur d'élongation type EF-1A/EF-Tu spécifique : SELB chez *E. coli*, et inconnu pour le moment chez les eucaryotes; il faut également un mécanisme de charge spécifique de l'ARNt<sub>SeI</sub>, celui-ci est d'abord chargé par une sérine, qui sera ensuite modifiée en sélénocystéine (Forchhammer *et al.* 1991).

Chez les procaryotes, ce qui différencie un codon stop UGA d'un codon Sélénocystéine est une tige-boucle située juste après le codon UGA (Zinoni *et al.* 1990; Berg *et al.* 1991; Heider *et al.* 1992). Cette tige-boucle est capable de redéfinir de la même manière un codon UAG, UAA (Heider *et al.* 1992) ou UGC (Baron *et al.* 1990) dans des systèmes artificiels. Chez les mammifères, il faut également une structure secondaire nommée SECIS, mais elle est située dans les régions 3' non traduites de l'ARNm. Une trop grande proximité entre SECIS et le codon stop, ce qui est le cas juste avant la région 3' non traduite, fait que celui-ci est reconnu comme un réel codon stop provoquant la terminaison de la traduction (Williams *et al.* 1992; Vendeland *et al.* 1995; Martin *et al.* 1996). Un élément SECIS peut redéfinir jusqu'à 10 codons UGA disséminés tout au long d'un messenger, et ceux-ci sont parfois séparés par un seul codon sens (Hill *et al.* 1993). Le mécanisme sous-jacent est très mal compris pour l'instant, mais il semble que la structure secondaire lie SELB pour amener le facteur d'élongation et l'ARNt<sub>SeI</sub> vers le ribosome. On ne comprend pas bien comment, dans le cas des mammifères, un seul SECIS peut amener ce complexe ternaire à tous les ribosomes en élongation, au niveau de chaque stop redéfini en même temps. On peut imaginer que les ribosomes sont redéfinis une fois pour toutes lors de l'initiation, et que SECIS ne sert qu'au moment de la terminaison. Autrement dit, ces ribosomes, reprogrammés pour cet ARNm, ne comprendraient pas les codons non-sens comme signal de terminaison mais comme des codons sélénocystéine, et ce n'est que la proximité de SECIS (dans la région 3' non traduite) qui reprogramme les ribosomes de manière à interpréter le dernier codon non-sens comme codon stop (Gesteland *et al.* 1996).

### **B.3.g. Les codons ochre**

D'une manière générale, les codons ochre UAA sont peu sujets à la suppression informationnelle. Les cellules de *E. coli* contenant un suppresseur d'ochre, même faiblement efficace, poussent très lentement (Tate *et al.* 1996). *A contrario*, des ARNt suppresseurs opale ou ambre

peuvent être présents dans des populations coliforme à l'état naturel (Tate *et al.* 1996), ou dans des souches de laboratoires telles HB101 ou JM105 (Sambrook *et al.* 1989), sans contre-sélection ou avantage sélectif. Les codons ochre ont rarement été détectés comme cible de translecture. Le seul exemple naturel connu est celui de *hdc* chez la drosophile (Steneberg *et al.* 1998). Chez *E. coli*, la reconnaissance de ce stop par les deux facteurs de terminaison augmente peut-être l'avantage de la terminaison par rapport à la suppression au niveau des codons ochre, et pourrait expliquer le comportement différent de UAA par rapport à la translecture. Cette explication n'est pas applicable aux eucaryotes, puisque eRF reconnaît les trois codons stop.

Pourtant la quasi inexistence de codons ochre translus n'est pas une contrainte fonctionnelle : dans nombre de systèmes de translecture, le codon stop impliqué (UAG ou UGA) peut être remplacé par un codon UAA. Un TMV mutant, dont le codon UAG a été remplacé par UAA, garde son potentiel sauvage d'infectivité et de réplication *in vitro* et *in vivo* (Ishikawa *et al.* 1986). Mais le codon UAA est trois fois moins bien supprimé *in vivo* et *ex vivo* dans des protoplastes de tabac en culture (Skuzeski *et al.* 1991), et *in vitro* dans un système hétérologue de réticulocytes de lapin (Valle *et al.* 1992). Chez la levure, un codon ochre est translus dans le contexte issu du TMV avec presque la même efficacité que le codon ambre (Bonetti *et al.* 1995). De même les codons UGA de la sélénocystéine (Heider *et al.* 1992) ou UAG du MoMuLV (Feng *et al.* 1989; Wills *et al.* 1991) peuvent être remplacés par un codon ochre en gardant leur propriété de recodage. En revanche, le stop **UGAC** du Sindbis est codon spécifique. Si l'on remplace le codon opale par un autre codon stop, le produit de translecture disparaît complètement lors d'une traduction *in vitro* en réticulocytes de lapin (Li *et al.* 1993). Ceci s'explique sans doute par le contexte minimal de ce site de recodage, et par le fait que des contraintes différentes existent entre les trois codons stops pour définir un stop fort (Bonetti *et al.* 1995).

Pour expliquer cette quasi-absence de système de recodage faisant intervenir des codons ochre, il a été suggéré que la cellule a pu utiliser, ou utilise toujours, les codons de terminaisons avec des sens différents : un codon UAA voudrait dire "toujours stop", alors que les codons UGA et UAG pourraient permettre la production d'extensions C-terminales à certaines protéines dans certaines conditions. Cette fonction aurait pu précéder la fonction d'épissage alternatif (Geller *et al.* 1980).

Un autre système de recodage ressemble également à l'épissage, puisqu'il permet de ne pas traduire une partie d'un ARNm, en sautant une partie du message sur l'ARNm : le saut de ribosome.

#### **B.4. Saut de ribosome**

Le saut de ribosome (“hopping”) a été étudié chez le gène *60* du bactériophage T4. Le saut de ribosome est à différencier de l'épissage, puisqu'il se passe de manière dynamique, au niveau traductionnel. Il fait intervenir deux sites : un de décollage (“take off site”) et l'autre d'atterrissage (“landing site”). Ce sont deux codons identiques séparés par une distance de 50 nucléotides, ce qui permet d'éviter un codon stop dans cette séquence sautée, et de changer de phase. La fréquence de saut de ribosome est de l'ordre de 100% (Weiss *et al.* 1990a). Cette efficacité requière 4 éléments :

- 2 codons glycine identiques GGA
- un stop entre les deux, ainsi qu'une tige-boucle
- une distance optimale de 50 nucléotides
- la séquence de la chaîne peptidique en croissance (16 aa) est également importante, ce qui suggère une interaction avec le ribosome, le tunnel de sortie du peptide contenant jusqu'à 30 aa.

Ce système ne semble pas être totalement marginal, puisque d'autres cas potentiels ont été rapportés, pour *carA* de *Pseudomonas aeruginosa* (Wong *et al.* 1990), pour une lectine-like de *Prevotella loescheii* (Manch-Citron *et al.* 1994), une endoglucanase de *Bactéroïdes ruminicola* (Matsushita *et al.* 1991), ainsi que dans certains systèmes artificiels de mesure (Gesteland *et al.* 1996), ou aux alentours de certains codons affamés (John Gallant, communication personnelle).

On peut imaginer comme dernier cas de recodage les erreurs faux sens, que je n'évoquerai pas car elles restent très peu documentées étant très difficiles à identifier expérimentalement. Une des approches envisagées repose sur les gels de protéines bi-dimensionnels différentiels; du matériel satellite autour des points caractérisés de certaines protéines provient de la migration différentielle de produits mal-traduits. Cette méthode permet de mettre en évidence les changements de charge non conservatifs dans une protéine, soit un tiers des erreurs faux sens (pour

revue Gallant *et al.* 1997). Une autre approche expérimentale utilisant la spectrométrie de masse, permet de quantifier des produits de traduction d'une protéine purifiée dont le contenu en acides aminés diffère ne serait-ce qu'à une seule position (Calderone *et al.* 1996). Cet outil n'a pas été développé dans ce sens pour le moment.

Pour résumer cette introduction exposant les différents types de recodage, on peut remarquer que l'appareil traductionnel réagit à un contexte particulier que l'on peut assez bien définir. Il s'agit d'un site favorable au recodage : codon stop faible pour la translecture, heptamère glissant pour le décalage de phase en -1, codon décodé par un ARNt glissant ou occluant pour le décalage de phase en +1. Ces événements de recodage sont cinétiquement favorisés par rapport au décodage grâce à une pause du ribosome : stop faiblement reconnu en ce qui concerne la translecture, structure secondaire en ce qui concerne le décalage de phase en -1, codon affamé pour le décalage de phase en +1. La compréhension de ces systèmes de traduction alternative a été rendue possible grâce à la mesure de l'efficacité des sites sauvages de recodage, et des variations mesurées pour des mutants de ces sites. Dans un premier temps, ces taux de recodage ont été mesurés grâce à la détection en Western-Blot de produits de traduction synthétisés *in vitro*. Ceci a permis de définir les grandes règles sus-citées. Mais l'utilisation de systèmes enzymatiques, plus sensibles, permet de détecter des variations plus fines du taux de recodage, comme on le verra au deuxième chapitre. L'utilisation *in vivo* de tels systèmes permet en outre de mesurer directement l'effet, même faible, de facteurs *trans* (drogues, mutants,...), comme on le verra dans le troisième chapitre. C'est à ces fins que j'ai entrepris la construction chez la levure *S. cerevisiae* d'un système de mesure du recodage, que je vais décrire maintenant.



# Chapitre I. Construction d'un outil de mesure du recodage

## I.A. État des lieux en 1993

Dans une revue sur le recodage, à propos des études portant sur la quantification du décalage de phase -1, John Atkins écrivait : “*Ces résultats devraient être interprétés avec précaution, vu que toutes les études du décalage de cadre de lecture des eucaryotes supérieurs avec séquences altérées ont été réalisées dans des systèmes acellulaires avec un lysat de réticulocytes; il est probable qu'il y ait des différences dans le nombre de ribosomes chargés par message, et peut-être dans la balance des ARNt, de cellules moins spécialisées in vivo. Des expériences avec des cellules en culture sont clairement nécessaires, surtout puisque des expériences de synthèse protéique en lysats de réticulocytes, ou des études dans des cellules de levure, ont montré peu ou pas du tout d'effet des séquences 3' sur le décalage de phase du VIH (Wilson et al. 1988)*” (Atkins et al. 1990)

En effet, l'article cité par John Atkins présente des résultats de mesure de recodage du VIH-1 dans des systèmes de traduction *in-vitro* de mammifères ou *in vivo* de levure, par révélations immunologiques. Dans ces études, les auteurs n'ont pas détecté de différence quantitative entre le décalage de phase en -1 de la séquence heptamérique U UUU UUA seule, ou avec la séquence palindromique susceptible de former une tige-boucle en aval (Wilson et al. 1988b). D'autres expériences semblaient confirmer cela par l'absence d'effets sur le décalage de phase de mutations même importantes de la tige-boucle (Yelverton et al. 1994). Des études ultérieures ont montré que dans un système enzymatique *ex vivo* dans des cellules en culture, cette séquence avait en fait un rôle stimulateur (Parkin et al. 1992). De plus, d'une manière générale, les résultats de quantification *in vivo* sont toujours moins élevés que ceux obtenus *in vitro* (Reil et al. 1990; Parkin et al. 1992; Kollmus et al. 1994; Miller et al. 1995; Dinman et al. 1997b; Grentzmann 1998). Nous avons donc envisagé la construction de systèmes de mesure *in vivo*, grâce à l'utilisation d'un système rapporteur enzymatique extrêmement sensible : le gène *luc*, produisant la luciférase de *Photinus pyralis* (Cassan et al. 1990; Reil et al. 1990). Nous disposions déjà d'outils de ce type, fonctionnels dans les cellules de mammifère en culture (Cassan et al. 1990; Berteaux et al. 1991; Cassan et al. 1994). Nous avons voulu exploiter un organisme permettant d'aborder des études génétiques; j'ai

donc entrepris la construction d'un système de mesure similaire utilisable en parallèle dans les cellules de mammifères en culture et dans la levure *S. cerevisiae*.

## **I.B. Le système de mesure**

Les spécificités de notre système de mesure sont les suivantes :

(i) le gène reporter principal est *luc*; l'activité enzymatique de la luciférase est mesurée par la production de photons en présence de son substrat naturel, la luciférine, et d'ATP; elle est exprimée en unités arbitraires,

(ii) cette activité enzymatique nous permet de calculer un taux de recodage en la rapportant à l'activité obtenue dans un témoin en phase, dont la séquence est conçue en tenant compte du système de recodage étudié : la protéine de fusion obtenue aura exactement la même séquence protéique que celle servant à la mesure en (i), mais pour 100% des produits de traduction,

(iii) les deux activités luciférase, construction test (i) et témoin en phase (ii), sont normalisées par rapport à un témoin interne, *lacZ*, qui permet de s'affranchir de tout ce qui pourrait modifier les activités mesurées : nombre de copies du plasmide, efficacité de traduction, de transfection, et surtout stabilité différentielle des messagers avec ou sans stop précoce. Les deux cistrons *lacZ* et *luc* sont présents sur le même ARNm, mais dans deux phases différentes, mimant la disposition des gènes *gag* et *pol* des rétrovirus (Fig. 1 de l'article I).

Je vais d'abord expliquer ma démarche et les raisons des choix que nous avons faits. Les résultats obtenus ont été publiés dans un article joint. Je commenterai à la suite plus en détail les résultats obtenus. Dans une quatrième partie j'exposerai des résultats non publiés obtenus avec le même système chez *E. coli*.

## **I.C. Validation du système rapporteur**

Pour valider cette approche dans *S. cerevisiae*, nous avons utilisé ce système pour quantifier le taux de décalage de phase provoquée par la

séquence heptamérique du transposon de levure Ty1, et en le comparant à des chiffres précédemment publiés (Belcourt *et al.* 1990).

## **I.D. Étude de cibles de recodage hétérologues chez *S. cerevisiae***

Pour savoir si la levure *S. cerevisiae* était utilisable pour étudier des cibles de recodage issues de système eucaryotes supérieurs, nous avons mesuré la translecture d'un codon UAG issu du TMV, et le taux de décalage de phase en -1 issu du VIH-1. Pour ces deux cibles, ainsi que pour certains mutants, les taux de recodage *in vivo* sont connus dans des cellules de mammifères en culture (Cassan *et al.* 1994; J.-P. Rousset et M. Cassan, manuscrit en préparation).

La levure *S. cerevisiae* peut interpréter un heptamère glissant comme cible de recodage, provoquant un décalage en -1 et le taux de ce glissement est dépendant d'un pseudo-nœud en aval pour le ScV (Dinman *et al.* 1991; Dinman *et al.* 1992). Mais les mesures obtenues précédemment chez la levure *S. cerevisiae* en ce qui concerne le VIH-1 avaient mené à la conclusion que le site heptamérique était suffisant (Wilson *et al.* 1988b). Dans tous les autres systèmes où la dépendance du décalage de phase vis-à-vis de la tige-boucle a été mesurée : lignées aviaires, simiennes (Parkin *et al.* 1992), murines (Rousset et Cassan, résultats non publiés), humaines (Cassan *et al.* 1994; Reil *et al.* 1994), le rôle de la tige-boucle comme stimulateur du décalage de phase d'un facteur deux à cinq a été démontré. Nous avons voulu vérifier à l'aide de notre système si la levure était un cas particulier par rapport à cet événement de recodage.

Dans les deux cas, TMV et VIH-1, nous avons mesuré le taux de recodage de ces cibles, ainsi que de certains mutants, en parallèle dans les cellules de mammifères en culture et la levure *S. cerevisiae*, en utilisant les mêmes vecteurs.

## **I.E. ARTICLE I**

<http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m&form=6&uid=95303640&Dopt=b>

## I.F. Décalage de phase en +1 chez *S. cerevisiae*

Les valeurs que nous calculons pour le taux de décalage de phase en +1 du Ty1 sont semblables à celles précédemment publiées (Belcourt *et al.* 1990), ce qui valide l'utilisation de notre double système de rapporteurs *lacZ-luc*. Ce type de construction a de plus l'avantage de mettre le site de recodage dans une position similaire à celle que nous voulons modéliser : *lacZ* et *luc* miment l'organisation des gènes *gag-pol*. Malgré cette cohérence entre les différents systèmes de mesure, une mesure de l'efficacité du décalage de phase en +1 au niveau du Ty1 lui-même a révélé un taux de 3% (Kawakami *et al.* 1993). Plusieurs différences entre le système endogène et le système rapporteur pourrait expliquer cette "surestimation" mesurée dans les systèmes rapporteurs :

(i) un élément inhibiteur actif à distance, dont l'existence n'aurait pas été mise en évidence durant les travaux initiaux (Clare *et al.* 1988),

(ii) une influence du taux de traduction sur le décalage de phase +1 du Ty1, avec des ribosomes s'accumulant au niveau d'un codon rare (Wolin *et al.* 1988; ou cf. § II.D); la traduction au niveau du site de glissement pourrait être perturbée en fonction du nombre de ribosomes bloqués en amont sur l'ARNm; cette hypothèse pourrait être testée en faisant varier le taux de traduction de notre ARNm rapporteur, en modifiant le contexte de l'AUG initiateur, ou en créant des codons rares au long de la séquence de l'ARNm,

(iii) une influence du taux d'expression du gène rapporteur comportant le codon rare AGG; on peut imaginer qu'une production importante d'un ARNm contenant le codon rare AGG va titrer l'ARNt correspondant, le rendant de ce fait encore moins disponible pour décoder AGG au niveau du site glissant (Morris *et al.* 1997); des expériences sont en cours pour étudier cette cible Ty1 sur un plasmide analogue, mais à haut nombre de copies dans la levure, ce qui devrait pouvoir permettre de tester cette hypothèse,

(iv) un effet du biais de codons "local": le contenu en codons affamés AGG de notre système rapporteur est nul, alors que le Ty1 endogène en comporte 4; il a été montré dans des cellules perméabilisées que les ARNt sont conservés et aminoacylés dans un environnement local (Stapulionis *et al.* 1997). Ce phénomène dit de "channeling" permettrait le recyclage local de tous les acteurs de la traduction, comme les ARNt déacylés et les

aminoacyl-ARNt-synthétases. On peut imaginer que la traduction d'un ARNm contenant plusieurs codons limitants identiques ne sera limité dans son utilisation que lors de la première rencontre de ce codon. Un second codon serait immédiatement "nourri" par l'ARNt issu du channeling, et n'aura donc plus un effet de codon "affamé". Le taux de décalage de phase en +1 du Ty1 endogène, stimulé par un codon affamé, mais fréquent dans sa séquence codante, pourrait alors être moins élevé que celui que nous mesurons sur un ARNm ne possédant qu'un seul codon affamé de ce type. Ceci peut-être testé en augmentant le nombre de codons rares AGG dans notre système rapporteur. Cette construction répondrait dans le même temps aux hypothèses (ii) et (iii) ci-dessus. Ces hypothèses, très spéculatives, vont à l'encontre des résultats obtenus chez *E. Coli*, ou au contraire plus un codon rare est présent sur un ARNm, plus cela provoque des erreurs faux-sens au niveau de ces codons (Calderone *et al.* 1996). Les différences de compartimentalisation, et donc peut-être l'absence de "channeling" chez *E. coli* limitent la portée de ces contradictions. Si une construction contenant plus ou moins de codons rares permet moins ou plus de décalage de cadre, cela permettrait une approche génétique pour étudier le phénomène de channelling, qui s'étudie pour le moment dans des systèmes complexes (Stapulionis *et al.* 1997).

(v) on ne peut pas exclure que la mesure par détection immunologique du ratio TYA-TYB/TYA soit peu précise, ou que les techniques soient inadaptées (transfert des protéines en Western-blot différent selon la taille et/ou la charge) (Kawakami *et al.* 1993).

## I.G. Levures et cibles issues d'eucaryotes supérieurs

### I.G.1. TMV

La séquence CAA **UAG** CAA du TMV est fortement translue chez *S. cerevisiae* (20%). La souche Fy1679-18B utilisée pour ces mesures est la même que celle ayant servi au séquençage systématique du génome de *S. cerevisiae*. Ce séquençage n'a révélé aucun suppresseur connu, et ne présentait aucun phénotype connu d'une souche [psi+] (eRF3 sous forme prion, présentant un phénotype suppresseur). Le contexte du stop semble en être la seule explication, puisque le mutant TMG CAG **UAG** CAG a une translecture diminué d'un facteur 10.

Depuis ces mesures, une analyse des contextes des codons stops de la levure a révélé pour cette même cible CAA **stop** CAA des taux de terminaison du même ordre (Bonetti *et al.* 1995). L'ARNt impliqué dans cette forte suppression pourrait être l'ARNt<sub>Gln</sub> avec anticodon 5'CAG<sub>3</sub>', puisque celui-ci peut supprimer des codons ambre à l'état sauvage (Weiss *et al.* 1987b). J'exposerai également dans le troisième chapitre de cette thèse un crible pour trouver de nouveaux facteurs influant sur cette translecture.

### **I.G.2. VIH-1**

L'heptamère glissant est susceptible à lui seul de faire changer le ribosome de phase à un taux non négligeable, ce qui confirme son caractère exceptionnellement glissant (Jacks *et al.* 1988a; Wilson *et al.* 1988b), dû sans doute à sa nature monotone, qui facilite le réappariement en phase -1 des ARNt. En opposition avec les résultats précédemment publiés (Wilson *et al.* 1988b), nous démontrons ici que la levure interprète comme les cellules de mammifères en culture les messages contenus au niveau du site de décalage en -1 du VIH-1. Le décalage de phase est augmenté par une séquence palindromique en aval, susceptible de prendre une conformation de type tige-boucle. De sa stabilité dépend sans doute sa faculté à freiner le ribosome, et donc à provoquer le décalage de phase en -1. C'est ce que nous avons voulu étudier par la suite, dans le deuxième chapitre de cette thèse.

### **I.H. Le recodage chez les procaryotes**

Nous avons également mesuré, avec les mêmes outils, le taux de recodage chez *E. coli*, souche RM4102. Le principe est exactement le même, puisque le promoteur SV40 permet l'expression chez *E. coli* (Camonis *et al.* 1990). Les résultats n'ont pas été intégrés à l'article I, et sont résumés ci-dessous.

La mesure de translecture des deux variants du stop issu du TMV donne des valeurs d'environ  $10^{-4}$ , ce qui est une terminaison efficace. Pourtant le C en aval du stop devrait en faire un stop faible (cf. introduction ou Poole *et al.* 1995); la faible efficacité de translecture peut donc indiquer plus une déficience dans la suppression qu'une très bonne terminaison.

Cette séquence, qui provoque une forte translecture chez tous les eucaryotes étudiés, n'est donc pas interprétée de la même façon chez *E. coli*. Ceci peut être dû à plusieurs raisons. La terminaison semble être une étape peu conservée entre procaryotes et eucaryotes, tant au niveau du nombre de facteurs de terminaison, que de la présence d'une étape supplémentaire de relargage des ribosomes de l'ARNm par RRF (ribosome release factor, Janosi *et al.* 1998), inconnu chez les eucaryotes à ce jour. Mais ceci peut refléter autant notre ignorance que des différences réelles de l'appareil de traduction. Une autre différence peut provenir de l'existence ou non d'un ARNt suppresseur naturel, comme celui qui a été identifié chez la levure (Weiss *et al.* 1987b). Le système est encore trop peu connu chez les eucaryotes pour pouvoir expliquer ce non recodage chez *E. coli*.

Le décalage de phase en +1 de la séquence glissante du Ty1 est de l'ordre de  $10^{-3}$  chez *E. coli*, ce qui reflète sans doute des différences dans la rareté de l'ARNt cognat pour AGG (le codon lui-même est extrêmement rare chez *E. coli*), puisque l'ARNt glissant de CUU à UUA l'est aussi chez *E. coli* (Curran 1993).

En ce qui concerne le VIH-1, le décalage de phase est, comme chez la levure et les mammifères, dépendant de la séquence située en 3'. Sans cette séquence, le taux de décalage de phase en -1 est de 1% (pAC89) au lieu de 6,5% (pAC1789). Une séquence dont le palindrome est en partie détruit (pAC1815) diminue le décalage de phase en -1 de 40% (4,5%). D'après ces résultats, *E. coli* interprète donc ce site de décalage de phase en -1 essentiellement de la même façon que les eucaryotes. Cette conservation du rôle des structures secondaires issues des eucaryotes dans le recodage chez *E. coli* n'est pas observé pour les pseudo-nœuds de l'IBV et du BYDV. Chez les eucaryotes, le décalage de phase en -1 de l'IBV est dépendant d'un pseudo-nœud; chez *E. coli*, cette structure peut-être remplacée par une tige-boucle de même longueur (Brierley *et al.* 1997). Pour le décalage de phase en -1 du BYDV, une mutation dans la deuxième tige du pseudo-nœud ne semble pas affecter le taux de décalage de phase en -1 chez *E. coli* (Garcia *et al.* 1993).

Finalement, nous ne pouvons, avec les mutants testés ici, prédire quel type de décalage de phase, en tandem ou en singleton, est utilisé par *E. coli*. Des expériences indépendantes ont montré que des mutants du premier triplet homopolymérique réduisent d'un facteur 4 le décalage de phase du VIH-1 (UUUUUA plus tige-boucle) chez *E. coli*, et qu'une

mutation dans le deuxième triplet homopolymérique le réduisait d'un facteur supérieur à 10 (Weiss *et al.* 1989). Il s'agirait donc dans le cas du VIH majoritairement d'un décalage de phase en tandem des deux ARNt présents au site glissant, ce qui est le cas aussi chez les eucaryotes supérieurs *in vitro* (Jacks *et al.* 1988b), ou dans *E. coli* sans la séquence 3' stimulatrice (Yelverton *et al.* 1994). Dans le cas de l'IBV, il s'agit d'un décalage d'un seul ARNt et non des deux en tandem, puisqu'un mutant U CUA AAG provoque un décalage au même taux que la séquence U UUA AAG (Brierley *et al.* 1997). Il existe également des décalages de phase en -1 programmés chez *E. coli*, comme ceux de *dnaX*, codant pour une sous-unité de l'ADN polymérase, ou pour des séquences insertionnelles *IS* (pour revue Farabaugh 1996).

Le site de décalage de phase en -1 de *dnaX* ressemble fortement à un site eucaryote : heptamère glissant A AAA AAG, suivi d'une tige-boucle. Des mutants de l'heptamère ont montré que le décalage a majoritairement lieu par décalage en tandem (Blinkowa *et al.* 1990; Tsuchihashi *et al.* 1990; Tsuchihashi *et al.* 1992). Il était connu que cet heptamère monotone était particulièrement glissant chez *E. coli* (Weiss *et al.* 1989), mais ceci n'explique pas le fort taux de décalage observé (50%). Il a été démontré qu'il existe en plus une séquence type Shine-Dalgarno située en 5' qui stimule le recodage (cf. § III.A.2.c).

L'analyse de la séquence des *IS* révèle des sites consensus de décalage de phase en -1 par rapport aux eucaryotes : séquence glissante et structure secondaire en aval. Mais des études par mutagenèse et séquençage protéique ont démontré, dans le cas de *IS1*, que le décalage de phase avait en fait lieu grâce à un seul ARNt qui glisse sur un demi-site consensus A AAA (Sekine *et al.* 1992a; Sekine *et al.* 1992b). Toutes les structures secondaires prévisibles ont été mutagénisées, et aucune ne s'est révélée indispensable. Pourtant la séquence située en aval stimule le décalage de phase -1 d'un facteur 4. C'est un stop situé 18 nucléotides plus loin qui a ce rôle stimulateur (Sekine *et al.* 1992b). On peut imaginer qu'un premier ribosome stoppé en cours de terminaison bloquerait un deuxième ribosome sur le site de glissement. Les autres *IS* (*IS150*, *IS3*, *IS911*) n'ont pas été mutagénisées à ma connaissance pour savoir si le glissement se faisait en tandem ou grâce à un seul ARNt. Toutes ces *IS* pourraient faire appel à une deuxième particularité traductionnelle : le couplage traductionnel (ré-initiation immédiate après terminaison de la traduction) (Farabaugh 1996). Il semble donc que le décalage de phase chez *E. coli*, bien qu'en apparence similaire à celui des eucaryotes, soit

beaucoup plus diversifié que ne le laissent supposer les séquences impliquées.

Les outils construits s'avèrent efficaces, et continuent à être développés au laboratoire : des plasmides dérivés possédant d'autres sites de clonage ont été élaborés : le pAC98 dont le site NheI en 5' de la -galactosidase a été détruit, et dont le site BclI (sensible à la méthylation) a été remplacée par NotI (Stéphanie Boissard, Laure Bidou, communication personnelle); le pAC99 contient un unique site de clonage bouts francs MscI (Olivier Namy, communication personnelle); les vecteurs pACU contiennent le marqueur URA3 (Hongmei Liu, Olivier Namy, communication personnelle), les vecteurs pHI sont présents en multicopie chez la levure (Isabelle Hatin, communication personnelle), et les constructions obtenues peuvent être insérées par recombinaison homologue au locus ARG4 grâce à des homologies en 5' et en 3'.

La levure semble être un bon modèle pour l'étude des cibles de recodage issus des eucaryotes supérieurs. Nous avons voulu le confirmer en faisant une étude sur l'influence de la stabilité de la tige-boucle du VIH-1 sur le décalage de phase en -1, en parallèle entre les cellules de mammifères et la levure.

# Chapitre II. Étude de l'influence de la stabilité de la structure secondaire du VIH-1 sur le taux de décalage de phase en -1

## II.A. État des lieux en 1995

L'explication de l'influence des structures secondaires sur le recodage comme barrière énergétique arrêtant les ribosomes au niveau d'un heptamère glissant, et privilégiant ainsi cinétiquement un événement normalement défavorisé, est évoquée depuis que l'on a découvert le phénomène de décalage de phase en -1 programmé (Jacks *et al.* 1985). Cette explication est confortée par le fait que la plupart des gènes faisant appel au décalage de phase en tandem possèdent une structure secondaire en 3' de leur séquence glissante (Brierley 1995; Miller *et al.* 1995). Néanmoins les analyses de stabilité des structures secondaires par mutation des appariements des tiges impliquées ne sont pas toujours concluantes : parfois l'identité de nucléotides non appariés semblent importants, ou des mutations compensatoires ne restaurent pas un taux de décalage de phase sauvage (Jacks *et al.* 1988a; Brierley *et al.* 1989; Brierley *et al.* 1991; Morikawa *et al.* 1992; ten Dam *et al.* 1994; Chen *et al.* 1995; ten Dam *et al.* 1995, Kollmus *et al.* 1996) Les arguments pour expliquer ces comportements erratiques existent, comme la méconnaissance d'une partie des facteurs influant sur la formation et/ou la stabilité d'une structure secondaire (surtout en ce qui concerne les pseudo-nœuds), ou encore que ces résultats sont obtenus *in vitro*, en quantifiant sur gel les protéines obtenues et ne donnent qu'une indication qualitative de l'effet observé,... Mais le rôle des structures secondaires comme barrières énergétiques n'a jamais été prouvé de manière définitive, ce qui laisse la possibilité pour d'autres explications (cf. introduction, § B.1.g) : facteurs *trans* reconnaissant les structures secondaires, interaction avec des constituants du ribosome,... On ne peut d'ailleurs pas exclure un double rôle : freiner le ribosome dans un premier temps, puis avoir une action sur le ribosome. Ceci est suggéré par le fait que la plupart des pseudo-nœuds efficaces pour stimuler le décalage de phase en -1 présentent une structure particulière n'ayant pas d'incidence sur la stabilité, mais sur la forme (cf. introduction, § B.1.b), et par le fait que certaines structures secondaires provoquent une pause du ribosome semblable celle induite par une structure sauvage, mais sans provoquer de décalage de phase en -1 (Tu *et al.* 1992; Somogyi *et al.* 1993).

En ce qui concerne les tige-boucles stimulatrices du décalage de phase -1, la conformation tridimensionnelle n'a jamais été étudiée à ma connaissance, et les résultats obtenus par mutagenèse ne présentent pas pour l'instant de corrélation entre la stabilité et la capacité à provoquer un décalage de phase (Kollmus *et al.* 1994; Kollmus *et al.* 1996). Nous avons voulu profiter de notre outil pour mesurer précisément le taux de décalage de phase en -1 du VIH-1 en fonction de la stabilité théorique de la tige-boucle. Ce type d'analyse, avec un système rapporteur quasiment identique avait déjà été mené *in vivo* dans des cellules de rein de hamster BHK21 en ce qui concerne la tige-boucle stimulatrice du décalage de phase en -1 du HTLV-2 (Kollmus *et al.* 1994). Parmi les 6 mutants analysés, les stabilités de la tige-boucle calculées (de 0 à -13 kcal.mol<sup>-1</sup> pour le sauvage), et les variations mesurées de décalage de phase ne semblent pas corrélées. Les variants de la structure secondaire analysés avaient au minimum 3 nucléotides de différence avec la séquence sauvage ou entre eux. Vu les difficultés de prédiction de stabilité, il me semble difficile de tenter de corréler décalage de phase et stabilité calculée dans ces expériences. L'absence de témoin en phase d'activité spécifique et l'importance de la variabilité de ces expériences sont peut-être les explications du manque de corrélation obtenu.

Pour essayer de faire varier la stabilité de la tige de manière la plus contrôlée possible, nous avons construit des mutants de la tige sauvage dans lesquelles nous avons remplacé des appariements G-C par des appariements A-T et réciproquement, sans changer la longueur de la tige ni les appariements terminaux. Nous essayons de cette façon de ne faire varier que le minimum de paramètres, puisque des facteurs tels que la distance au site de glissement, la forme de la structure secondaire (tige-boucle, tige-boucle avec boucles internes, pseudo-nœud, pseudo-nœud coudé) peuvent faire varier le taux de décalage de phase indépendamment de la stabilité calculée de la structure. Avec cette façon de procéder, les changements d'une construction à l'autre peuvent permettre de hiérarchiser les mutants selon leur stabilité, indépendamment des calculs de stabilité absolus. Nous avons testé tous ces variants, ainsi que pour chacun un témoin en phase, en parallèle dans la levure *S. cerevisiae* et les cellules NIH3T3. Si les résultats entre levure et cellules de mammifères sont comparables, cela nous permet de renforcer la conclusion de l'article précédant quant à la conservation des règles de recodage entre levure et mammifères (Stahl *et al.* 1995). S'il existe des différences entre les deux organismes, ce pourrait être très informatif quant au rôle de la tige-boucle.

## **II.B. ARTICLE II**

<http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m&form=6&uid=97465799&Dopt=b>

## **II.C. Existence de la tige-boucle en aval du site de glissement de VIH-1**

À la publication de cet article, la structure putative en tige-boucle de la séquence palindromique située en aval de l'heptamère glissant du VIH-1 n'avait pas été prouvée de manière directe. Le test fonctionnel du rôle de cette séquence en fonction de son contenu en A-T versus G-C est sans doute la meilleure manière de prouver indirectement l'existence de cette structure. Nos résultats (Fig. 2 de l'article II) sont un argument supplémentaire pour l'existence de cette structure. Des preuves plus directes viennent d'être obtenues grâce à des sondes enzymatiques (Kang 1998). Finalement, une approche à base de RMN, comme celles utilisées pour élucider la forme des pseudo-nœuds, serait encore plus informative (Shen *et al.* 1995; Chen *et al.* 1996; Kang *et al.* 1997).

En ce qui concerne nos analyses, l'augmentation du contenu en GC de cette séquence fait augmenter en parallèle le taux de décalage de phase en -1 du VIH-1, et ce jusqu'à un maximum obtenu pour la séquence sauvage (pour le mutant plus stable mais provoquant moins de décalage de phase, voir le paragraphe ci-dessous). Depuis la publication de cet article, une autre analyse du même type a été menée pour la tige-boucle en aval du site de glissement de *dnaX in vivo* chez *E. coli* (Larsen *et al.* 1997). Une délétion totale de la moitié supérieure de la tige-boucle n'a qu'un effet mineur sur le taux de décalage de phase (87% du sauvage). Cette structure est pourtant en théorie deux fois moins stable que la structure sauvage (-6,8 kcal/mol comparé à -14,5). Cette partie de la tige est plus sensible aux sondes chimiques que le bas de la tige. La délétion du reste de la tige-boucle ne permet plus que 27% du décalage de phase du sauvage, soit 14% en valeur absolue. Ce fort taux de décalage résiduel est explicable par les qualités glissantes de l'heptamère A AAA AAG chez *E. coli* et par la présence d'une séquence SD stimulatrice en amont (Larsen *et al.* 1994). Pour tous les autres mutants testés, il y a une très bonne corrélation entre la stabilité calculée et le taux de décalage de phase, ce qui corrobore nos résultats.

## II.D. Mutant plus stable que le sauvage

Le cas du mutant 1830 a constitué une surprise. Ce mutant est censé avoir une tige-boucle très stable, constitué presque exclusivement d'appariements GC. Ce mutant provoque pourtant un décalage de phase à un taux moins fort que le sauvage (75-80%). Plusieurs hypothèses pourraient expliquer ce phénomène.

(i) Il pourrait s'agir d'une différence de stabilité de l'ARNm ou d'activité spécifique de la protéine de fusion. Notre système de mesure permet d'infirmer cette hypothèse, étant donné que tous les mutants testés sont comparés à un témoin en phase qui ne diffère que par un A inséré après le site heptamérique.

(ii) Il pourrait s'agir d'un blocage de la traduction, de "drop-off", dû à la présence d'une structure trop forte, qui stoppe complètement le ribosome, et provoque le relargage du peptidyl-ARNt. Il y aurait donc synthèse uniquement de -galactosidase, si celle-ci n'est pas alors dégradée. Mais cet effet de blocage de la traduction devrait être le même sur le témoin en phase, grâce auquel nous calculons le taux de décalage de phase, et cet effet serait pris en compte dans nos calculs.

(iii) On peut imaginer que si l'on allonge trop le temps de pause, les ARNt reviennent en phase 0 après un aller-retour : il y aurait une "erreur dans le recodage". Le temps de pause du sauvage serait le temps de pause optimal pour permettre aux ARNt de glisser en phase -1, sans pour autant revenir en phase 0. On peut toutefois augmenter le taux de décalage de phase en -1 du VIH-1 par d'autres structures secondaires (tige-boucle du HTLV-2, double tige-boucle, cf. ci-après). Si ces structures jouent seulement le rôle de barrière énergétique, ceci pourrait infirmer cette hypothèse.

(iv) Une autre hypothèse est envisageable, à la lumière des expériences qui ont mis en évidence que derrière un ribosome bloqué (à l'initiation, au cours de la terminaison, derrière une structure secondaire,...) il y avait accumulation de ribosomes accolés les uns aux autres (Wolin *et al.* 1988). On peut imaginer qu'une structure secondaire plus stable ait tendance à accumuler plus de ribosomes que la structure sauvage. Si le premier ribosome déroule la tige-boucle avant de poursuivre l'élongation, il se peut que la structure secondaire n'ait pas le temps de se reformer, et qu'un deuxième ribosome ne soit pas bloqué à son tour, et donc ne subisse pas de décalage de phase.

Plusieurs expériences sont envisageables pour tester cette dernière hypothèse. On peut envisager de vérifier que les pauses provoquées par les tige-boucles sont en corrélation avec leur stabilité, et cela même pour le mutant le plus fort. Ceci prouverait que les ribosomes subissent une pause plus longue derrière cette structure 1830. On devrait également pouvoir visualiser des ribosomes supplémentaires bloqués pour le mutant 1830. Une autre expérience plus indirecte serait de ralentir artificiellement le taux de traduction de cet ARNm (en modifiant le contexte de son codon initiateur), ou en ralentissant la vitesse d'élongation, soit en disposant des codons rares en amont du site de décalage, soit avec des inhibiteurs de l'élongation traductionnelle. La corrélation observée entre stabilité de la tige boucle et taux de décalage de phase devrait alors s'étendre à la tige la plus stable, qui n'accumulerait plus les ribosomes.

Dans d'autres laboratoires, des expériences tendant à modifier la stabilité de structures favorisant le décalage de phase en -1 ont été menées, avec parfois l'obtention de mutants plus stables provoquant des décalages de phase plus importants que le sauvage (Brierley *et al.* 1991; ten Dam *et al.* 1995). Mais ces études portaient sur des pseudo-nœuds, *in vitro*. Des expériences ont montré que pour des mutants hybrides, contenant le site de glissement du VIH-1 et la tige-boucle stimulatrice du HTLV-2 le taux de décalage de phase augmente d'un facteur 6 (de 3% avec le VIH-1 sauvage à 17% pour le site hybride) dans des cellules de hamster en culture (Kollmus *et al.* 1994). Ce résultat est très étonnant, vu la similarité des structures : une tige d'une dizaine de nucléotides de hauteur, majoritairement à composition GC. Les différences notables entre les deux structures sont : une tétraboucle pour le VIH-1, une pentaboucle pour le HTLV-2; la distance entre le site de glissement et la tige-boucle, qui est de 7 nucléotides pour le HTLV-2 et qui pourrait être de 8 nucléotides pour le VIH; la composition de la partie 5' de la tige qui est mixte pour le VIH-1, mais composée seulement de purines pour le HTLV-2. Ce résultat demande à être vérifié vu le facteur énorme d'augmentation du taux de décalage de phase mesuré. Le site de décalage de phase du HTLV-2 nécessite apparemment une structure stimulatrice particulière, puisqu'elle ne peut pas être remplacée ni par celle du VIH-1, ni par une autre tige-boucle IRE (Iron Response Element) (Kollmus *et al.* 1994; Kollmus *et al.* 1996). L'addition d'une deuxième tige-boucle coaxiale avec celle du VIH-1 (structure secondaire de 50,6 kcal.mol<sup>-1</sup>) provoque un décalage de phase en -1 de plus de 30% (Vickers *et al.*

1992). Les mécanismes d'action d'une telle structure secondaire ne sont toutefois pas connus.

Les expériences les plus proches de celles que nous avons réalisées ont été rapportées pour la tige-boucle de *dnaX*, *in vivo* chez *E. coli*. Dans ce cas, des mutants plus stables permettent d'augmenter le taux de décalage de phase d'un facteur significatif par rapport au sauvage : 120% en créant des appariements sur un bulge A G-> C-G; 136% en supprimant un petit bulge C; 176% pour la combinaison des deux mutants précédents; 126% pour une augmentation de trois bases de la tige-boucle (Larsen *et al.* 1997). Dans ce système procaryote, on observe donc une corrélation entre décalage de phase et stabilité jusqu'à presque deux fois celle du sauvage. Les règles ne sont peut-être pas les mêmes en ce qui concerne les deux règnes, ou alors la limite n'est pas la même ( G de -23,4 kcal.mol<sup>-1</sup> pour le mutant le plus fort) dans ce système que dans la notre (-22,8 pour notre mutant 1830). Des facteurs tels qu'efficacité d'initiation ou vitesse d'élongation dans l'hypothèse de l'accumulation de ribosomes bloqués, ou fidélité intrinsèque du système traductionnel dans l'hypothèse de type "aller-retour", peuvent faire varier l'optimum de stabilité. Je me propose de mesurer le taux de décalage de phase de nos mutants du VIH-1 chez *E. coli* pour tester cette hypothèse, en contexte sauvage et *rpsL* (mutant plus fidèle, dont la vitesse de traduction est plus faible, cf. § III.A.2.a).

En tout état de cause il faudrait, pour être certain de la réalité de ce résultat, étudier un mutant encore plus stable que le mutant 1830, ou alors comparer dans des systèmes homogènes (mêmes systèmes rapporteurs, même organisme) les résultats des deux expériences les plus informatives (Bidou *et al.* 1997; Larsen *et al.* 1997) pour pouvoir conclure quant aux causes de ce phénomène.

## **II.E. Comparaison entre levure et cellules de mammifère, facteurstrans**

Nous observons exactement le même type de résultats dans les cellules de mammifères en culture et les cellules de levure. Nous confirmons donc les résultats obtenus dans le premier article, où nous observions une conservation des règles de recodage entre levure et mammifères. Ceci

nous indique également qu'il n'y a pas de facteur *trans* spécifique d'un organisme. On peut arguer contre cela que les fibroblastes de souris NIH3T3 ne sont pas l'hôte du VIH dans la nature, mais les résultats sont identiques pour au moins trois constructions dans des cellules de lymphome T (Cassan *et al.* 1994). Il n'y a sans doute pas non plus d'intervention d'un facteur viral, puisque ces résultats ont été testés avant et après infection par le VIH-1 (Cassan *et al.* 1994; Reil *et al.* 1994).

Le fait que le changement de la boucle (partie la plus exposée à une interaction séquence spécifique) dans le mutant 1804 n'ait qu'un effet marginal sur le taux de décalage de phase, et que nous modifions de façon non négligeable la séquence primaire de la tige-boucle dans les différents mutants va dans le sens de l'absence d'un facteur *trans* séquence spécifique.

Dans le prochain chapitre je vais décrire des approches pour identifier des facteurs *trans* modulant le recodage.

## Chapitre III. Facteurs Trans modulant le recodage

Ce chapitre est organisé en deux parties : je ferai d'abord une revue des facteurs qui ont un effet spécifique sur les différents types de recodage, puis j'exposerai l'approche génétique que j'ai élaborée pour identifier de nouveaux facteurs *trans*. Les informations sur le recodage et sur la fidélité générale de la traduction que peuvent apporter des études croisées entre facteurs connus de la traduction, et éléments *cis* bien caractérisés du recodage seront discutées dans la conclusion de cette thèse.

### III.A. Facteurs *trans* identifiés du recodage

Je n'exposerai dans cette partie que les facteurs identifiés qui modulent spécifiquement le recodage, c'est-à-dire les facteurs criblés spécifiquement par rapport à des cibles du recodage, ou les facteurs généraux de la fidélité de la traduction qui ont été testés sur des cibles du recodage. Il est probable qu'une majorité des facteurs contrôlant la fidélité générale de la traduction (revue dans Parker 1989; Atkins *et al.* 1991; Hinnebusch *et al.* 1991), auraient un effet sur certaines cibles du recodage, mais ces expériences restent pour la plupart à faire.

À ma connaissance, seuls trois cribles spécifiques ont été menés pour découvrir des facteurs modulant le recodage :

- un crible mené pour identifier des facteurs surexprimés inhibant la transposition du Ty chez *S. cerevisiae* a permis d'isoler un ARNt réduisant son taux de décalage, et donc son taux de transposition (Xu *et al.* 1990),
- un crible a été mené chez *E. coli* pour identifier des facteurs modulant le hopping du gène 60 du bactériophage T4 et a permis d'identifier la protéine ribosomale L9 (Herbst *et al.* 1994),
- un crible utilisé pour identifier des mutants de *S. cerevisiae* augmentant le décalage de phase en -1 du ScV (mutants *mof* pour "maintenance of frame"), a permis d'identifier 9 mutants appartenant à des groupes de complémentation différents (Dinman *et al.* 1994). Parmi les gènes identifiés, *mof2-1* est muté dans le gène *SUI1*, facteur d'initiation de la traduction (cf. § III.A.2.b ou Cui *et al.* 1998), *mof4* est un mutant affecté

dans la stabilité des ARNm contenant un stop précoce (Non-sense mediated mRNA decay, NMD) (cf. § III.A.2.b ou Cui *et al.* 1996), *mof9* est un mutant de l'ARNr 5S (cf. § III.A.2.c ou Dinman 1995). L'effet de ces mutants sur le recodage a été vérifié par des tests fonctionnels biologiques indirects : persistance du virus satellite M1 dépendant de la polymérase du ScV pour sa réplication, et donc du décalage de phase en -1; taux de transposition d'un transposon Ty modifié contenant un marqueur sélectif pour le décalage de phase en +1 (Dinman *et al.* 1998). Ces cribles permettent de différencier des facteurs affectant l'expression du gène rapporteur autrement que par un effet sur le recodage (transcription, stabilité des ARNm,...) sans avoir la facilité et la fiabilité du crible secondaire exposé ci-dessus (cf. § III.A.6).

### **III.A.1. Sélénocystéine**

Je ne détaillerai pas l'exemple le plus flagrant d'utilisation de facteurs *trans* pour stimuler le recodage qu'est l'incorporation de sélénocystéine. Ce système de recodage repose essentiellement sur des facteurs *trans* particuliers : facteur d'élongation, enzymes de modification des ARNt (cf. introduction, § B.3.f ou Berry *et al.* 1993)

## III.A.2. Gènes cellulaires

### III.A.2.a. protéines ribosomales

Des mutants des protéines ribosomiques de *E. coli* S4 et S5, dits *ram* (pour “ribosomal ambiguity”) sont affectés dans la fidélité de la traduction (pour revue Parker 1989). Il a d’ailleurs été mesuré que la constante cinétique  $k_3$  est deux fois plus élevée dans ces mutants, (cf. Fig.1, introduction § A) . Autrement dit, dans ces mutants le transfert peptidique se fait plus rapidement, et les ARNt non-cognats ont plus de chance d’être incorporés au lieu de se dissocier (Thompson 1988). Des mutants *rpsD* (S4) ou *rpsE* (S5) ne modifient pas le taux de décalage du bactériophage T7 (Siple *et al.* 1991). De même des mutants de *SUP44* (S4, équivalent de S5 chez *E. coli*) et *SUP46* (S13, équivalent de S4 chez *E. coli*) sont sans effet sur le taux de décalage de phase en -1 du ScV chez *S. cerevisiae* (Dinman *et al.* 1994).

Des mutants *rpsL/strA* (S12), au contraire, présentent une fidélité accrue de la traduction et (en général) une diminution de la vitesse de traduction ( $k_3$  diminué de moitié, Thompson 1988). Cette modification de la fidélité et/ou de la vitesse de traduction ne modifient pas le taux de décalage de phase en -1 de *dnaX* mesuré *in vitro* (Tsuchihashi 1991) ou du bactériophage T7 (Siple *et al.* 1991). En revanche, des mutants *rpsL*, ont un effet sur le taux de décalage de phase en +1 d’une cible modifiée issue de RF2 où le codon stop affamé UGA est remplacé par UGG : 13% de décalage de phase chez le sauvage, cinq fois plus pour le mutant *rpsL* le plus fidèle (streptomycine dépendant), avec des taux intermédiaires pour les mutants streptomycine pseudo-dépendants ou résistants (Curran *et al.* 1989). Ces mutants n’ont aucun effet sur une cible contenant un codon non affamé CUG (décalage de phase en +1 < 2%) (Siple *et al.* 1993), ou sur le codon UGG nourri par un excès d’ARNt<sub>Trp</sub> cognat (7% de décalage de phase chez le sauvage et le mutant *rpsL* le plus fidèle) (Siple *et al.* 1993). De même, une souche *rpsL* streptomycine résistante n’a pas d’effet sur le taux de décalage de phase sur des cibles “RF2”, dont la séquence de glissement est UUU Ynn (décalage de phase en +1 < 5%) (Schwartz *et al.* 1997). Le décalage de phase en +1 de RF2 est donc dépendant de la fidélité du ribosome au niveau du choix de l’ARNt au site A. Les mutants *rpsL* influent sur le stimulateur du décalage de phase en +1, la rareté de l’ARNt au site A; si l’on modifie la fidélité de sélection de l’ARNt, on modifie également la présence d’un ARNt (cognat ou presque

cognat) au codon affamé. Il pourrait être intéressant d'observer l'effet inverse avec les mutants *ram*.

La traduction du gène *60* du phage T4 chez *E. coli* provoque un "saut de ribosome" (en fait un glissement de l'ARNm) de 50 nucléotides avec des contraintes en *cis* très précises, à un taux de 100%. Un mutant dont la tige-boucle est allongée de 36 nucléotides a une efficacité de saut qui chute à 0,035%. Ce mutant a été utilisé pour isoler un suppresseur extragénique, *hop-1*, qui s'est avéré être un mutant substitué de la protéine ribosomale L9 (Herbst *et al.* 1994).

### **III.A.2.b. Facteurs de la traduction**

#### **Initiation**

Le mutant *mof2-1* est touché dans le gène *SUI1*, identifié au préalable comme acteur du choix du codon initiateur. Les mesures de décalage de phase montrent que *mof2-1* et *sui1-1* augmentent le taux de décalage de phase programmé en -1 du ScV, mais pas le décalage de phase en +1 du Ty1 ni des décalages de phases anonymes en +1 ou -1 (Cui *et al.* 1998). Le fait que d'autres mutants de la phase d'initiation *sui2* et *SUI3* ne soient pas affectés dans le décalage de phase en -1, que *mof2-1* soit affecté tout comme *sui1-1* dans la reconnaissance d'un codon initiateur correct (AUG vs UUG), suggère un rôle général pour *SUI1* dans la fidélité de la traduction au cours de l'initiation et de l'élongation (Cui *et al.* 1998). Ce type d'effet pourrait être médié par une activité stimulatrice sur les activités GTPasiques de eIF-2 et EF-1A, ce qui affecterait l'efficacité du "proofreading" médié par ces deux facteurs.

#### **Élongation**

Le facteur d'élongation EF-1A (EF-Tu chez les procaryotes) est codé chez *S. cerevisæ* par deux gènes *TEF1* et *TEF2* (Nagata *et al.* 1984; Schirmaier *et al.* 1984; Cottrelle *et al.* 1985). Des mutants ne contenant qu'un gène fonctionnel sont moins suppresseurs que le sauvage au niveau d'allèles non-sens (Song *et al.* 1989), et augmentent le décalage de phase en -1 du ScV (Dinman *et al.* 1997a). Il était connu chez *E. coli* que des mutants *tufA* ou *tufB* modifient la fidélité de la traduction. Tout comme les mutants *ram*, des mutants *tufAr* ont un k3 multiplié par deux, (cf. Fig.1,

introduction § A et Thompson 1988). Chez *S. cerevisiae*, des mutants de *TEF2* ont été produits *in vitro*, puis criblés *in vivo* pour un phénotype [SUF] dominant (suppression de décalage de cadre anonymes). Certains de ces mutants sont également capables de supprimer des allèles non-sens (Sandbaken *et al.* 1988). Sur 9 allèles analysés pour leur effet spécifique sur différentes cibles du recodage, 2 sont affectés dans le décalage de phase de lecture en -1 du ScV, alors que deux autres le sont au niveau du décalage de phase en +1 du Ty1 (Dinman *et al.* 1997a). Certains de ces allèles ont été également testés dans le contexte Ty3 : GCG AGU U. Le codon GCG a été remplacé par chacun des 63 autres codons, ce qui produit divers niveaux de décalage de phase (cf. introduction § B.2.b). Dans les deux mutants testés (dont l'un déjà identifié pour son effet sur Ty1 ci-dessus), tous les niveaux de décalage de phase sont augmentés d'un facteur 3 à 5, ce qui implique qu'il n'y a pas d'effet spécifique de ces mutants sur le décalage de phase favorisé par un ARNt glissant ou occluant (Farabaugh *et al.* 1998).

### Terminaison

Comme on l'a vu en introduction, le taux de décalage de phase en +1 de RF2 est modulé par la capacité de RF2 à reconnaître un codon UGA en phase. La concentration cellulaire de RF2 est ainsi régulée par un boucle de retro-contrôle négative : RF2 est le principal facteur *trans* de son propre recodage. Mais tout facteur faisant varier l'efficacité de terminaison de RF2 fera varier le taux de décalage de phase nécessaire à son expression et sera donc un facteur *trans* indirect de ce recodage (voir par exemple Arkov *et al.* 1995; Grentzmann *et al.* 1995; Mottagui Tabar *et al.* 1997).

Chez la levure, un mutant omnipotent *SUP45* (eRF1), ou le facteur [ ] (forme suppresseur prion de eRF3) n'ont pas d'effet sur le décalage de phase en -1 du ScV (Dinman *et al.* 1994). L'influence de la terminaison sur le recodage pourrait donc être limitée aux translectures et au rétro-contrôle de RF2. Des mutants de eRF3 ont pourtant été isolés comme *sup*, suppresseurs de décalage de cadre de lecture anonymes (Culbertson *et al.* 1982; Wilson *et al.* 1988a).

## Les facteurs NMD

Il existe chez les eucaryotes, un système spécifique de dégradation des ARNm contenant un stop précoce, c'est à dire situé en 5' dans la phase codante (Peltz *et al.* 1994; Maquat 1995). Ce système ("non-sense mediated decay" - NMD) est utilisé pour dégrader les ARNm mal transcrits, mal épissés, ou transcrits à partir de gènes contenant une mutation non-sens. Des ARNm contenant un décalage de cadre présentent toujours, en pratique, un codon stop dans la phase 0. Dans le cas des premiers cribles utilisés chez la levure pour tenter d'isoler des allosuppresseurs de décalage de cadre (*upf* pour "up frameshift suppressor"; Culbertson *et al.* 1980), les gènes utilisés comme crible contiennent un décalage de cadre de lecture, et donc un codon stop précoce. Les mutants *upf* sont touchés dans la voie NMD, et augmentent la stabilité et non le taux de décalage de phase de ces ARNm (Leeds *et al.* 1991; Leeds *et al.* 1992; Lee *et al.* 1995a). D'autres cribles visant à identifier chez *S. cerevisiae* les mutants *mof* au niveau du ScV (Dinman *et al.* 1994) et *ifs* (pour "increased frameshift") du MMTV (Lee *et al.* 1995b) ont mené à l'identification de mutants dans cette voie NMD (Lee *et al.* 1995b; Cui *et al.* 1996). Des expériences quantitatives d'expression et de stabilité des ARNm, suggèrent que certains allèles mutants des gènes *UPF/NMD* seraient également altérés dans la terminaison *per se* (Weng *et al.* 1996a; Weng *et al.* 1996b). De plus le mutant *mof4-1*, touché dans le gène *UPF1*, n'est pas affecté seulement dans la voie NMD : des ARNm contenant un stop précoce sont deux à trois fois moins dégradés dans des mutants *mof4-1*, *upf1-Δ*, *ifs2-1* que dans un sauvage *UPF1+*. Cette stabilisation d'un facteur 2 s'accompagne dans tous les cas d'une augmentation d'expression attendue d'un facteur 2, à part pour *mof4-1*, dont le taux d'expression est mutliplié par 4,5. Cette augmentation du taux d'expression de *lacZ* pourrait tenir à deux facteurs : la stabilisation d'un facteur 2 de l'ARNm *lacZ* hors de phase, et l'augmentation du taux de recodage au niveau du ScV (Cui *et al.* 1996). Ces résultats sont très prometteurs, puisqu'en plus de ces phénotypes pléiotropes, *mof4-1* est le premier mutant dans lequel le taux de décalage de phase est variable en fonction de la concentration en paromomycine *in vivo* (cf. ci dessous, § III.A.6 ou Cui *et al.* 1996).

### **III.A.2.c. ARNr**

De nombreux mutants des ARNr affectant la fidélité de la traduction (fidélité, maintien de cadre de lecture, terminaison) ont été isolés, spécialement chez *E. coli* (von Ahsen 1998), mais aussi chez *S. cerevisiae* (Liebman *et al.* 1995). Leur effet sur des cibles de recodage n'a pas été évalué, et je ne présenterai ci-dessous que les effets avérés sur des cibles du recodage ou assimilées.

#### **23S**

Un mutant résistant au chloramphénicol de l'ARNr 21S (équivalent de l'ARNr 23S dans la mitochondrie de *S. cerevisiae*), a des taux de décalage de phase en +1 et en -1 diminués au niveau d'un homopolymère U<sub>6</sub> (Weiss-brummer *et al.* 1995). Bien que la séquence du site de glissement soit semblable à celui utilisé par les rétrovirus, le mécanisme de décalage de phase n'est pas connu à ce jour.

## 16S

La séquence Shine-Dalgarno (SD) intervient en amont de plusieurs recodages : le décalage de phase en -1 de *dnaX* (à 10 nucléotides en amont) et le décalage de phase en +1 de RF-2 (3 nucléotides en amont, cf. introduction § B.2.c). Ces séquences agissent par interaction avec l'ARNr 16S : des mutants de ces séquences SD diminuent le taux de décalage de phase, et parallèlement des mutants compensatoires dans la séquence anti-SD de l'ARNr 16S restaurent un taux de recodage sauvage (Weiss *et al.* 1988; Larsen *et al.* 1994). La distance de la séquence SD au site de recodage est également cruciale pour son effet stimulateur. Le site de décalage de RF2 n'est efficace que si la séquence SD est localisée précisément à 3 nucléotides en amont (Weiss *et al.* 1987a; Larsen *et al.* 1994), une séquence SD trop proche du site de décalage de *dnaX* (2-4 nucléotides en amont) devient inhibitrice du décalage de phase en -1 (Larsen *et al.* 1994). Il a été suggéré que l'ARNr 16S positionnait précisément le site de décodage par rapport à la séquence SD (comme au moment de l'initiation), et pourrait décaler l'ARNm par rapport au ribosome. Un rôle plus simple de ralentissement du ribosome ne peut être exclu, et n'a pas été étudié à ce jour.

## 5S

Chez *S. cerevisiae*, l'ARNr est codé au locus *RDN1*, qui est une répétition directe de plus d'une centaine de copies des gènes codant l'ARNr 35S, précurseur des ARNr 23S, 18S et 5,8S, et l'ARNr 5S. Le phénotype du mutant *mof9*, qui augmente le décalage de phase en -1 du ScV et le décalage de phase en +1 du Ty1, peut être supprimé par une copie plasmidique de l'ARNr 5S. D'autres mutants de l'ARNr 5S épisomiques, ainsi que des insertions dans le locus *RDN1* (*RDN1::LEU2* ou *RDN1::URA3*) ont des phénotypes similaires à *mof9* (Dinman *et al.* 1995).

Un crible utilisé au laboratoire a permis d'isoler un clone d'ADN génomique en multicopie modifiant le taux de translecture du TMV. Ce clone contient une partie du locus *RDN1*, dont le seul gène entier est le gène codant pour 5S (Olivier Namy, communication personnelle). Aucun rôle précis n'a pu être attribué à l'ARNr 5S, et des études sont en cours pour déterminer l'implication de l'ARNr 5S dans la terminaison ou le maintien du cadre de lecture.

### III.A.2.d. ARNt

Les ARNt sont au premier rang des facteurs impliqués dans le recodage. Ils sont nécessaires à la suppression lors de la translecture et interviennent dans le maintien du cadre de lecture (Atkins *et al.* 1991; Hinnebusch *et al.* 1991).

#### **Décalage de phase +1**

Le site de décalage de phase en +1 dont la dépendance vis-à-vis des ARNt a été le plus étudié est l'heptanucléotide du Ty1 CUU AGG C. Tous les ARNt susceptibles de décoder tout ou partie de ce site, en phase 0 ou en phase +1 sont cruciaux :

(i) une compétition entre l'ARNt<sub>Leu</sub> (anticodon 3'GAU<sub>5</sub>' venant décoder CUU, et susceptible de se ré-apparier en phase +1 avec le codon UUA) et un ARNt<sub>Leu</sub> modifié dont l'anticodon 3'GAA<sub>5</sub>' est capable de décoder CUU mais incapable de reconnaître UUA abolit le décalage de phase en +1 (Belcourt *et al.* 1990),

(ii) l'ARNt rare, qui rend le codon AGG affamé, permet à l'ARNt<sub>Leu</sub> de se recalculer en phase +1; sa surexpression inhibe le décalage de phase en +1 (Belcourt *et al.* 1990; Xu *et al.* 1990), et son absence augmente le taux de décalage de phase à plus de 90% (Kawakami *et al.* 1993).

(iii) l'ARNt nécessaire à la suite du décodage influence également le taux de décalage de phase en +1. C'est en général un ARNt abondant, mais dans le cas étudié (artificiel), on peut surexprimer un ARNt rare correspondant, et augmenter le décalage de phase (Pande *et al.* 1995). Ceci suggère que le premier ARNt de la phase +1 fixe ou favorise (par une arrivée au site A hors de phase, par exemple) cette erreur dans le maintien de la phase.

Des expériences similaires ont été menées sur le décalage de phase en +1 de RF2 chez *E. coli* (Sipley *et al.* 1993), et sur le site de décalage en +1 du Ty3 (cf. introduction § B.2.b ou Farabaugh *et al.* 1993; Vimaladithan *et al.* 1994; Pande *et al.* 1995).

Un site de décalage de phase en +1 UUU CGG (modifié de RF2) dépend du glissement de l'ARNt<sub>Phe</sub>. Cet ARNt a les modifications 2-méthylthio et

6-isopentenyl. La mutation du gène *miaA*, responsable de ces modifications chez *Salmonella typhimurium* augmente le taux de décalage de phase d'un facteur 1,5 à 2,5 en présence ou en absence de séquence SD stimulatrice (Schwartz *et al.* 1997). Ces modifications sont sans doute nécessaires à une stabilisation du duplex ARNt-ARNm, et leur absence permettrait à l'ARNt<sup>Phe</sup> de se désapparier plus facilement, et donc de se décaler en +1.

### Décalage de phase -1

Parmi les séquences connues de décalage de phase en -1, les deuxièmes codons de l'heptamère glissant X XXY YYZ sont limités à 5 chez les eucaryotes (AAC, AAU, UUA, UUC, UUU) et 2 chez les procaryotes (AAA, AAG) (Farabaugh 1996). Tous les ARNt capables de décoder ces codons possèdent des hypermodifications dans la boucle anticodon (Farabaugh 1996; Brierley *et al.* 1997). Les modifications sont résumées ci-dessous en fonction des codons concernés :

- ARNt<sup>Asn</sup> (AAC, AAU) : la base juste en 3' de l'anticodon est une queuine (Q),
- ARNt<sup>Leu</sup> (UUA) : la troisième base est une 2-méthyl-5-formyl-cytidine,
- ARNt<sup>Phe</sup> (UUC, UUU) la base juste en 3' de l'anticodon est une wyebutoxine (Y),
- ARNt<sup>Lys</sup> (AAA, AAG) : la troisième base est une 5-methoxy-carbonylméthyl-2-thiouridine chez les eucaryotes et une 5-méthylaminométhyl-2-thiouridine chez les procaryotes; chez les eucaryotes, il existe un autre iso-accepteur avec un anti-codon 3'UUC<sub>5'</sub> non modifié, et il a été suggéré que cette différence expliquait la forte tendance à glisser d'un site A AAG chez *E. coli* (*dnaX*, par exemple) et pas chez les eucaryotes (Tsuchihashi 1991; Tsuchihashi *et al.* 1992).

Les ARNt hyper modifiés pourraient changer la structure de la boucle anticodon et ainsi stabiliser l'interaction codon-anticodon; *a contrario*, l'absence de modifications pourrait promouvoir le décalage de phase (Hatfield *et al.* 1992). Parallèlement à ces observations structurales, les ARNt impliqués dans le décodage de certains sites de décalage de phase en -1 (VIH-1, BLV, MMTV) ont été étudiés par chromatographie. Le niveau d'expression de ces ARNt ne semble pas modifié, mais certains d'entre

eux sont majoritairement non modifiés au niveau d'une base de la boucle anticodon, comme la wyebutoxine pour l'ARNt<sup>Phe</sup> lors de l'infection par le VIH, ou la queuosine pour l'ARNt<sup>Asn</sup> lors de l'infection par le HTLV-1 ou le BLV (Hatfield *et al.* 1989). Il a été suggéré que l'infection par le virus, provoquait une inhibition de la modification des ARNt, avec une influence sur le taux de décalage de phase (Hatfield *et al.* 1990). Le taux de décalage de phase en -1 du VIH-1 ne varie pas dans des cellules infectées par le VIH-1 par rapport à des cellules non-infectées dans deux expériences indépendantes menées dans deux systèmes différents (Cassan *et al.* 1994; Reil *et al.* 1994). Ces études, et d'autres menées sur l'effet de l'infection du MoMuLV sur sa translecture (Panganiban 1988; Berteaux *et al.* 1991), suggèrent qu'aucun facteur viral n'influe sur le taux de recodage. L'observation faite sur la diminution de l'état modifié des ARNt dans des cellules infectées pourrait être plutôt mise en rapport avec le choix de ces ARNt comme amorce préférentielle pour la reverse transcriptase (Agris *et al.* 1997), ou même être une conséquence indirecte de l'infection virale.

D'autres expériences faites sur l'étude des modifications des ARNt ont mené à la conclusion que celles-ci ne sont pas directement impliquées dans le décalage de phase en -1 : le mutant *mod5*, ne possédant plus de modification isopentenylque ne présente pas de modifications du taux de décalage de phase du ScV (Dinman *et al.* 1994); l'étude du décalage de phase en -1 de l'IBV chez *E. coli* a montré que ce décalage de phase est peu sensible à l'absence de modifications des ARNt (Brierley *et al.* 1997); et une déficience dans la modification ms<sup>2</sup>io<sup>6</sup>A37 de l'ARNt<sup>Phe</sup> chez *Salmonella typhimurium* ne modifie pas le décalage de phase en -1 sur un site homopolymérique U<sub>6</sub>A (type VIH) (Q. Qian & G.R. Björk, manuscrit en préparation).

### **III.A.3. Facteurs de l'ARNm agissant entrans**

On peut imaginer qu'il y ait des séquences distantes sur les ARN viraux ayant un effet en *trans* sur le ribosome. On a par exemple évoqué le pseudo-nœud en 3' de beaucoup de virus ARN de plantes, qui peut prendre une forme d'ARNt, et qui est même aminoacétylable (Gesteland *et al.* 1996). De même, il a été suggéré que le pseudo-nœud du MMTV, présentait une conformation spatiale très semblable à celle de l'hélice formée par le codon et l'anti-codon (Shen *et al.* 1995). On peut imaginer que ce type de structure perturbe des activités biochimiques du ribosome

ou de facteurs de la traduction sensibles à l'existence d'un appariement codon-anticodon. Des structures secondaires très éloignées, comme l'élément SECIS responsable de l'insertion de sélénocystéine à de multiples codon stop, constituent sans doute le meilleur exemple d'éléments agissant en *trans*.

#### **III.A.4. Peptide en croissance**

Le peptide en croissance lors de la traduction du gène *60* du bactériophage T4 influence le hopping de 50 nucléotides en aval (Weiss *et al.* 1990b). Cet effet est certainement intra-ribosomal, puisque le tunnel de sortie du peptide du ribosome est susceptible de contenir une chaîne peptidique d'une trentaine d'acides aminés (Gesteland *et al.* 1996). Le mode d'action de ce peptide n'est pas connu, mais on peut penser à un frein pour l'élongation du ribosome, par analogie avec la pause provoquée par un peptide signal (Wolin *et al.* 1988). Ces effets sont à relier avec l'effet du dernier et du pénultième acide-aminé traduit sur l'efficacité de terminaison (Mottagui Tabar *et al.* 1997).

### **III.A.5. Polyamines**

Les polyamines spermine et spermidine, ont été impliquées dans de nombreuses fonctions biologiques, dont la fidélité de la traduction (Tabor *et al.* 1984). Chez la levure *S. cerevisiae*, la délétion du gène *SPE2* (codant la S-adenosylmethionine décarboxylase) rend la cellule dépendante de la concentration externe en spermidine. Il est possible dans ce mutant de mesurer l'influence des polyamines sur le taux de décalage de phase en -1 pour le ScV et en +1 pour le Ty1. Le décalage de phase en -1 n'est pas affecté. En revanche, le décalage de phase en +1 augmente fortement en fonction de la privation en spermidine dans le milieu de culture, jusqu'à 10 fois, tout en inhibant la transposition d'un Ty1 marqué (Balasundaram *et al.* 1994). Des études ultérieures ont montré que c'était la balance entre spermidine et putrescine (produit de dégradation des polyamines) qui influence l'appareil traductionnel dans sa manière de traiter le signal de recodage du Ty1. Cette influence n'est sensible que si le site A est inoccupé, puisque une surexpression de l'ARNt cognat inhibe ces effets stimulateurs (Balasundaram *et al.* 1994).

Comme nous l'avons déjà vu (cf. introduction § B.2.d, ou Rom *et al.* 1994), la production de l'antizyme de l'ornithine décarboxylase des eucaryotes supérieurs subit une boucle de rétrocontrôle négatif. Cette régulation est de type traductionnel. Tout comme le décalage de phase en +1 du Ty1 chez la levure est régulé par les polyamines, le décalage de phase en +1 de l'antizyme est stimulé en présence de spermidine, spermine ou putrescine (Matsufuji *et al.* 1995). Les polyamines n'agissent pas au niveau du pseudo-nœud ou du codon stop, puisqu'un site dépourvu de stop ou de pseudo-nœud garde cette régulation (Atkins *et al.* 1996). *In vitro*, les polyamines ne stimulent pas le décalage de phase en -1 du MMTV (résultats non publiés, cités dans Matsufuji *et al.* 1995). Le mécanisme de stimulation spécifique par les polyamines du décalage de phase en +1 n'est pas élucidé à ce jour, mais on pense qu'il peut s'agir d'une distorsion par les polyamines du site de décodage auquel les sites de décalage de phase en +1 (site A inoccupé) seraient particulièrement sensibles (Atkins *et al.* 1996).

### **III.A.6. Substances antibiotiques**

La cycloheximide, qui interfère avec la translocation, et la puromycine, qui interfère avec l'activité peptidyl transférase du ribosome diminuent

*in vitro* la quantité de produit de traduction, mais elles diminuent plus encore la quantité de produit de traduction nécessitant un décalage de phase pour une cible de décalage en -1 issue du HTLV-2 (Honigman *et al.* 1995). Deux inhibiteurs du transfert du peptidyl ont été testés *in vivo* chez *S. cerevisiae* et *in vitro* dans des systèmes de levure et de mammifères pour leur effets sur les décalages de phase en -1 du ScV. L'anisomycine, qui inhibe la liaison de l'aminocyl-ARNt au site accepteur du centre de transfert peptidyle, diminue le taux de décalage de phase en -1, alors que la sparsomycine, qui stimule la liaison du peptidyl-ARNt au site donneur, augmente le taux de décalage de phase en -1 (Dinman *et al.* 1997b). L'anisomycine et la sparsomycine pourraient être spécifiques du décalage de phase en -1 du ScV, puisque l'anisomycine n'a pas d'effets sur une cible issue du VIH-1 (pAC1789) à des doses sub-létales d'anisomycine dans des cellules NIH3T3 en culture ou dans la levure (Stéphanie Boissard, Laure Bidou, communication personnelle), et ni la sparsomycine, ni l'anisomycine n'ont d'effet *in vitro* sur cette même cible (Hung *et al.* 1998). De plus, l'anisomycine et la sparsomycine n'ont pas d'effet sur le décalage de phase en +1 du Ty1 (Dinman *et al.* 1997b).

Les aminoglycosides sont une classe d'antibiotiques qui altèrent la fidélité de la traduction, en augmentant le taux d'erreurs faux sens (Grant *et al.* 1989) ou non-sens (Cassan *et al.* 1990). Chez *E. coli*, la streptomycine, dans un contexte *rpsL*, diminue le taux de décalage de phase en +1 d'une cible issue de RF2 modifiée (codon rare UGG à la place de l'UAG, cf. § III.A.2.a, ou Siple *et al.* 1993). Par contre il n'y a pas modification du décalage de phase en -1 pour le bactériophage T7 (Siple *et al.* 1991). La paromomycine n'influence pas le décalage de phase *in vivo* dans des cellules de mamifère en culture au niveau de l'heptamère du VIH-1 (Cassan *et al.* 1990), ou du ScV chez la levure (Cui *et al.* 1998). Mais dans le contexte mutant *mof4-1*, le décalage de phase du ScV devient sensible à la paromomycine : il augmente jusqu'à plus de deux fois, en fonction des doses de paromomycine utilisées (Cui *et al.* 1998).

La Pokeweed Antiviral proteine (PAP) est une protéine de *Phytolacca americana* qui dépurine spécifiquement une adénine de l'hélice alpha-sarcine de l'ARNr 25S, et interfère ainsi avec la translocation ribosomique, et la réplication virale. La PAP surexprimée chez *S. cerevisiae* n'affecte pas le décalage de phase en -1 du virus ScV, bien que cette protéine soit connue pour inhiber la réplication du VIH-1 *ex vivo*. Par contre, elle fait chuter de 80% le décalage de phase en +1 du Ty1 *in vivo* et *in vitro* (Tumer *et al.* 1998).

Le RG501, identifié lors du crible de 56000 substances, inhibe le décalage de phase en -1 du VIH-1, en se liant spécifiquement au niveau de la tige-boucle. L'effet n'est pas le même en fonction des cibles retrovirales étudiées (tige-boucles de HIV-2, HTLV-1, HTLV-2, SIV). Ces résultats ont été confirmés *in vivo* par des tests immunologiques dans des cellules infectées chroniquement par VIH-1, avec un ratio gag/gag-pol augmenté d'un facteur deux à trois; et une diminution jusqu'à un facteur cinq de la réplication virale (Hung *et al.* 1998).

### **III.A.7. Conditions de croissance**

En général, les mesures des taux de recodage *in vivo* chez les micro-organismes se font en phase exponentielle de croissance, en milieu riche; ceci ne reflète pas les conditions de croissance des micro-organismes en milieu naturel (Kolter *et al.* 1993). Je relate ici les quelques expériences qui permettent d'indiquer quelles sont les influences des conditions générales de croissance (carence en sucre, type de substrat carboné, carence en acides aminés, chocs thermiques) sur différents systèmes de recodage. Les "senseurs" de ces différentes conditions de croissance ne sont pas connus, mais il pourrait s'agir de variations de facteurs influençant la fidélité de la traduction, comme les taux intracellulaires en RF (Adamski *et al.* 1994) ou en ARNt (Dong *et al.* 1996) en fonction des phases de croissance, ou la réponse stringente, médiée par le ppGpp (Ninio 1986).

#### **III.A.7.a. Phases de croissance**

Toutes les expériences relatées ci-dessous ont été menées chez *E. coli*. Un codon Phénylalanine UUU en 5' du gène *argI* provoque un décalage de phase en +1 à un taux de 3% à mi-phase exponentielle, et de 16% en phase exponentielle tardive (Fu *et al.* 1994). Ce décalage de phase provoque une erreur de processivité, une traduction avortée de *argI*, et n'a pas apparemment de rôle biologique. Il ne s'agit donc pas, selon la définition donnée en introduction, de recodage. Les taux de décalage de phase en +1 au niveau d'un site modifié de RF2 de type UUU CGG augmente 3 à 4 fois en phase stationnaire par rapport à la phase exponentielle, indépendamment de la présence de la séquence SD (Schwartz *et al.* 1997). Le taux de décalage en -1 provoqué par

l'heptamère U UUC AAG augmente d'un facteur 20 en phase stationnaire (Barak *et al.* 1996).

Des études spécifiques de l'influence de la phase de croissance sur la translecture de différents stops, d'un site de décalage de phase en +1 et en -1 au niveau d'un homopolymère U<sub>9</sub> ont été menées chez *E. coli*. Le taux de translecture diminue d'un facteur 2 à 5 entre le milieu de phase exponentielle et la phase stationnaire tardive. L'effet est plus immédiat et plus marqué pour un stop fortement translucide. Les taux de décalage de phase en +1 et -1 augmentent tout au long de la phase exponentielle, puis diminuent à partir de l'entrée en phase stationnaire, jusqu'à revenir à un taux semblable à celui de milieu de phase exponentielle (Wentzel *et al.* 1998).

Ces données, contradictoires, attendent des vérifications, avec une homogénéisation des conditions et milieux de culture, de sites de recodage, ou même de façons de calculer l'augmentation du taux de recodage : activité spécifique par rapport à un témoin en phase (Fu *et al.* 1994; Schwartz *et al.* 1997), activité normalisée par un témoin interne de traduction (Wentzel *et al.* 1998), augmentation de l'activité spécifique normalisée par le taux différentiel de synthèse de l'enzyme (*sic*) (Barak *et al.* 1996). Ces expériences indiquent en tous cas qu'il y a un changement radical du comportement de la synthèse protéique et de la fidélité de la traduction au moment de la transition entre phase exponentielle et phase stationnaire.

### **III.A.7.b. Carence en acides aminés**

La traduction est dépendante de la concentration du milieu en acides aminés, par le biais des ARNt chargés en acides aminés ou non. Il a été prouvé que des ARNt non chargés, en présence d'un analogue d'acide aminé (tryptophane hydroxamate), sont capable d'inhiber légèrement le décalage de phase en +1 d'un site RF2 modifié (codon affamé tryptophane UGG à la place du stop) (Gao *et al.* 1995). En ce qui concerne le décalage de phase en -1, les heptamères U<sub>6</sub>A ou U<sub>7</sub> issus du VIH ont été étudiés chez *E. coli* en présence d'analogues d'acides aminés, simulant une carence en acides aminés spécifiques. La phénylalanine-hydroxamate et l'arginine-méthyl-ester n'ont pas d'effet sur le taux de décalage de phase de la cible U<sub>6</sub>A, alors que la leucine-méthyl-ester inhibe fortement ce décalage de phase. Le codon leucine UUA impliqué

dans ce décalage de phase, a été remplacé dans la construction U<sub>7</sub> par un codon Phe : la leucine-méthyl-ester n'a aucun effet sur le taux de décalage de phase en -1 dans ce contexte. Les protéines synthétisées après décalage de phase en carence de leucine ou sans carence ont été séquencées. La limitation en leucine favorise un décalage d'un seul ARNt au niveau du premier codon de l'heptamère (Yelverton *et al.* 1994). Il eût été intéressant de connaître l'effet de la phénylalanine-hydroxamate en présence de leucine-méthyl-ester sur la cible U<sub>6A</sub> (décalage en singleton d'un seul ARNt<sub>Phe</sub>), ou l'effet de la phénylalanine-hydroxamate seule sur la cible U<sub>7</sub> (décalage de deux ARNt<sub>Phe</sub>).

### **III.A.7.c. Source de carbone**

Le décalage de phase en +1 du Ty1 dans une souche sauvage est plus de deux fois plus élevé quand elle est cultivée en galactose qu'en raffinose, alors que le décalage de phase en -1 (virus ScV) n'est pas modifié (Tumer *et al.* 1998).

L'efficacité de suppression de mutants *SUP35* et *SUP45* (eRF3 et eRF1) chez *S. cerevisiae* est variable en fonction du substrat carboné utilisé. La croissance en milieux non-fermentescibles (éthanol, glycérol, galactose) augmente l'efficacité de suppression par rapport à un milieu glucose. La terminaison est sans doute sensible à la repression catabolique, puisque l'ajout d'une faible quantité de glucose dans un milieu non-fermentescible ramène la suppression à son niveau de base en glucose (Inge Vechtomov *et al.* 1994). La suppression augmente également chez des mutants sans ADN mitochondrial, incapables de pousser en milieu non-fermentescible (Inge Vechtomov *et al.* 1994).

### **III.A.7.d. Sénescence**

Il ne semble pas à ce jour que la sénescence des cellules d'eucaryotes supérieurs soit accompagnée d'une différence de fidélité de la traduction *ex vivo* ou *in vivo* (Parker *et al.* 1981), même si des polémiques persistent à ce sujet (pour revue Gallant *et al.* 1997, et papiers efférents). Mais réciproquement, de nombreux mutants de l'appareil traductionnel influent sur la longévité (Belcour *et al.* 1991) et la différenciation (Coppin-Raynal *et al.* 1988) chez *Podospora anserina*. L'un des locus identifiés, *AS4* code le facteur EF-1A (Silar *et al.* 1994). EF-1A est codé par

plusieurs gènes dans de nombreux organismes, dont l'expression varie selon les tissus de manière apparemment mutuellement exclusive : une des deux formes est trouvée dans la majorité des tissus somatiques, alors que l'autre est limitée aux tissus en phase terminale de différenciation (neurones, cellules cardiaques, os,...); les deux formes identifiées chez le lapin n'ont pas la même constante de dissociation vis-à-vis du GDP (Kahns *et al.* 1998). Ces faisceaux de données pourraient indiquer un rôle d'EF-1A dans la sénescence et/ou la différenciation, mais rien n'indique à ce jour que cette influence se fasse par le biais de la fidélité de la traduction. Au contraire, chez *P. anserina*, des mutants plus ou moins fidèles peuvent augmenter ou diminuer la longévité; ce n'est donc pas le taux d'erreur *per se* qui contrôle la sénescence (Belcour *et al.* 1991).

### **III.A.8. Ferritine**

Une protéine liée à une tige boucle peut augmenter le taux de décalage de phase en -1 : une tige-boucle IRE (iron response element) en 5' de l'ARNm de la chaîne H de la ferritine est la cible de protéines cytoplasmiques IRP uniquement en présence de fer, et régule en conséquence la traduction de cet ARNm. Cette tige-boucle a été placée en aval de la séquence glissante du VIH-1, dans un système de mesure du même type que celui décrit au chapitre I. Dans des cellules de hamster BHK-21 en culture, le taux de décalage de phase en -1 est augmenté pour ces constructions en absence de fer, ou lorsque les IRPs sont surexprimées. Cette augmentation n'est pas sensible pour une tige-boucle mutée connue pour être déficiente dans la liaison aux IRPs (Kollmus *et al.* 1996). Ces expériences démontrent qu'il est possible à une protéine se liant à une structure secondaire d'augmenter le décalage de phase en -1. L'effet de cette protéine est sans doute uniquement d'ordre stérique, puisqu'elle provient d'un système non relié au décalage de phase.

Un autre système dérivé de l'IRE devrait permettre l'isolement de facteurs *trans* liant une structure secondaire. Une structure secondaire est placée en 5' de la *gfp*. Si une protéine vient lier cette structure, l'initiation de la traduction est inhibée, ce qui peut être détecté par la mesure de l'activité fluorescente (Paraskeva *et al.* 1998). Cette technique devrait permettre d'isoler des facteurs *trans* spécifiques des structures secondaires agissant sur le recodage.

### III.B. Crible de facteurs modulant la translecture du TMV

Après avoir caractérisé les séquences *cis* du recodage *in vivo* chez la levure, et validé l'utilisation de cet organisme par des expériences menées en parallèle dans les cellules de mammifère en culture, nous avons décidé de chercher à identifier chez la levure des facteurs *trans* du recodage, décalage de phase en -1 du VIH-1 ou translecture de codons stops.

Trois approches différentes sont envisageables :

(i) l'étude de l'effet de facteurs connus comme étant des modulateurs généraux de la fidélité de la traduction. Cette approche permet de se baser sur la fonction biochimique connue des facteurs étudiés, et donc d'interpréter l'effet sur le décalage de phase avec des bases fonctionnelles,

(ii) l'étude de facteurs paralogues à des facteurs déjà connus, ce qui permet d'identifier des nouveaux facteurs, tout en se basant dans une certaine mesure sur les avantages du (i),

(iii) entreprendre une analyse génétique sans autre *a priori* que celui du crible utilisé pour isoler les mutants d'intérêt.

En profitant de notre système de mesure efficace, nous avons entrepris plusieurs de ces approches en parallèle. J'ai étudié chez la levure l'effet de la surexpression du facteur EF4 (YNR163c), similaire à EF2 (EF-G chez les procaryotes), cloné par hasard lors d'une PCR (Franco Fasiolo, communication personnelle), du facteur *HBS1*, présentant des similitudes de séquence avec EF-1A (Nelson *et al.* 1992) en collaboration avec Sidonie Barbeau et Olivier Jeanjean, et étudié l'effet de 5 mutants, dont *sup35* (eRF3) et *sup45* (eRF1), isolés dans un crible d'allosuppression (Erbs 1996). Aucune de ces études n'a indiqué clairement un effet de ces mutants sur le décalage de phase en +1 (Ty), en -1 (VIH-1) ou sur la translecture du TMV. En parallèle, j'ai entrepris le crible de mutants de levure affectés dans la translecture du TMV.

*“...une analyse génétique des séquences provoquant une faible terminaison en utilisant des systèmes plus faciles à manier (par exemple S. cerevisiæ ou Arabidopsis thaliana) pourrait aider à identifier d'autres composants cellulaires qui participent au processus de terminaison”*

conclue le papier décrivant l'analyse des séquences *cis* nécessaires à la translecture du TMV (Skuzeski *et al.* 1991). En effet, les limites de l'analyse des séquences *cis* atteintes, et ces études ne suffisant pas à comprendre complètement le recodage, quelques groupes ont envisagé des cribles génétiques permettant d'isoler de nouveaux facteurs responsables de ces taux élevés du recodage. Je vais décrire la méthodologie que nous exploitons au laboratoire, puis les résultats préliminaires d'un crible génétique que j'ai entrepris afin d'isoler des mutants génomiques modulant l'efficacité de translecture du TMV.

### **III.B.1. Principe**

- (i) Un gène rapporteur est construit de telle sorte que son expression soit dépendante de la translecture du codon stop issu du TMV,
- (ii) Des mutants exprimant plus ou moins ce gène rapporteur sont criblés,
- (iii) Ces mutants sont retestés avec le système de mesure *lacZ-luc* permettant de sélectionner spécifiquement les mutants affectés au niveau de la translecture,
- (iv) Après analyse génétique, les mutants seront clonés sur la base de phénotypes secondaires.

### **III.B.2. Gène rapporteur**

La mesure de l'activité luciférase de la luciole ne peut se faire que sur des extraits protéiques. Les vecteurs pAC décrits précédemment, portant un gène de fusion *lacZ-luc* ne sont pas utilisables *in vivo* pour un crible chez la levure. Nous avons alors élaboré, avec Hongmei Liu et Olivier Namy, un crible utilisable chez *S. cerevisiae* sur boîte de Petri. Après plusieurs essais infructueux avec la *gfp* (Green Fluorescent Protein), nous avons décidé d'utiliser le gène *lacZ* qui avait déjà servi pour des quantifications *in vivo* chez la levure, et spécialement dans un crible similaire portant sur la cible du décalage de phase en -1 du ScV (mutants *mof*, Dinman *et al.* 1994).

Nous avons choisi deux types de recodage comme cibles : le site complet du décalage de phase en -1 du VIH-1, et le codon ambre soumis à

translecture du TMV dans son contexte. Ces deux types de recodage sont efficaces à un taux de 8% et 20% respectivement chez la levure (Stahl *et al.* 1995). Ces hauts taux de recodage nous permettent d'envisager la recherche de facteurs augmentant ou diminuant ce taux. Des mutants fidélisant la traduction ont été jusqu'à présent identifiés par recherche de suppresseurs de suppresseurs (anti-suppresseurs), c'est-à-dire de révertants plus fidèles de mutants infidèles. L'identification de tels "fidélisateurs" de la terminaison dans un fond sauvage est une approche nouvelle. En ce qui concerne le maintien du cadre de lecture, des "anti-suppresseurs de décalage de phase" n'ont jamais été recherchés directement ou indirectement. L'obtention directe de mutants plus fidèles est donc une des nouveautés du crible que nous avons envisagé.

Pour pouvoir, à partir d'un niveau de base, détecter des mutants plus ou moins exprimés, les cibles de recodage doivent être suffisamment exprimées chez le sauvage. Nous avons donc choisi d'exprimer la -galactosidase sous la dépendance d'un promoteur fort, non régulé, celui d'une des deux isoformes de l'asparagine synthétase de *S. cerevisiae*, le gène *ASN1* (Monique Bolotin-Fukuhara, communication personnelle). La -galactosidase a été fusionnée en phase avec les 417 premiers codons de la phase codante d'*ASN1* sous le contrôle de son propre promoteur (pXZL6, Hongmei Liu, communication personnelle). Nous avons inséré à la jonction entre *asn1* et *lacZ* les cibles de recodage du VIH-1 ou du TMV. Ces constructions ont par la suite été insérées au locus *ASN1* par recombinaison homologue après co-transformation avec un plasmide sélectionnable, et en criblant les transformants pour la présence d'activité -galactosidase sur filtre dans la souche Y349. Les souches obtenues, vérifiées par Southern Blot (O. Namy & H. Liu, communication personnelle) sont : YAL1 pour la construction *asn1-lacZ* en phase, YONX7 pour la construction *asn1-TMV-lacZ*, Y9L2 pour la construction *asn1-VIH-lacZ* (Fig. 14).

### **III.B.3. Construction de la souche de mutagenèse**

Le fragment Pst1 issu du plasmide pSGZ-TMV, qui porte le gène de fusion promoteur pSV40-*lacZ-TMV-luc* (Stahl *et al.* 1995) et est flanqué à ses extrémités des parties 5' et 3' du locus ARG4, a été intégré au locus ARG4 par recombinaison homologue (même technique que ci-dessus pour le locus *ASN1*, Marie-Claire Fontyne, communication personnelle) dans la

souche ORT128. Le transformant YTMV obtenu a été vérifié pour son auxotrophie pour l'arginine.

La souche YTMV a été croisée avec la souche YONX7 pour obtenir la souche diploïde YTX. Après sporulation et dissection de tétrades, la souche

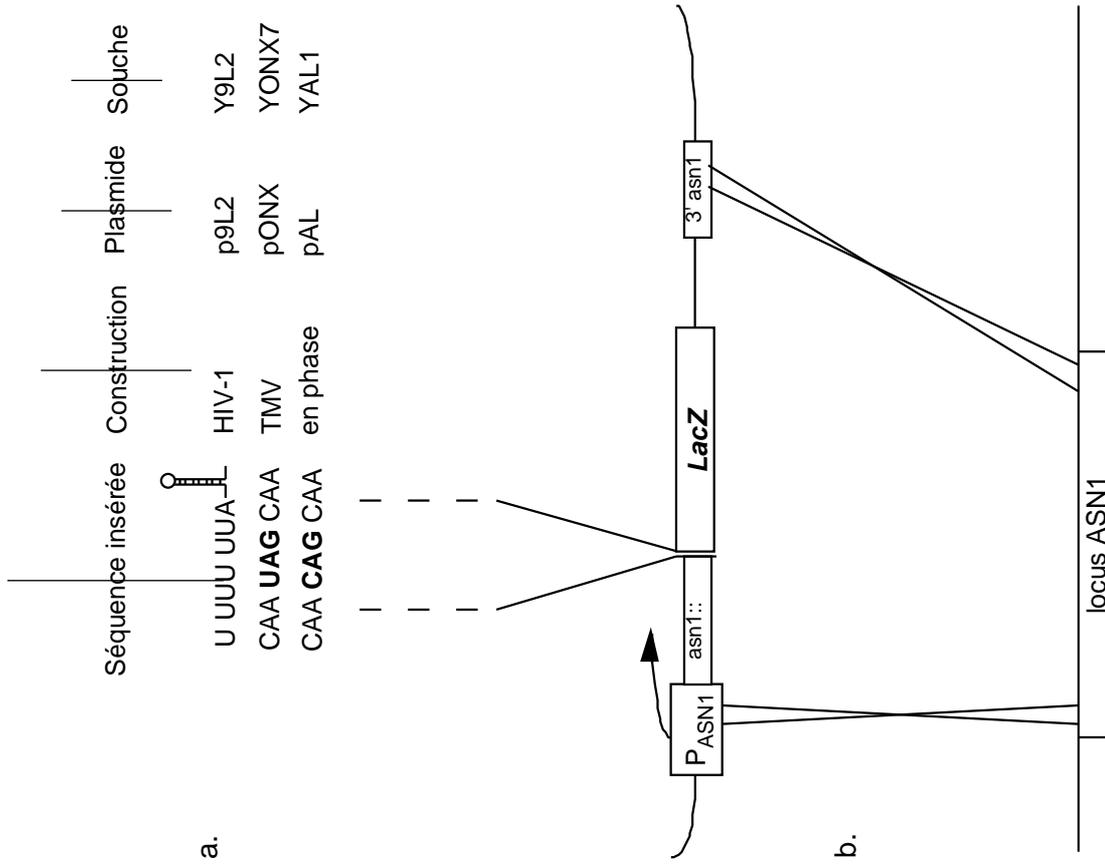
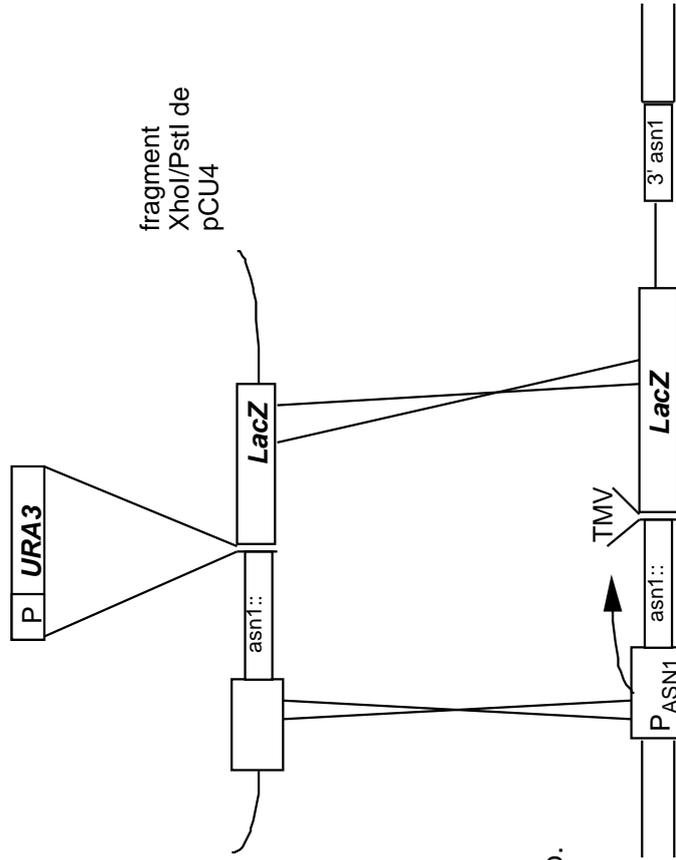


Figure 14 : Construction des souches contenant le gène rapporteur *LacZ* derrière une cible de recodage  
 a- Construction d'un plasmide intégratif  
 b- Intégration au locus ASN1 par recombinaison homologue

a.



b.

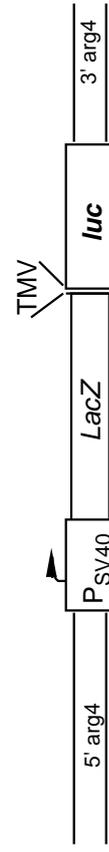


Figure 15 : Souche YTX8b et disruption du locus ASN1 par recombinaison homologue

a- Plasmides pCU4/8  
 b- Génotype partiel de la souche YTX8b

YTX8b a été obtenue, avec le génotype suivant : *Mat a* , *arg4Δ2060::pACTMV*, *asn1-TMV-lacZ*, *ade2*, *trp1-289*, *leu2,3-112*, *lys2-Δ201*, *ura3-52*, *his3Δ* (Fig. 15). Cette souche a été utilisée lors de la mutagenèse.

### **III.B.4. Mutagenèse**

J'ai choisi d'utiliser les ultraviolets comme agent mutagène. Le taux de survie de chaque mutagenèse est évaluée par rapport à un échantillon non irradié; le taux de survie moyen est de 3%. On peut estimer qu'à ce taux de survie, 30000 mutants permettraient d'avoir touché au moins une fois chaque gène (François Lacroute, communication personnelle).

### **III.B.5. Crible primaire**

Le crible primaire se fait sur la base de la coloration bleue des colonies de levure en présence de X-Gal. Deux cribles successifs sont entrepris. Le premier crible se fait par réplique sur papier filtre, immersion dans l'azote liquide, puis révélation en tampon Z + X-Gal. Les mutants intéressants sont prélevés directement sur le filtre. Après ré-isolément, les mutants sont déposés en gouttes sur boîte et remis à pousser, puis testés pour leur activité -galactosidase par la technique du Top-Agarose, après perforations des cellules au chloroforme. L'utilisation successive de deux techniques différentes devrait permettre d'optimiser ce crible primaire.

### **III.B.6. Crible secondaire**

La variation d'expression d'un gène peut être due à des paramètres aussi divers que son taux d'expression (transcription et traduction), la stabilité de son ARNm, des modifications post-transcriptionnelles, des variations de co-facteurs nécessaires à son activité enzymatique, etc... Ce sont autant de mutants de ces étapes différentes qui vont être isolés dans le premier crible. Mais en utilisant un crible secondaire reposant sur une construction avec un promoteur et un rapporteur différent, et possédant un témoin interne de normalisation (la -galactosidase), on peut discriminer *a posteriori* les mutants affectant spécifiquement le recodage,

des autres mutants d'expression : l'identification d'un mutant du recodage se fait en mesurant une variation du rapport des activités luciférase (mesure) et  $\beta$ -galactosidase (normalisation pour l'expression de l'ARNm). Ce type d'approche devrait nous permettre notamment d'éviter l'analyse de mutants tels que ceux impliqués dans la voie de dégradation des ARNm contenant un stop précoce (Non-sense mediated mRNA decay, NMD) (cf. § III.A.2.b ou Leeds *et al.* 1991; Leeds *et al.* 1992; Lee *et al.* 1995b).

Pour cela, la construction pAC-TMV au locus ARG4 va nous servir de crible secondaire. Mais l'utilisation de ce système nécessite le dosage de l'activité enzymatique du témoin interne, la  $\beta$ -galactosidase. Le crible primaire repose lui aussi sur cette enzyme. Le gène de fusion *asn1-TMV-lacZ* produit beaucoup plus de  $\beta$ -galactosidase que le gène pSV40-*lacZ-TMV-luc* (plus de trois ordres de grandeur de différence) ce qui permettait de négliger l'expression de la  $\beta$ -galactosidase du pSV40 lors du crible primaire, mais le contraire est impossible. J'ai inséré, à la jonction entre le début de la phase codante de *ASN1* et *lacZ* du plasmide pAL, un gène de sélection *URA3* (fragment BglII issu du pFL44L, Bonneaud *et al.* 1991). Il a été vérifié que ce plasmide pALURA9 ne produisait pas d'activité  $\beta$ -galactosidase détectable chez *S. cerevisiae*. Un fragment XhoI-ClaI du pALURA9 contenant la construction *asn1-URA3-lacZ* a été sous-cloné dans le plasmide pBCSK+ (*Stratagène*), ne contenant pas d'origine de répllication fonctionnelle chez la levure (plasmide pCU4). La transformation de la souche YTX8b par ce plasmide pCU4 digéré par XhoI et ClaI inactive dans la majorité des transformants [Ura<sup>+</sup>] obtenus l'activité  $\beta$ -galactosidase au locus ASN1. (Fig. 15)

Les mutants intéressants issus du premier crible sont transformés de cette façon, et deux transformants sont ré-isolés. Les activités luciférase et  $\beta$ -galactosidase sont mesurées à partir d'extraits protéiques bruts.

### **III.B.7. Phénotypes secondaires**

Les supresseurs précédemment caractérisés chez *S. cerevisiae* ont la particularité de présenter en général des phénotypes associés secondaires pléiotropiques, comme une croissance lente, une thermo/cryosensibilité (Wakem *et al.* 1990), parfois associée au phénotype *cdc* (Dinman *et al.* 1994), une déficience de respiration, des défauts de sporulation et de germination, des défauts de caryogamie et de cytoduction (Inge

Vechtomov *et al.* 1994; Liebman *et al.* 1995), une sensibilité aux conditions hyperosmotiques (Singh 1977; Wakem *et al.* 1990; Inge Vechtomov *et al.* 1994), et une sensibilité accrue à différentes drogues : paromomycine (Wakem *et al.* 1990), cycloheximide, trichodermine (Inge Vechtomov *et al.* 1994) novobiocine (Pocklington *et al.* 1990), benomyl (Inge Vechtomov *et al.* 1994). Certains de ces phénotypes nous permettent de différencier les différents suppresseurs ou anti-suppresseurs isolés, et surtout de pouvoir cloner le gène sauvage par complémentation fonctionnelle grâce à une sélection et non pas avec le crible originel (-galactosidase) plus lourd à manier. La présence des mutations *ade2-101* (ochre) et *trp1-289* (ambre) permettent éventuellement de différencier des suppresseurs codons spécifiques de suppresseurs omnipotents.

### **III.B.8. Résultats**

mutagenèses indépendantes ont été effectuées pour un total de 42000 colonies criblées. La stratégie est rappelée Fig. 16. Le premier crible a permis d'isoler 368 mutants plus bleus et 129 moins bleus. Lors du deuxième passage, j'ai retenu 175 et 58 mutants respectivement. Je n'ai pour l'instant que des résultats préliminaires en ce qui concerne le taux de translecture (rapport activité luciférase/ -galactosidase, après transformation par pCU4). Un tableau récapitulatif de ces résultats préliminaires est présenté en annexe.

La majorité des clones moins bleus sont [Gly<sup>-</sup>] (déficientes pour la respiration mitochondriale). Le phénotype [Gly<sup>-</sup>] n'est sans doute pas un phénotype secondaire spécifique des anti-suppresseurs (moins bleus). J'ai mesuré le nombre de colonies [Gly<sup>-</sup>] après mutagenèse, qui est de l'ordre de 50% du total. Ces colonies [Gly<sup>-</sup>] poussent également moins vite en milieu riche glucose, et je suggère que la différence de taille des colonies provoque une coloration bleue moins intense. La majorité (environ 90%) des colonies isolées comme moins bleues sont [Gly<sup>-</sup>], sans doute des mutants *rho*<sup>-</sup> (ayant perdu une partie de leur ADN mitochondrial), et donc non relatives à notre crible. On ne peut pas non plus les écarter d'office, puisque les capacités respiratoires des mutants est un des phénotypes secondaires que l'on s'attend à trouver dans des mutants affectés dans la terminaison (cf. ci-dessus, § III.B.7). L'analyse de ces mutants est compliquée par le fait que le taux apparent de translecture (le rapport activité luciférase/ -galactosidase) est dix fois plus faible en

contexte [Gly<sup>-</sup>] qu'en contexte [Gly<sup>+</sup>] . Par ailleurs il a été rapporté qu'une souche déficiente pour la respiration aurait un taux de suppression augmenté



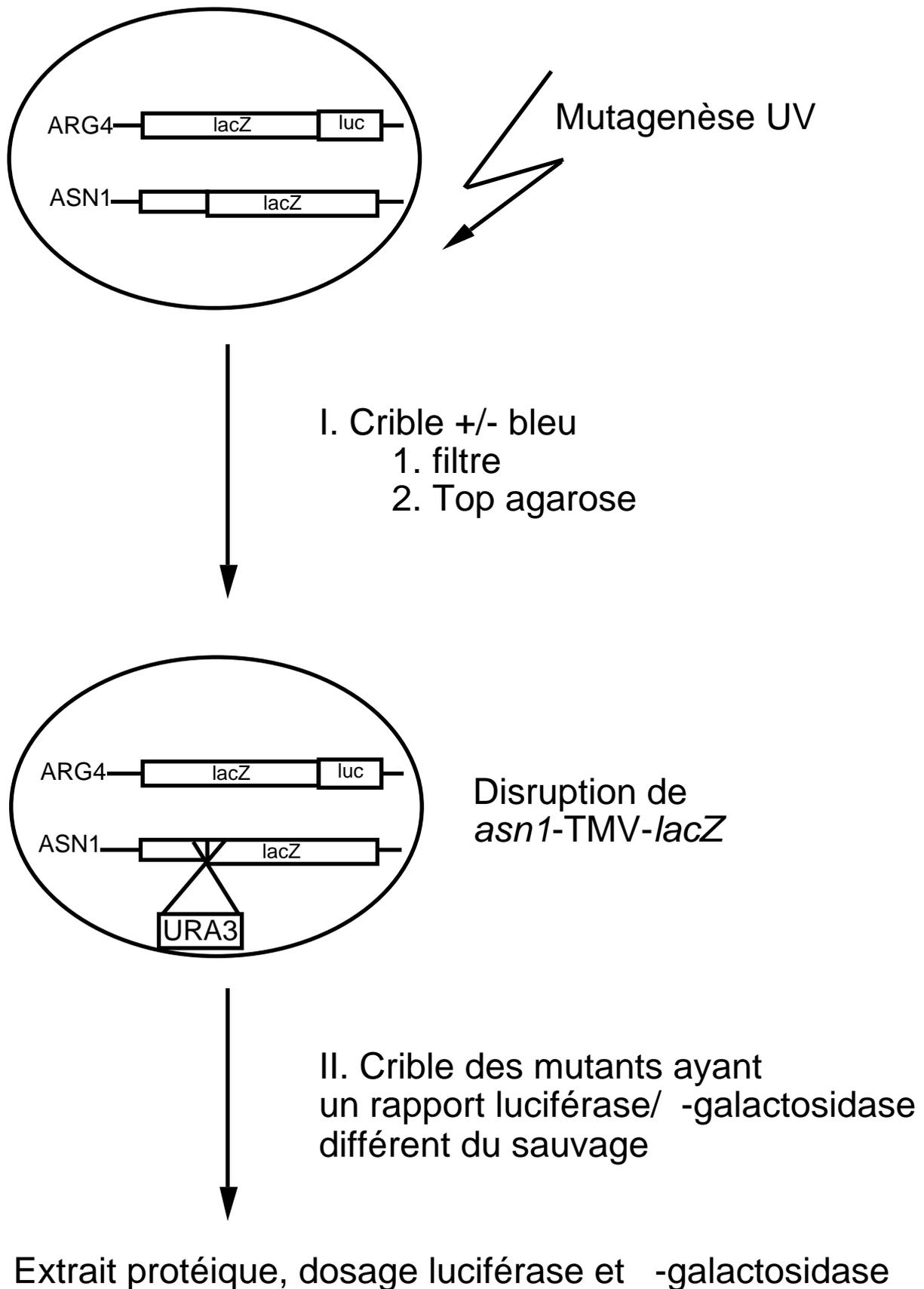


Figure 16 : Crible de facteurs trans sur la transfection du TMV

(Inge Vechtomov *et al.* 1994). Une souche sauvage YTX8b [Gly<sup>-</sup>] a été obtenue récemment afin de servir de référence.

Dans une expérience préliminaire portant sur 27 mutants, un tiers sont affectés de plus de 15% dans leur rapport d'activité luciférase/ -galactosidase par rapport au sauvage. Le plus important écart au sauvage est d'un facteur 1,5. L'un des mutants isolé comme étant plus bleu présente un rapport 2 fois moins élevé que le sauvage; ce phénotype peut être expliqué si ce mutant est affecté dans l'activité de -galactosidase produite par la cellule (cofacteur,...) : il sera alors plus bleu au premier crible, mais le rapport luciférase/ -galactosidase chutera (ici d'un facteur 2) dans le second crible. Cette hypothèse n'a pu être testée. Il est difficile de discuter des ces résultats en n'ayant pas d'information définitive sur le rapport d'activités luciférase/ -galactosidase de la majorité des mutants, mais le crible primaire est déjà bien maîtrisé, et celui-ci peut donc être discuté. La difficulté dans le premier crible pour détecter des colonies nettement différenciables du sauvage, et le fait que sur les 300 mutants plus bleus isolés, aucun ne soit nettement plus bleu que le sauvage lors du deuxième crible, m'a mené à comparer les colorations de la souche sauvage YONX7 comportant une fusion *asn1-TMV-lacZ* et la souche YAL1 comportant une fusion en phase *asn1-lacZ* (Olivier Namy, communication personnelle, cf. Fig. 14). Les deux souches ne présentent pas une grande différence de coloration dans le deuxième crible coloré (Top-Agarose). On attend une différence d'un facteur 5, puisque le codon ambre en contexte TMV est translucide à un taux de 20%. Des résultats préliminaires obtenus sur des cellules perméabilisées au chloroforme indiquent que la souche YONX7 produit environ 50% de l'activité -galactosidase (rapporté à la DO<sub>600</sub>) de la souche YAL1. Plusieurs raisons peuvent expliquer cette observation :

(i) à cause du promoteur fort du gène *ASN1*, la cellule est saturée pour l'expression de la -galactosidase (au niveau transcriptionnel, post-transcriptionnel ou post-translationnel) et ne produit pas 5 fois plus d'activité -galactosidase dans la souche contenant la construction en phase que dans la construction contenant le stop du TMV,

(ii) la souche YONX7 contient un suppresseur d'ambre; ceci est envisageable, puisque l'intégration au locus *ASN1* a été sélectionnée sur boîte pour son phénotype bleu, et la souche YONX7 est la seule obtenue sur plus de 1000 co-transformants; un biais a pu avoir lieu dans ce crible qui a mené à l'isolement de cette seule souche bleue,

(iii) la forte expression de la cible TMV est capable de provoquer la titration d'un facteur nécessaire à la terminaison dans ce contexte.

Afin de tester ces hypothèses, j'ai disrupté dans la souche YONX7 le transgène *asn1-TMV-lacZ* à l'aide du plasmide pCU4 (cf. ci-dessus), pour quantifier dans cette souche la translecture grâce aux vecteurs pAC (nécessitant une souche *lacZ*-), et d'identifier un éventuel suppresseur évoqué au (ii). La mesure, dans cette souche YONX7::pCU4, du taux de translecture obtenu à partir du pACTMV par rapport au témoin en phase pACTQ est de 32%. Ce taux de translecture est proche de celui obtenu dans la souche mère Y349, et ne suffit pas à expliquer le taux apparent de 50% de translecture mesuré en comparant la souche YONX7 à la souche YAL1.

Pour mesurer une titration d'un éventuel facteur dans la terminaison de la cible TMV (3ème hypothèse), j'ai également mesuré l'activité luciférase dans les souches YONX7 et YAL1, en la normalisant par la quantité de protéines totales de l'extrait brut (technique de Bradford, Biorad). Des résultats préliminaires obtenus sur un seul transformant YONX7 [pACTMV] m'indiquent un taux de translecture de plus de 50%. Ce résultat, qui reste à confirmer, irait dans le sens d'une titration d'un facteur trans nécessaire à la terminaison au codon CAA **UAG** CAA.

### **III.B.9. Faisabilité du crible**

Les contraintes de ce crible sont de deux ordres, en plus de la difficulté expérimentale liée au crible coloré.

La première contrainte est liée au fort taux de translecture apparent utilisé dans les souches YONX7 et YTX8b. L'obtention de mutants avec un taux de translecture encore plus élevé semble difficile, au vu des résultats déjà obtenus. Les expériences visant à comprendre ces forts taux de translecture sont en cours.

La deuxième contrainte est sur l'obtention de mutants plus fidèles dans la terminaison, et donc moins bleus dans notre crible. Une majorité des colonies moins bleues sont [Gly<sup>-</sup>], ce qui reflète certainement une perte de tout ou partie du génome mitochondrial. Le phénotype moins bleu a l'air lié au phénotype [Gly<sup>-</sup>], ce qui noie les mutants potentiellement intéressants (anti-suppresseurs) parmi ces mutants non pertinents. On peut envisager, afin de sélectionner plus facilement des mutants moins

bleus sur milieu Glycérol, ou dans une souche récemment obtenue YTX8b [Gly<sup>-</sup>]. Mais les conditions de suppression ne sont pas les mêmes sur milieu glycérol, et surtout la plupart des mutants supprimeurs sont déficients dans la respiration (Inge Vechtomov *et al.* 1994). Des expériences sont prévues afin de quantifier la translecture dans cette souche, et de mesurer l'impact de la croissance en milieu glycérol sur la translecture du TMV.

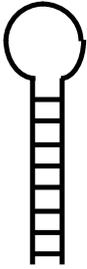
|                         | Site d'action                               | Stimulateur (pause)   |
|-------------------------|---|---|
| Décalage de phase en -1 | X XXY YYZ                                   |  |
| Décalage de phase en +1 | ARNt spécial<br><u>        </u><br>ABC DEFG | ARNt rare<br><u>        </u><br>ABC DEF G   |
| Translecture            | Stop  | Contexte défavorable  |

Figure 17 : Récapitulatif des mécanismes de recodage

## Conclusion

Même si par la diversité des sites et des systèmes, le recodage semble peu homogène, l'analyse des environnements et des conditions favorables au recodage nous permet de dégager un mécanisme commun : deux éléments *cis* présents sur l'ARNm (i) un site favorable mécaniquement à un événement de recodage, et (ii) un élément induisant une pause qui favorise cinétiquement le recodage par rapport au décodage au niveau de ce site (Fig. 17). Comme le reflète le chapitre III, peu de recherches systématiques ont été entreprises pour caractériser les facteurs solubles ou intrinsèques de l'appareil traductionnel qui modulent le recodage. La raison de ce manque d'étude des facteurs *trans* tient au fait que ce type d'approche a attendu que la caractérisation des éléments *cis* soit approfondie, bien que la recherche de facteurs *trans* ne puisse à mon sens que renforcer la compréhension du rôle des éléments *cis*. Une deuxième raison est peut-être le manque de systèmes de mesure simples, sensibles et fiables qui sont indispensables à cette phase d'études. Les autres systèmes *in vivo* n'utilisant qu'un seul gène rapporteur font appel à des phénotypes secondaires pour valider les effets observés (cf. début du § III.A ou pour revue Dinman *et al.* 1998). Ces phénotypes secondaires ne sont sensibles qu'à un changement supérieur à un facteur 2. Je pense que l'utilisation des vecteurs que j'ai construit et testé au cours de ma thèse est un atout dans cette recherche de facteurs *trans*. Je vais en rappeler ici brièvement les caractéristiques, avantages et inconvénients.

### Systeme de mesure, les éléments *cis*

La principale caractéristique de ces vecteurs pAC est de bénéficier d'une mesure interne de tous les facteurs influençant l'expression de l'ARNm considéré avant d'atteindre la cible de recodage. Cette valeur "recodage/expression" nous permet de comparer le recodage entre deux conditions différentes (typiquement, l'effet de facteurs *trans* par rapport aux conditions sauvages). Pour des mesures précises du taux de recodage *per se* nous disposons dans tous les cas d'une construction en phase dont les activités spécifiques sont identiques à celles de la construction test après recodage. Dans la plupart des cas, l'écart-type mesuré dans des expériences indépendantes est inférieur à 10% de la valeur mesurée, ce

qui permet de mettre en évidence des influences de faible amplitude sur le recodage. Nous pourrions par exemple vérifier et quantifier précisément les inter-relations entre initiation et maintien du cadre de lecture pour le mutant *mof2-1* (Cui *et al.* 1998), et les inter-relations entre terminaison, décalage de cadre de lecture et voie NMD (Cui *et al.* 1996; Weng *et al.* 1996b; Weng *et al.* 1996a), puisque le témoin interne est sensible à tous ces facteurs indépendamment du recodage.

Les autres avantages de ces vecteurs sont, en résumé :

(i) un site de clonage permettant d'insérer de façon orientée une cible de recodage et son témoin en phase à des fins d'analyse quantitative.

(ii) leur utilisation versatile dans de nombreux organismes grâce au promoteur SV40 : *E. coli*, *S. cerevisiae*, les cellules de mammifères en culture, des cellules de drosophile en culture (Jean-Pierre Rousset, communication personnelle), et sans doute beaucoup d'autres sont envisageables.

(iii) une utilisation aisée chez *S. cerevisiae* : les plasmides pAC sont fondés sur la famille de plasmides pFL, qui comprend des plasmides répliatifs multicopie (2 $\mu$ m), répliatifs mono-copies stables (ARS-CEN), des marqueurs de sélection divers comme TRP1, URA3, LEU2. Les marqueurs d'auxotrophies et les origines de réplication sont interchangeable grâce à un système de cassettes (Bonneaud *et al.* 1991). La présence de séquences issues du locus ARG4 de chaque côté de la construction permet d'utiliser la recombinaison homologue afin d'intégrer les constructions dans le génome de la levure.

D'autres approches ne sont pas, ou difficilement, abordables avec ce système rapporteur :

(i) ces plasmides ne sont pas utilisables *in vitro*, afin d'analyser plus finement l'influence de tel ou tel facteur : il a été suggéré que l'ARNm codant pour la -galactosidase était trop long pour être traduit efficacement *in vitro* (Grentzmann 1998). Il existe depuis peu des plasmides utilisables *in vitro* et *in vivo* dans des cellules de mammifères en culture, basés sur le même système que nos vecteurs (témoin interne/rapporteur de recodage), avec deux gènes *luc* d'origines différentes (méduse et ver luisant), dont les activités sont quantifiables séparément (Grentzmann 1998). Ce système est dans son utilisation en

tout point comparable au nôtre, mais n'a pas été adapté pour l'instant à l'utilisation dans la levure.

(ii) l'étude de conditions de croissance, ou d'allèles conditionnels, se heurte à un écueil majeur, qui est celui de la persistance des protéines synthétisées avant le passage aux conditions de mesure. Si les protéines synthétisées sont très stables, les variations de recodage n'étant mesurables que sur les protéines néo-synthétisées, on ne pourra sans doute pas mesurer ces effets avant la disparition de la majorité des protéines de fusions présentes avant l'application des nouvelles conditions. Ce problème méthodologique pourrait être résolu par l'utilisation d'un promoteur inductible : on induirait alors la synthèse de l'ARNm rapporteur au moment même de la transition entre deux conditions. Le choix d'un tel promoteur doit tenir compte du fait que les conditions d'induction ne doivent pas interférer avec la fidélité de traduction. Il est connu par exemple que le type de substrat carboné ou la température peuvent influencer sur les erreurs de traduction (Kawakami *et al.* 1993; Inge Vechtomov *et al.* 1994; Tumer *et al.* 1998).

(iii) la présence en 5' de la cible de recodage de la  $\beta$ -galactosidase ne permet pas d'obtenir la séquence protéique au niveau de la séquence de recodage; ceci peut s'avérer nécessaire pour distinguer des décalages en +1 et -2 (ou tout autre combinaison +/- 3 nucléotides) (Matsufuji *et al.* 1996), des décalages en -1 en singleton ou en tandem (Yelverton *et al.* 1994; Brierley *et al.* 1997), ou même entre la translecture et le hopping au dessus du stop (Weiss *et al.* 1987a). J'envisage de cloner un site de coupure endo-protéolytique juste en amont du site de clonage des cibles de recodage, afin de faciliter, après purification de la protéine (immunoprécipitation avec des anticorps dirigés contre la  $\beta$ -galactosidase, par exemple) le micro-séquençage de la jonction de la protéine de fusion  $\beta$ -galactosidase-luciférase, et de confirmer ainsi le lieu, le sens ou la nature du recodage.

Quelle que soit la précision de la mesure d'un taux de recodage, le système utilisé, la caractérisation fine d'un système de recodage dans un

hôte donné nécessite à mon avis plusieurs étapes indispensables à la compréhension du mécanisme utilisé :

(i) s'assurer qu'il s'agit d'un événement traductionnel, afin de le différencier de l'épissage alternatif ou de l'édition des ARN,

(ii) le séquençage protéique de la fusion renseigne très précisément sur le codon où s'est déroulé le recodage, et permet de différencier plusieurs systèmes similaires de recodage (cf. ci-dessus);

(iii) l'étude de mutants *cis* ponctuels et de délétions afin de cerner très précisément quel est le contexte et/ou le mécanisme utilisé pour effectuer et/ou stimuler le recodage

(iv) dans le cas d'études de structures secondaires, il faudrait analyser la structure tri-dimensionnelle par des sondes enzymatiques et chimiques, ou par RMN. Par exemple l'existence de la tige-boucle en aval de l'heptamère glissant du VIH-1, qui était postulé depuis 1985, n'a été étudié que très récemment grâce à des expériences de protection enzymatiques (Kang 1998). En cela, les études que nous avons mené sur l'influence de la stabilité de la tige-boucle en aval de la séquence glissante du VIH-1 ont fourni des éléments supplémentaires pour comprendre l'influence des structures secondaires sur le décalage de phase en -1. Nous disposons avec ces constructions de matériel utilisable pour des expériences structurales ou cinétiques.

Pour ce qui est de la découverte de facteurs modulant en *trans* le recodage, la caractérisation des cibles utilisées n'a pas besoin d'être complète, et j'essayerai même de montrer que, par un système d'aller-retour entre recodage et fonctions connues dans l'appareil traductionnel, on peut améliorer la compréhension de l'un et de l'autre.

## **Les facteurs *trans*, l'avenir ?**

Comme je l'ai défini au début de ma conclusion, le maintien de la phase de lecture est, tout comme la fidélité de lecture et la terminaison, un événement favorisé du décodage classique, qui dépend des caractéristiques cinétiques de l'appareil traductionnel. Les alternatives sont défavorisées cinétiquement lors d'une élongation normale; mais

lorsqu'une pause (ARNt rare, structure secondaire, codon stop dans un environnement faible) survient au niveau de séquences "anormales" (codons glissants, codons faisant appel à des ARNt occluant, codons stop favorables à la reconnaissance par un ARNt suppresseur), le ribosome peut changer de fonctionnement (Fig. 17). Des erreurs de traduction, d'un taux basal d'environ  $10^{-5}$ , peuvent ainsi être amplifiées à un taux particulièrement élevé de  $10^{-2}$  à  $10^{-1}$ . Il en découle trois questions :

(i) le recodage est-il un re-codage, une erreur de traduction, une fidélité exacerbée de la traduction ?

(ii) les facteurs trans, spécifiques ou issus de l'étude de la traduction usuelle modulant le recodage, peuvent-ils nous aider à comprendre les systèmes de recodage ?

(iii) réciproquement, l'analyse des systèmes du recodage peuvent-ils nous éclairer sur les mécanismes intrinsèques de la traduction qu'il utilise à ses fins ?

### **(i) Recodage, infidélité ou fidélité ?**

Cette question est un débat sémantique polémique, qui existe depuis que les bases du domaine du recodage (Gesteland *et al.* 1992), de la lecture alternative du code génétique (Farabaugh 1993) ont été posées. Quelques années auparavant, en analysant le rôle de EF-Tu comme standard interne de la fidélité de la traduction, Thompson évoquait déjà cette possibilité d'amplification des erreurs de traduction à un niveau compatible avec l'expression génique : "*Il est intéressant de spéculer sur le fait que de tels événements de faible probabilité soient toujours pensés en termes d'erreurs. [...] Il ne serait peut-être pas très surprenant qu'il y ait des cas où la synthèse d'une protéine importante dépende d'un événement de décodage qui, du point de vue du biologiste moléculaire, est une erreur.*" (Thompson 1988). On peut effectivement se demander si le recodage est un phénomène actif, ou s'il n'est qu'un opportunisme vis-à-vis des faiblesses de l'appareil traductionnel : "*...mais la différence entre ces événements et les erreurs aléatoires est-elle de niveau, de probabilité des erreurs, ou de nature ? Vu que les changements programmés de l'élongation sont phénoménologiquement divers, il n'y a pas de réponse à cette question*" (Farabaugh 1993). Mais aujourd'hui il apparaît que les événements de recodage pourraient être médiés par les

caractéristiques cinétiques intrinsèques de l'appareil traductionnel, caractéristiques qui sont essentiellement conservées au cours de l'évolution. À ce jour aucun facteur protéique modulant spécifiquement le recodage n'a été isolé, sauf en ce qui concerne le cas particulier de l'incorporation de la sélénocystéine. Ceci, et la diversité des systèmes étudiés, de manière hétérologue ou homologue, *in vivo* ou *in vitro*, la conservation des règles générales des bactéries à l'homme, semble suggérer que le recodage est une capacité intrinsèque de la cellule; qu'outre les facteurs *cis*, l'appareil traductionnel est nécessaire et suffisant pour permettre un décalage de phase ou une translecture. On peut alors se demander si ces événements sont des erreurs de traduction imposées par les éléments *cis*, ou au contraire une fidélité exacerbée imposée par l'appareil traductionnel lui-même. Le fait qu'un mutant *rpsL* hyper-fidèle pour la reconnaissance codon-anticodon augmente le taux de décalage de phase en +1 de RF2 montre que cet événement est dépendant de la fidélité de traduction, de la capacité de l'appareil traductionnel à discriminer entre décodage et erreurs de traduction par le proofreading. Les mécanismes de contrôle de la fidélité de la traduction, surtout en ce qui concerne la terminaison et le maintien du cadre de lecture, sont encore trop peu connus pour pouvoir prétendre de discriminer entre erreur de traduction ou fidélité exacerbée. Le terme recodage a l'avantage de définir une discipline, de délimiter un champ de recherche particulier dans le domaine du contrôle traductionnel. Je vais essayer de démontrer que les analyses croisées entre facteurs de la traduction connus et recodage permettront d'approfondir les connaissances du recodage, et du contrôle de la fidélité de traduction.

## **(ii) Facteurs *trans* comme révélateurs des mécanismes du recodage**

L'utilisation de facteurs *trans* connus de la traduction, et l'étude de leurs effets sur le recodage ont permis d'infirmer l'influence supposée des ARNt modifiés sur le rephasage en -1, que de nombreux indices impliquaient dans ce processus. De même, j'espère qu'il sera possible de confirmer ou d'infirmer l'hypothèse émise au chapitre II concernant le mutant de VIH-1 contenant une tige-boucle plus stable, mais stimulant moins le décalage de phase en -1, en comparant chez *E. coli* les taux de rephasage entre une souche sauvage et une souche *rpsL* dont le taux de traduction est plus lent. L'utilisation de facteurs connus de l'appareil

traductionnel peut, par la connaissance phénotypique ou biochimique que nous en avons, nous informer sur les mécanismes précis du recodage. De nombreux facteurs restent ainsi à exploiter dans cette voie.

Certains facteurs, par des effets différents sur telle ou telle cible nous permettront peut-être d'utiliser ces différences comme base d'étude. L'analyse des effets de certains antibiotiques ont révélé des effets différents entre le décalage de phase en -1 du ScV et du VIH-1. Ceci est très prometteur pour comprendre ce qui différencie les deux cibles : pause plus longue dans un cas ou dans l'autre, séquence heptamérique plus ou moins glissante, influence spécifique du pseudo-nœud ou de la tige-boucle ? Les effets différents de mutants de EF-1A ou autres (*mof2-1*, *mof4-1*), sur le décalage de phase en +1 ou en -1 permettront certainement à terme de comprendre les différences entre les deux décalages programmés, qui sont sans doute en relation avec l'occupation différente des sites A et P dans les deux cas.

La caractérisation de facteurs *trans* modulant les taux de rephasage ou de redéfinition est une des voies pour mieux comprendre le fonctionnement de ces systèmes de recodage (et de l'appareil traductionnel lui-même, cf. ci-après). Peu de cribles génétiques ont été élaborés dans ce sens. Le crible que j'ai entrepris a pour but d'isoler les facteurs responsables de la très mauvaise terminaison/très forte translecture de la séquence issue du TMV. Ce crible pourra nous renseigner sur les facteurs connus affectés dans leur reconnaissance du codon stop dans ce contexte (eRF1, eRF3, RRF, ARNr, ARNt, protéine ribosomique...). Des cribles basés sur des cibles de recodage permettent d'espérer mieux comprendre le fonctionnement des systèmes de rephasage et de translecture, mais également de comprendre quelle fonction de l'appareil traductionnel est impliquée dans la terminaison efficace ou les mécanismes de maintien du cadre de lecture. En généralisant cette démarche, nous compléterons les connaissances que nous avons de la fidélité de la traduction ou de la traduction en général, et identifierons peut-être de nouveaux facteurs de la traduction.

### **(iii) Le recodage comme modèle d'étude de la traduction**

La recherche de mutants modulant le recodage permet également d'identifier de nouveaux facteurs de la traduction et d'en comprendre le

rôle dans le contrôle de la fidélité de la traduction. Ce type d'approche a permis d'isoler un mutant du gène codant pour la protéine L9 modulant le hopping du gène *60* du bactériophage T4, et de donner un modèle précis de son action dans l'orientation de certaines hélices de l'ARNr 23S par rapport au site P (Adamski *et al.* 1996). Un autre crible a permis de suggérer que l'ARNr 5S aurait un rôle dans le maintien du cadre de lecture (Dinman *et al.* 1995). La découverte de mutants de l'initiation et de la voie NMD influant aussi sur le recodage montre l'intégration du processus traductionnel à un point inattendu (Cui *et al.* 1996), et qui nécessite des investigations supplémentaires.

D'une manière générale, les conditions très particulières (un taux élevé à un site précis) des événements tels que le décalage de phase ou la translecture de codons stops, permettent une quantification spécifique de cette erreur *versus* le taux d'erreurs de base sur l'ARNm. On peut imaginer ainsi mieux comprendre le rôle de l'ARNr dans la terminaison ou le maintien du cadre de lecture, comme dans le cas des mutants *mof9*, ou du rôle des séquences SD. La connaissance précise des conditions dans lesquelles se trouve l'appareil traductionnel au moment où se produit le recodage (comme l'occupation connue des sites A et P) permet de détailler très précisément les effets de tel ou tel facteur sur la traduction. Par exemple les séquences permettant la redéfinition de codons stops nous renseignent sur les environnements favorables ou défavorables pour une bonne terminaison. On peut aussi quantifier et analyser précisément *in vivo* l'effet d'antibiotiques sur des cibles connues (site A vide ou occupé), au lieu de constater leur effet global sur la traduction des ARNm. De même pour l'effet différentiel de différents mutants de EF-1A sur le décalage de phase en +1 ou en -1 : par un système d'aller-retour, on pourra comprendre ce qui est affecté dans ces mutants EF-1A (arrivée moins rapide du complexe ternaire au site A, discrimination d'un complexe ternaire arrivant hors de phase,...). Des modèles nouveaux du rôle de EF-1A sur la fidélité de la traduction sont élaborés sur les bases des effets de mutants *TEF2* sur le recodage (Farabaugh *et al.* 1998). En cela le recodage peut aider à comprendre les mécanismes/effeteurs de l'appareil traductionnel dans des conditions de traduction normale. Pour l'heure, les données d'inter-relation sont trop fragmentaires pour pouvoir définir des règles simples, mais je pense que ces études ne peuvent qu'être informatives. De nombreux facteurs de la traduction restent encore à analyser dans le cadre du recodage.

Toutes ces connaissances qui augmentent notre compréhension du décalage de phase et de son contrôle traductionnel seront applicables à des stratégies anti-virales agronomiques ou thérapeutiques. En effet, aucun gène nécessaire à une fonction eucaryote et dépendant d'un décalage de phase en -1 n'a été découvert à ce jour. De plus, le recodage des virus s'appuyant sur l'appareil traductionnel de l'hôte, et non sur les fonctions codées par le génome viral, la probabilité de sélectionner des mutants viraux résistants est plus faible (Dinman *et al.* 1998). Des études ont déjà permis d'évaluer l'influence de composés anti-recodage sur la propagation *ex vivo* du VIH-1 (Hung *et al.* 1998).

Enfin, l'influence sur le recodage de facteurs généraux comme la température, les phases de croissance ouvrent des perspectives encore plus larges pour relier l'état physiologique général de la cellule au contrôle de la fidélité de la traduction.

Chez *S. cerevisiæ*, une cellule mère ne peut subir qu'un nombre limité de divisions, jusqu'à ce que l'accumulation de copies d'ADNr sous forme épisomique ne provoque la sénescence (Sinclair *et al.* 1997). L'augmentation du nombre de copies d'ADNr pourrait provoquer une modification de la fidélité de traduction. D'un autre côté, la longueur des télomères est contrôlée notamment par *EST3*, dont l'expression nécessite un décalage de phase en +1 (Morris *et al.* 1997). La fidélité de la traduction pourrait donc avoir une influence sur la longueur des télomères de la levure.

La longueur des télomères influe sur la sénescence en modulant la recombinaison de l'ADNr (Sinclair *et al.* 1997). Si les deux hypothèses précédentes sont vérifiées, la fidélité de la traduction chez *S. cerevisiæ* pourrait donc indirectement moduler la sénescence, dont l'une des conséquences serait une modification de la fidélité de la traduction. Ceci permet de faire l'hypothèse de régulations complexes à la croisée entre fidélité de la traduction et sénescence.



# Annexes

## Matériels et Méthodes

Voir les articles I et II

### *Escherichia coli*

Les souches, milieux et méthodes utilisées en biologie moléculaire ont été décrits (Sambrook *et al.* 1989).

### *Saccharomyces cerevisiae*

#### Techniques

Les techniques utilisées ont été décrites dans “Methods in Yeast Genetics, A laboratory course Manual”, C. Kaiser, A. Adams, D. Gottschling, 1995, Cold Spring Harbor Laboratory Press

#### **Tests $\beta$ -galactosidase sur boîte**

##### sur filtre

Les boîtes de Petri sont répliquées directement sur du papier *Whatman* (#1450082, grade 50). La réplique est immergée 30 secondes dans l’azote liquide, puis mise à sécher une minute dans une boîte de Petri vide. 450  $\mu$ l de Tampon Z additionné de X-gal 400  $\mu$ g/ml final sont ajoutés sous le filtre à la micropipette en évitant les bulles d’air. La coloration est révélée à température ambiante. Les colonies intéressantes sont prélevées directement sur le filtre et ré-isolées.

##### en Top-Agarose

La technique est inspirée de Duttweiler, 1996. Contrairement à ce qui est indiqué, les colonies sont rarement récupérables après cette révélation. Une réplique au préalable est nécessaire.

## Souches

| Nom        | génotype  | origine                                |
|------------|---|--|
| Fy1679-18B | Mat <sup>-</sup> , his3-200, leu2-1, trp1-63, ura3-52   | (Dujon <i>et al.</i> 1994)             |
| Y349       | Mat <sup>-</sup> , his3-200, leu2,3-112, lys2-201, ura3-52  | Van-Dinh Dong, thèse Paris 11, 1996    |
| YAL1       | Mat <sup>-</sup> , asn1-lacZ, his3-200, leu2,3-112, lys2-201, ura3-52   | Y349                                   |
| Y9L2       | Mat <sup>-</sup> , asn1-HIV-lacZ, his3-200, leu2,3-112, lys2-201, ura3-52   | Y349                                   |
| YONX7      | Mat <sup>-</sup> , asn1-TMV-lacZ, his3-200, leu2,3-112, lys2-201, ura3-52   | Y349                                   |
| ORT128     | MAT <sup>a</sup> , arg4 <sup>2060::URA3</sup> , ade2, his3 <sup>1</sup> , trp1-289, ura3-52                                       | MGD131-102a (Rocco <i>et al.</i> 1992) |
| YTMV       | MAT <sup>a</sup> , arg4 <sup>2060::pACTMV</sup> , ade2, his3 <sup>1</sup> , trp1-289, ura3-52                                     | ORT128                                 |
| YTX        | YTMV x YONX7  |  |
| YTX8b      | Mat <sup>a</sup> , arg4 <sup>2060::pACTMV</sup> , asn1-TMV-lacZ, ade2, trp1-289, leu2,3-112, lys2-201, ura3-52, his3 <sup>?</sup> | spore de YTX                           |

## Tableau récapitulatif des résultats de mutagenèse

| # | Taux de survie | Nombre colonies criblées | Nombre bleues foncées <sup>a</sup> | Nombre bleues claires <sup>a</sup> | [Gly <sup>-</sup> ] | [rose] <sup>b</sup> | mutant + ou - <sup>c</sup> |
|---|----------------|--------------------------|------------------------------------|------------------------------------|---------------------|---------------------|----------------------------|
| 0 | <i>nd</i>      | 1000                     | (12)                               |                                    | <i>nd</i>           | <i>nd</i>           | 0+/12<br>1-/12             |
|   |                |                          |                                    | (5)                                | <i>nd</i>           | <i>nd</i>           | 1-/5<br>2+/5               |
| 1 | 5%             | 4000                     | 143 (41)                           |                                    | 3/7                 | 3-/7                | 4+ /10<br>3- / 10          |
|   |                |                          |                                    | 50 (31)                            | 24/26               | 24-/26              | <i>nd</i>                  |
| 2 | 1%             | 20000                    | 50 (26)                            | <i>nd</i>                          | 0/26                | <i>nd</i>           | <i>nd</i>                  |
| 3 | 1%             | 2000                     | 9 ( <i>nd</i> )                    |                                    | 0/9                 | <i>nd</i>           | <i>nd</i>                  |
|   |                |                          |                                    | 4 (0)                              | -                   | -                   | -                          |
| 4 | 4%             | 9000                     | 96 (38)                            |                                    | 0/96                | 6++/3<br>8          | <i>nd</i>                  |
|   |                |                          |                                    | 67 (19)                            | 15/19               | 15-/19              | <i>nd</i>                  |
| 7 | 2,5%           | 4500                     | 24 ( <i>nd</i> )                   | 3 blancs                           | <i>nd</i>           | <i>nd</i>           | <i>nd</i>                  |
| 8 | 5%             | 1700                     | 34 ( <i>nd</i> )                   | <i>nd</i>                          | <i>nd</i>           | <i>nd</i>           | <i>nd</i>                  |
|   |                | 42200                    | 368 (175)                          | 129<br>(58)                        | <i>nd</i>           | <i>nd</i>           | <i>nd</i>                  |

en *italique*, les mutants perdus lors d'une infection acarienne (#0, 1, 2 et 3).

*nd* : non déterminé

<sup>a</sup> entre parenthèses, le nombre de colonies ayant passé le deuxième crible coloré avec succès.

<sup>b</sup> Le phénotype [rose] est un phénotype coloré que prennent les levures *ade2*; ici sont indiqués les phénotypes plus roses (++) ou blancs (-).

<sup>c</sup> sont considérés arbitrairement comme mutants affectés dans la translecture du TMV les clones ayant plus de 15% de variation d'activités luciférase/ -galactosidase par rapport au sauvage, en plus (+) ou en moins (-).



## Code génétique

|           |           |           |           |
|-----------|-----------|-----------|-----------|
| TTT phe F | TCT ser S | TAT tyr Y | TGT cys C |
| TTC phe F | TCC ser S | TAC tyr Y | TGC cys C |
| TTA leu L | TCA ser S | TAA OCH Z | TGA OPA Z |
| TTG leu L | TCG ser S | TAG AMB Z | TGG trp W |
| CTT leu L | CCT pro P | CAT his H | CGT arg R |
| CTC leu L | CCC pro P | CAC his H | CGC arg R |
| CTA leu L | CCA pro P | CAA gln Q | CGA arg R |
| CTG leu L | CCG pro P | CAG gln Q | CGG arg R |
| ATT ile I | ACT thr T | AAT asn N | AGT ser S |
| ATC ile I | ACC thr T | AAC asn N | AGC ser S |
| ATA ile I | ACA thr T | AAA lys K | AGA arg R |
| ATG met M | ACG thr T | AAG lys K | AGG arg R |
| GTT val V | GCT ala A | GAT asp D | GGT gly G |
| GTC val V | GCC ala A | GAC asp D | GGC gly G |
| GTA val V | GCA ala A | GAA glu E | GGA gly G |
| GTG val V | GCG ala A | GAG glu E | GGG gly G |

## Références bibliographiques

1 Adamski, F. M., J. F. Atkins & R. F. Gesteland (1996). "Ribosomal protein L9 interactions with 23 S rRNA: the use of a translational bypass assay to study the effect of amino acid substitutions." J Mol Biol **261**(3): 357-371.

2 Adamski, F. M., K. K. McCaughan, F. Jorgensen, C. G. Kurland & W. P. Tate (1994). "The concentration of polypeptide chain release factors 1 and 2 at different growth rates of Escherichia coli." J Mol Biol **238**(3): 302-308.

3 Agris, P. F., R. Guenther, P. C. Ingram, M. M. Basti, J. W. Stuart, E. Sochacka & A. Malkiewicz (1997). "Unconventional structure of tRNA(Lys)SUU anticodon explains tRNA's role in bacterial and mammalian ribosomal frameshifting and primer selection by HIV-1." RNA **3**(4): 420-428.

4 Arkov, A. L., S. V. Korolev & L. L. Kisselev (1995). "5' contexts of Escherichia coli and human termination codons are similar." Nucleic Acids Res **23**(22): 4712-4716.

5 Atkins, J. F. & R. F. Gesteland (1996). Regulatory Recoding. Translational Control. Cold Spring Harbor, Cold Spring Harbor Laboratory Press: 653-684.

6 Atkins, J. F., R. B. Weiss & R. F. Gesteland (1990). "Ribosome gymnastics-Degree of difficulty 9.5, style 10.0." Cell **62**: 413-423.

7 Atkins, J. F., R. B. Weiss, S. Thompson & R. F. Gesteland (1991). "Towards a genetic dissection of the basis of triplet decoding, and its natural subversion: programmed reading frame shifts and hops." Annu Rev Genet **25**: 201-228.

8 Balasundaram, D., J. D. Dinman, R. B. Wickner, C. W. Tabor & H. Tabor (1994). "Spermidine deficiency increases +1 ribosomal frameshifting efficiency and inhibits Ty1 retrotransposition in Saccharomyces cerevisiae." Proc Natl Acad Sci USA **91**(1): 172-176.

- 9 Barak, Z., J. Gallant, D. Lindsley, B. Kwieciszewski & D. Heidel (1996). "Enhanced ribosome frameshifting in stationary phase cells." J Mol Biol **263**(2): 140-148.
- 10 Baron, C., J. Heider & A. Bock (1990). "Mutagenesis of selC, the gene for the selenocysteine-inserting tRNA- species in E. coli: effects on in vivo function." Nucleic Acids Res **18**(23): 6761-6766.
- 10b Belcour, L., O. Begel & M. Picard (1991). "A site-specific deletion in mitochondrial DNA of *Podospira* is under the control of nuclear genes." Proc Natl Acad Sci USA **88**(9): 3579-3583.
- 11 Belcourt, M. F. & P. J. Farabaugh (1990). "Ribosomal frameshifting in the yeast transposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site." Cell **62**(2): 339-352.
- 12 Benard, L., N. Mathy, M. M. Grunberg, B. Ehresmann, C. Ehresmann & C. Portier (1998). "Identification in a pseudoknot of a U.G motif essential for the regulation of the expression of ribosomal protein S15." Proc Natl Acad Sci USA **95**(5): 2564-2567.
- 13 Berg, B. L., C. Baron & V. Stewart (1991). "Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. II. Evidence that a mRNA stem-loop structure is essential for decoding opal (UGA) as selenocysteine." J Biol Chem **266**(33): 22386-22391.
- 14 Berry, M. J. & P. R. Larsen (1993). "Recognition of UGA as a selenocysteine codon in eukaryotes: a review of recent progress." Biochem Soc Trans **21**(4): 827-832.
- 15 Berteaux, V., J. P. Rousset & M. Cassan (1991). "UAG readthrough is not increased *in vivo* by Moloney murine leukemia virus infection." Biochimie **73**: 1291-1293.
- 16 Bidou, L., G. Stahl, B. Grima, H. Liu, M. Cassan & J. P. Rousset (1997). "In vivo HIV-1 frameshifting efficiency is directly related to the stability of the stem-loop stimulatory signal." RNA **3**(10): 1153-1158.
- 17 Blinkowa, A. L. & J. R. Walker (1990). "Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III gamma subunit from within the tau subunit reading frame." Nucleic Acids Res **18**(7): 1725-1729.

- 18 Bonetti, B., L. W. Fu, J. Moon & D. M. Bedwell (1995). "The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in *Saccharomyces cerevisiae*." J Mol Biol **251**(3): 334-345.
- 19 Bonneaud, N., O. Ozier-Kalogeropoulos, G. Li, M. Labouesse, L. Minvielle-Sebastia & F. Lacroute (1991). "A family of low and high copy replicative, integrative and single-stranded *S. cerevisiae*/*E. coli* shuttle vectors." Yeast **7**: 609-615.
- 20 Brault, V. & W. A. Miller (1992). "Translational frameshifting mediated by a viral sequence in plant cells." Proc Natl Acad Sci USA **89**(6): 2262-2266.
- 21 Brierley, I. (1995). "Ribosomal frameshifting on viral RNAs." J Gen Virol **76**( Part 8): 1885-1892.
- 22 Brierley, I., M. E. G. Boursnell, M. M. Blinns, B. Bilimoria, V. C. Blok, T. D. K. Brown & S. C. Inglis (1987). "An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV." Embo J **6**(12): 3779-3785.
- 23 Brierley, I., P. Digard & S. C. Inglis (1989). "Characterization of an efficient coronavirus ribosomal frameshift signal: Requirement for an RNA pseudoknot." Cell **57**: 537-547.
- 24 Brierley, I., A. J. Jenner & S. C. Inglis (1992). "Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshift signal." J Mol Biol **227**: 463-479.
- 25 Brierley, I., M. R. Meredith, A. J. Bloys & T. G. Hagervall (1997). "Expression of a coronavirus ribosomal frameshift signal in *Escherichia coli*: influence of tRNA anticodon modification on frameshifting." J Mol Biol **270**(3): 360-373.
- 26 Brierley, I., N. J. Rolley, A. J. Jenner & S. C. Inglis (1991). "Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal." J Mol Biol **220**(4): 889-902.
- 27 Brown, C. M., P. A. Stockwell, C. N. Trotman & W. P. Tate (1990a). "The signal for the termination of protein synthesis in procaryotes." Nucleic Acids Res **18**(8): 2079-2086.

- 28 Brown, C. M., P. A. Stockwell, C. N. Trotman & W. P. Tate (1990b). "Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes." Nucleic Acids Res **18**(21): 6339-6345.
- 29 Calderone, T. L., R. D. Stevens & T. G. Oas (1996). "High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in Escherichia coli." J Mol Biol **262**(4): 407-412.
- 30 Camonis, J. H., M. Cassan & J. P. Rousset (1990). "Of mice and yeast: versatile vectors which permit expression in both budding yeast and higher eukaryotic cells." Gene **86**: 263-268.
- 31 Cassan, M., V. Berteaux, P. O. Angrand & J. P. Rousset (1990). "Expression vectors for quantitating *in vivo* translational ambiguity: their potential use to analyse frameshifting at the HIV gag-pol junction." Res Virol **141**: 597-610.
- 32 Cassan, M., N. Delaunay, C. Vaquero & J. P. Rousset (1994). "Translational frameshifting at the gag-pol junction of human immunodeficiency virus type 1 is not increased in infected T-lymphoid cells." J Virol **68**(3): 1501-1508.
- 33 Chamoro, M., N. Parkin & H. E. Varmus (1992). "An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA." Proc Natl Acad Sci USA **89**: 713-717.
- 34 Chen, X., M. Chamorro, S. I. Lee, L. X. Shen, J. V. Hines, I. J. Tinoco & H. E. Varmus (1995). "Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting." Embo J **14**(4): 842-852.
- 35 Chen, X., H. Kang, L. X. Shen, M. Chamorro, H. E. Varmus & I. J. Tinoco (1996). "A characteristic bent conformation of RNA pseudoknots promotes -1 frameshifting during translation of retroviral RNA." J Mol Biol **260**(4): 479-483.
- 36 Clare, J. J., M. Belcourt & P. J. Farabaugh (1988). "Efficient translational frameshifting occurs within a conserved sequence of the overlap between the two genes of a yeast Ty1 transposon." Proc Natl Acad Sci USA **85**: 6816-6820.

- 36b Coppin-Raynal, E., M. Dequard-Chablat & M. Picard (1988). Genetics of Ribosomes and Translation Accuracy in *Podospira anserina*. Genetics of translation. M. F. Tuite, M. Picard and M. Bolotin-Fukuhara. Springer-Verlag, Berlin Heidelberg, NATO ASI Series **H14**.
- 37 Cottrelle, P., D. Thiele, V. L. Price, S. Memet, J. Y. Micouin, C. Marck, J. M. Buhler, A. Sentenac & P. Fromageot (1985). "Cloning, nucleotide sequence, and expression of one of two genes coding for yeast elongation factor 1 alpha." J Biol Chem **260**(5): 3090-3096.
- 38 Craigen, W. J. & C. T. Caskey (1986). "Expression of peptide chain release factor 2 requires high-efficiency frameshifting." Nature **322**: 273-275.
- 39 Craigen, W. J., R. G. Cook, W. P. Tate & C. T. Caskey (1985). "Bacterial peptide chain release factors: conserved primary structure and possible frameshift regulation of release factor 2." Proc Natl Acad Sci USA **82**(11): 3616-20.
- 40 Cui, Y., J. D. Dinman, T. G. Kinzy & S. W. Peltz (1998). "The Mof2/Sui1 protein is a general monitor of translational accuracy." Mol Cell Biol **18**(3): 1506-1516.
- 41 Cui, Y., J. D. Dinman & S. W. Peltz (1996). "Mof4-1 is an allele of the UPF1/IFS2 gene which affects both mRNA turnover and -1 ribosomal frameshifting efficiency." Embo J **15**(20): 5726-5736.
- 42 Culbertson, M. R., R. F. Gaber & C. M. Cummins (1982). "Frameshift suppression in *Saccharomyces cerevisiae*. V. Isolation and genetic properties of nongroup-specific suppressors." Genetics **102**(3): 361-78.
- 43 Culbertson, M. R., K. M. Underbrink & G. R. Fink (1980). "Frameshift suppression *Saccharomyces cerevisiae*. II. Genetic properties of group II suppressors." Genetics **95**(4): 833-53.
- 44 Curran, J. F. (1993). "Analysis of effects of tRNA:message stability on frameshift frequency at the *Escherichia coli* RF2 programmed frameshift site." Nucleic Acids Res **21**(8): 1837-1843.
- 45 Curran, J. F. & M. Yarus (1989). "Rates of aminoacyl-tRNA selection at 29 sense codons in vivo." J Mol Biol **209**(1): 65-77.
- 46 Dinman, J. D. (1995). "Ribosomal frameshifting in yeast viruses." Yeast **11**(12): 1115-1127.

- 47 Dinman, J. D., T. Icho & R. B. Wickner (1991). "A -1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein." Proc Natl Acad Sci USA **88**: 174-178.
- 48 Dinman, J. D. & T. G. Kinzy (1997a). "Translational misreading: mutations in translation elongation factor 1alpha differentially affect programmed ribosomal frameshifting and drug sensitivity." RNA **3**(8): 870-881.
- 49 Dinman, J. D., E. M. Ruiz & S. W. Peltz (1998). "Translating old drugs into new treatments: ribosomal frameshifting as a target for antiviral agents." Trends Biotechnol **16**(4): 190-196.
- 50 Dinman, J. D., M. J. Ruiz-Echevarria, K. Czaplinski & S. W. Peltz (1997b). "Peptidyl-transferase inhibitors have antiviral properties by altering programmed -1 ribosomal frameshifting efficiencies: development of model systems." Proc Natl Acad Sci USA **94**(13): 6606-6611.
- 51 Dinman, J. D. & R. B. Wickner (1992). "Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation." J Virol **66**(6): 3669-3676.
- 52 Dinman, J. D. & R. B. Wickner (1994). "Translational maintenance of frame: mutants of *Saccharomyces cerevisiae* with altered -1 ribosomal frameshifting efficiency." Genetics **136**: 75-86.
- 53 Dinman, J. D. & R. B. Wickner (1995). "5 S rRNA is involved in fidelity of translational reading frame." Genetics **141**(1): 95-105.
- 54 Dong, H., L. Nilsson & C. G. Kurland (1996). "Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates." J Mol Biol **260**(5): 649-663.
- 55 Donly, B. C., C. D. Edgar, F. M. Adamski & W. P. Tate (1990). "Frameshift autoregulation in the gene for *Escherichia coli* release factor 2: partly functional mutants result in frameshift enhancement." Nucleic Acids Res **18**(22): 6517-22.
- 56 Dujon, B. *et al.* (1994). "Complete DNA sequence of yeast chromosome XI." Nature **369**: 371-377.
- 57 Duttweiler, H. M. (1996). "A highly sensitive and non-lethal beta-galactosidase plate assay for yeast." Trends Genet **12**(9): 340-341.

58 Erbs, P. (1996). L'aspartyl-tRNA synthétase de *Saccharomyces cerevisiae*: sélection *in vivo* de mutations létales, complémentation intragénique hétérodimérique.

Allosuppresseurs dans le système de codage du résidu aspartyl. UFR des Sciences de la Vie et de la Terre. Strasbourg, Louis Pasteur: 235.

59 Farabaugh, P. J. (1993). "Alternative readings of the genetic code." Cell **74**(4): 591-596.

60 Farabaugh, P. J. (1995). "Post-transcriptional regulation of transposition by Ty retrotransposons of *Saccharomyces cerevisiae*." J Biol Chem **270**(18): 10361-10364.

61 Farabaugh, P. J. (1996). "Programmed translational frameshifting." Microbiol Rev **60**(1): 103-134.

62 Farabaugh, P. J. & A. Vimaladithan (1998). "Effect of frameshift-inducing mutants of elongation factor 1alpha on programmed +1 frameshifting in yeast." RNA **4**(1): 38-46.

63 Farabaugh, P. J., H. Zhao & A. Vimaladithan (1993). "A novel programmed frameshift expresses the POL3 gene of Retrotransposon-Ty3 of yeast - frameshifting without transfer RNA slippage." Cell **74**(1): 93-103.

64 Felsenstein, K. M. & S. P. Goff (1988). "Expression of the gag-pol fusion protein of Moloney murine leukemia virus without gag protein does not induce virion formation or proteolytic processing." J Virol **62**(6): 2179-2182.

65 Felsenstein, K. M. & S. P. Goff (1992). "Mutational analysis of the gag-pol junction of Moloney murine leukemia virus: requirements for expression of the gag-pol fusion protein." J Virol **66**(11): 6601-6608.

66 Feng, Y. X., J. G. Levin, D. L. Hatfield, T. S. Schaefer, R. J. Gorelick & A. Rein (1989). "Suppression of UAA and UGA termination codons in mutant murine leukemia viruses." J Virol **63**(6): 2870-2873.

67 Feng, Y. X., H. Yuan, A. Rein & J. G. Levin (1992). "Bipartite signal for read-through suppression in murine leukemia virus mRNA: an eight-nucleotide purine-rich sequence immediately downstream of the gag termination codon followed by an RNA pseudoknot." J Virol **66**(8): 5127-5132.

- 68 Fichant, G. A. & Y. Quentin (1995). "A frameshift error detection algorithm for DNA sequencing projects." Nucleic Acids Res **23**(15): 2900-2908.
- 69 Forchhammer, K. & A. Bock (1991). "Selenocysteine synthase from Escherichia coli. Analysis of the reaction sequence." J Biol Chem **266**(10): 6324-6328.
- 70 Fu, C. & J. Parker (1994). "A ribosomal frameshifting error during translation of the argI mRNA of Escherichia coli." Mol Gen Genet **243**(4): 434-441.
- 71 Gallant, J., C. Kurland, J. Parker, R. Holliday & R. Rosenberger (1997). "The error catastrophe theory of aging. Point counterpoint." Exp Gerontol **32**(3): 333-346.
- 72 Gao, W., H. Jakubowski & E. Goldman (1995). "Evidence that uncharged tRNA can inhibit a programmed translational frameshift in Escherichia coli." J Mol Biol **251**(2): 210-216.
- 73 Garcia, A., D. J. van & C. W. Pleij (1993). "Differential response to frameshift signals in eukaryotic and prokaryotic translational systems." Nucleic Acids Res **21**(3): 401-406.
- 74 Geller, A. I. & A. Rich (1980). "A UGA termination suppression tRNA<sup>Trp</sup> active in rabbit reticulocytes." Nature **283**(5742): 41-46.
- 75 Gesteland, R. F. & J. F. Atkins (1996). "Recoding: dynamic reprogramming of translation." Annu Rev Biochem **65**(741): 741-768.
- 76 Gesteland, R. F., R. B. Weiss & J. F. Atkins (1992). "Recoding: Reprogrammed genetic decoding." Science **257**: 1640-1643.
- 76b Goodsell, D. S. (1991). "Inside a living cell." Trends Biochem Sci **16**(6): 203-206.
- 77 Grant, C. M., M. Firoozan & M. F. Tuite (1989). "Mistranslation induces the heat-shock response in the yeast Saccharomyces cerevisiae." Mol Microbiol **3**(2): 215-20.
- 78 Grentzmann, G., J. A. Ingram, P. J. Kelly, R. F. Gesteland & J. F. Atkins (1998). "A dual-luciferase reporter system for studying recoding signals." Rna **4**(4): 479-486.

79 Grentzmann, G., D. Brechemierbaey, V. Heurguehamard & R. H. Buckingham (1995). "Function of polypeptide chain release factor RF-3 in *Escherichia coli* - RF-3 action in termination is predominantly at UGA-containing stop signals." J Biol Chem **270**(18): 10595-10600.

80 Hatfield, D., Y. X. Feng, B. J. Lee, A. Rein, J. G. Levin & S. Oroszlan (1989). "Chromatographic analysis of the aminoacyl-tRNAs which are required for translation of codons at and around the ribosomal frameshift sites of HIV, HTLV-1 and BLV." Virology **173**: 736-742.

81 Hatfield, D. & S. Oroszlan (1990). "The where, what and how of ribosomal frameshifting in retroviral protein synthesis." Trends Biochem Sci **15**(5): 186-190.

82 Hatfield, D. L., J. G. Levin, A. Rein & S. Oroszlan (1992). "Translational suppression in retroviral gene expression." Adv Virus Res **41**(193): 193-239.

83 Heider, J., C. Baron & A. Bock (1992). "Coding from a distance: dissection of the mRNA determinants required for the incorporation of selenocysteine into protein." Embo J **11**(10): 3759-3766.

84 Herbst, K. L., L. M. Nichols, R. F. Gesteland & R. B. Weiss (1994). "A mutation in ribosomal protein L9 affects ribosomal hopping during translation of gene 60 from bacteriophage T4." Proc Natl Acad Sci USA **91**(26): 12525-12529.

85 Hill, K. E., R. S. Lloyd & R. F. Burk (1993). "Conserved nucleotide sequences in the open reading frame and 3' untranslated region of selenoprotein P mRNA." Proc Natl Acad Sci USA **90**(2): 537-541.

86 Hinnebusch, A. G. & S. W. Liebman (1991). Protein synthesis and translational control in *Saccharomyces cerevisiae*. The molecular and cellular biology of the yeast Saccharomyces. C. S. H. L. Press. Cold Spring Harbor. **1**: 627-735.

87 Honigman, A., H. Falk, N. Mador, T. Rosental & A. Panet (1995). "Translation efficiency of the human T-cell leukemia virus (HTLV-2) gag gene modulates the frequency of ribosomal frameshifting." Virology **208**(1): 312-318.

- 88 Hopfield, J. J. (1974). "Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity." Proc Natl Acad Sci USA **71**(10): 4135-4139.
- 89 Horsfield, J. A., D. N. Wilson, S. A. Mannering, F. M. Adamski & W. P. Tate (1995). "Prokaryotic ribosomes recode the HIV-1 gag-pol-1 frameshift sequence by an E/P site post-translocation simultaneous slippage mechanism." Nucleic Acids Res **23**(9): 1487-1494.
- 90 Hughes, D., J. F. Atkins & S. Thompson (1987). "Mutants of elongation factor Tu promote ribosomal frameshifting and nonsense readthrough." Embo J **6**(13): 4235-4239.
- 91 Hung, M., P. Patel, S. Davis & S. R. Green (1998). "Importance of ribosomal frameshifting for human immunodeficiency virus type 1 particle assembly and replication." J Virol **72**(6): 4819-4824.
- 92 Inge Vechtomov, S. G., L. N. Mironova & M. D. Ter Avanesian (1994). "Ambiguity of translation: a eukaryotic version?" Genetika **30**(8): 1022-1035.
- 93 Ishikawa, M., T. Meshi, F. Motoyoshi, N. Takamatsu & Y. Okada (1986). "In vitro mutagenesis of the putative replicase genes of tobacco mosaic virus." Nucleic Acids Res **14**(21): 8291-8305.
- 94 Jacks, T., H. D. Madhani, H. D. Masiarz & H. E. Varmus (1988a). "Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region." Cell **55**: 447-458.
- 95 Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr & H. E. Varmus (1988b). "Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression." Nature **331**: 280-283.
- 96 Jacks, T., K. Townsley, H. E. Varmus & J. Majors (1987). "Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins." Proc Natl Acad Sci USA **84**(12): 4298-4302.
- 97 Jacks, T. & H. E. Varmus (1985). "Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting." Science **230**: 1237-1242.

- 98 Jalajakumari, M. B., C. J. Thomas, R. Halter & P. A. Manning (1989). "Genes for biosynthesis and assembly of CS3 pili of CFA/II enterotoxigenic *Escherichia coli*: novel regulation of pilus production by bypassing an amber codon." *Mol Microbiol* **3**(12): 1685-1695.
- 98b Janosi, L., S. Mottagui-Tabar, L. A. Isaksson, Y. Sekine, E. Ohtsubo, S. Zhang, S. Goon, S. Nelken, M. Shuda & A. Kaji (1998). "Evidence for in vivo ribosome recycling, the fourth step in protein biosynthesis." *Embo J* **17**(4): 1141-1151.
- 99 Jorgensen, F., F. M. Adamski, W. P. Tate & C. G. Kurland (1993). "Release factor-dependent false stops are infrequent in *Escherichia coli*." *J Mol Biol* **230**(1): 41-50.
- 100 Kahns, S., A. Lund, P. Kristensen, C. R. Knudsen, B. Clark, J. Cavallius & W. C. Merrick (1998). "The elongation factor 1 A-2 isoform from rabbit: cloning of the cDNA and characterization of the protein." *Nucleic Acids Res* **26**(8): 1884-1890.
- 101 Kang, H. (1998). "Direct structural evidence for formation of a stem-loop structure involved in ribosomal frameshifting in human immunodeficiency virus type 1." *Biochim Biophys Acta* **1397**(1): 73-78.
- 102 Kang, H. & I. J. Tinoco (1997). "A mutant RNA pseudoknot that promotes ribosomal frameshifting in mouse mammary tumor virus." *Nucleic Acids Res* **25**(10): 1943-1949.
- 103 Kawakami, K., S. Pande, B. Faiola, D. P. Moore, J. D. Boeke, P. J. Farabaugh, J. N. Strathern, Y. Nakamura & D. J. Garfinkel (1993). "A rare transfer RNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for ty1 retrotransposition in *saccharomyces cerevisiae*." *Genetics* **135**(2): 309-320.
- 104 Kollmus, H., M. W. Hentze & H. Hauser (1996). "Regulated ribosomal frameshifting by an RNA-protein interaction." *RNA* **2**(4): 316-323.
- 105 Kollmus, H., A. Honigman, A. Panet & H. Hauser (1994). "The sequences of and distance between two cis-acting signals determine the efficiency of ribosomal frameshifting in human immunodeficiency virus type 1 and human T-cell leukemia virus type II in vivo." *J Virol* **68**(9): 6087-6091.

- 106 Kolter, R., D. A. Siegele & A. Tormo (1993). "The stationary phase of the bacterial life cycle." Annu Rev Microbiol **47**(855): 855-874.
- 107 Kurland, C. G. (1992). "Translational accuracy and the fitness of bacteria." Annu Rev Genet **26**(29): 29-50.
- 108 Larsen, B., R. F. Gesteland & J. F. Atkins (1997). "Structural probing and mutagenic analysis of the stem-loop required for Escherichia coli dnaX ribosomal frameshifting: programmed efficiency of 50%." J Mol Biol **271**(1): 47-60.
- 109 Larsen, B., N. M. Wills, R. F. Gesteland & J. F. Atkins (1994). "rRNA-mRNA base pairing stimulates a programmed -1 ribosomal frameshift." J Bacteriol **176**(22): 6842-6851.
- 110 Lee, B. S. & M. R. Culbertson (1995a). "Identification of an additional gene required for eukaryotic nonsense mRNA turnover." Proc Natl Acad Sci USA **92**(22): 10354-10358.
- 111 Lee, S. I., J. G. Umen & H. E. Varmus (1995b). "A genetic screen identifies cellular factors involved in retroviral -1 frameshifting." Proc Natl Acad Sci USA **92**(14): 6587-6591.
- 112 Leeds, P., S. W. Peltz, A. Jacobson & M. R. Culbertson (1991). "The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon." Genes Dev: 2303-14.
- 113 Leeds, P., J. M. Wood, B. S. Lee & M. R. Culbertson (1992). "Gene products that promote mRNA turnover in Saccharomyces cerevisiae." Mol Cell Biol **12**(5): 2165-77.
- 114 Li, G. & C. M. Rice (1993). "The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon." J Virol **67**(8): 5062-7.
- 115 Liebman, S. W., Y. O. Chernoff & R. Liu (1995). "The accuracy center of a eukaryotic ribosome." Biochem Cell Biol **73**(11-12): 1141-1149.
- 116 Lustig, A. J. (1998). "DNA dynamics: different means to a common end?" Curr Biol **8**(5): R161-R164.

- 117 Major, L. L., E. S. Poole, M. E. Dalphin, S. A. Mannering & W. P. Tate (1996). "Is the in-frame termination signal of the Escherichia coli release factor-2 frameshift site weakened by a particularly poor context?" Nucleic Acids Res **24**(14): 2673-2678.
- 118 Manch-Citron, J. N. & J. London (1994). "Expression of the Prevotella loescheii adhesin gene (plaA) is mediated by a programmed frameshifting hop." J Bacteriol **176**(7): 1944-1948.
- 119 Maquat, L. E. (1995). "When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells." RNA **1**(5): 453-465.
- 120 Martin, G. W., J. W. Harney & M. J. Berry (1996). "Selenocysteine incorporation in eukaryotes: insights into mechanism and efficiency from sequence, structure, and spacing proximity studies of the type 1 deiodinase SECIS element." RNA **2**(2): 171-182.
- 121 Matsufuji, S., T. Matsufuji, Y. Miyazaki, Y. Murakami, J. F. Atkins, R. F. Gesteland & S. Hayashi (1995). "Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme." Cell **80**(1): 51-60.
- 122 Matsufuji, S., T. Matsufuji, N. M. Wills, R. F. Gesteland & J. F. Atkins (1996). "Reading two bases twice: mammalian antizyme frameshifting in yeast." Embo J **15**(6): 1360-1370.
- 123 Matsushita, O., J. B. Russell & D. B. Wilson (1991). "A Bacteroides ruminicola 1,4-beta-D-endoglucanase is encoded in two reading frames." J Bacteriol **173**(21): 6919-6926.
- 124 McCaughan, K. K., C. M. Brown, M. E. Dalphin, M. J. Berry & W. P. Tate (1995). "Translational termination efficiency in mammals is influenced by the base following the stop codon." Proc Natl Acad Sci USA **92**(12): 5431-5435.
- 125 Miller, W. A., S. P. Dinesh-Kumar & C. P. Paul (1995). "Luteovirus Gene Expression." Crit Rev Plant Sci **14**(3): 179-211.
- 126 Morikawa, S. & D. H. L. Bishop (1992). "Identification and analysis of the gag-pol ribosomal frameshift site of feline immunodeficiency virus." Virology **186**: 389-397.

- 127 Morris, D. K. & V. Lundblad (1997). "Programmed translational frameshifting in a gene required for yeast telomere replication." Curr Biol **7**(12): 969-976.
- 128 Mottagui Tabar, S. & L. A. Isaksson (1997). "Only the last amino acids in the nascent peptide influence translation termination in Escherichia coli genes." Febs Lett **414**(1): 165-170.
- 129 Nagata, S., K. Nagashima, Y. Y. Tsunetsugu, K. Fujimura, M. Miyazaki & Y. Kaziro (1984). "Polypeptide chain elongation factor 1 alpha (EF-1 alpha) from yeast: nucleotide sequence of one of the two genes for EF-1 alpha from Saccharomyces cerevisiae." Embo J **3**(8): 1825-1830.
- 130 Nelson, R. J., T. Ziegelhoffer, C. Nicolet, M. Werner Washburne & E. A. Craig (1992). "The translation machinery and 70 kd heat shock protein cooperate in protein synthesis." Cell **71**(1): 97-105.
- 131 Ninio, J. (1974). "A semi-quantitative treatment of missense and nonsense suppression in the strA and ram ribosomal mutants of Escherichia coli. Evaluation of some molecular parameters of translation in vivo." J Mol Biol **84**(2): 297-313.
- 132 Ninio, J. (1975). "Kinetic amplification of enzyme discrimination." Biochimie **57**(5): 587-595.
- 133 Ninio, J. (1986). "Fine tuning of ribosomal accuracy." Febs Lett **196**(1): 1-4.
- 133b Orgel, L. E. (1970). "The maintenance of the accuracy of protein synthesis and its relevance to ageing: a correction." Proc Natl Acad Sci USA **67**(3): 1476.
- 134 Pande, S., A. Vimaladithan, H. Zhao & P. J. Farabaugh (1995). "Pulling the ribosome out of frame by +1 at a programmed frameshift site by cognate binding of aminoacyl-tRNA." Mol Cell Biol **15**(1): 298-304.
- 135 Panganiban, A. T. (1988). "Retroviral gag amber codon suppression is caused by an intrinsic cis-acting component of the viral mRNA." J Virol **62**: 3574-3580.
- 136 Paraskeva, E., A. Atzberger & M. W. Hentze (1998). "A translational repression assay procedure (TRAP) for RNA-protein interactions in vivo." Proc Natl Acad Sci USA **95**(3): 951-956.

- 137 Park, J. & C. D. Morrow (1991). "Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production." J Virol **65**(9): 5111-5117.
- 138 Parker, J. (1989). "Errors and alternatives in reading the universal genetic code." Microbiol Rev **53**(3): 273-298.
- 139 Parker, J., J. Flanagan, J. Murphy & J. Gallant (1981). "On the accuracy of protein synthesis in *Drosophila melanogaster*." Mech Ageing Dev **16**(2): 127-39.
- 140 Parkin, N. T., M. Chamorro & H. E. Varmus (1992). "Human immunodeficiency virus type 1 *gag-pol* frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression *in vivo*." J Virol **66**(8): 5147-5151.
- 141 Pedersen, W. T. & J. F. Curran (1991). "Effects of the nucleotide 3' to an amber codon on ribosomal selection rates of suppressor tRNA and release factor-1." J Mol Biol **219**(2): 231-241.
- 142 Pelham, H. R. B. (1978). "Leaky UAG termination codon in tobacco mosaic virus RNA." Nature **272**: 469-471.
- 143 Pelham, H. R. B. (1979). "Translation of Tobacco Rattle Virus RNAs *in vitro* : four proteins from three RNAs." Virology **97**: 256-265.
- 144 Peltz, S. W., F. He, E. Welch & A. Jacobson (1994). "Nonsense-mediated mRNA decay in yeast." Prog Nucleic Acid Res Mol Biol: 271-98.
- 145 Pocklington, M. J., L. Johnston, J. R. Jenkins & E. Orr (1990). "The omnipotent suppressor SUP45 affects nucleic acid metabolism and mitochondrial structure." Yeast **6**(5): 441-450.
- 146 Poole, E. S., C. M. Brown & W. P. Tate (1995). "The identity of the base following the stop codon determines the efficiency of *in vivo* translational termination in *Escherichia coli*." Embo J **14**(1): 151-158.
- 147 Reil, H. & H. Hauser (1990). "Test system for determination of HIV-1 frameshifting efficiency in animal cells." Biochim Biophys Acta **1050**: 288-292.

- 148 Reil, H., M. Hoxter, D. Moosmayer, G. Pauli & H. Hauser (1994). "CD4 expressing human 293 cells as a tool for studies in HIV-1 replication: the efficiency of translational frameshifting is not altered by HIV-1 infection." Virology **205**(1): 371-375.
- 149 Ribas, J. C. & R. B. Wickner (1998). "The Gag domain of the Gag-Pol fusion protein directs incorporation into the L-A double-stranded RNA viral particles in *Saccharomyces cerevisiae*." J Biol Chem **273**(15): 9306-9311.
- 150 Robinson, D. N. & L. Cooley (1997). "Examination of the function of two kelch proteins generated by stop codon suppression." Development **124**(7): 1405-1417.
- 151 Rocco, V., B. de Massy & A. Nicolas (1992). "The *Saccharomyces cerevisiae* ARG4 initiator of meiotic gene conversion and its associated double-strand DNA breaks can be inhibited by transcriptional interference." Proc Natl Acad Sci USA **89**(24): 12068-12072.
- 152 Rom, E. & C. Kahana (1994). "Polyamines regulate the expression of ornithine decarboxylase antizyme in vitro by inducing ribosomal frameshifting." Proc Natl Acad Sci USA **91**(9): 3959-3963.
- 153 Sambrook, J., E. F. Fritsch & T. Maniatis (1989). Molecular cloning. A laboratory manual. New York, Cold Spring Harbor Laboratory Press.
- 154 Sandbaken, M. G. & M. R. Culbertson (1988). "Mutations in elongation factor EF-1 alpha affect the frequency of frameshifting and amino acid misincorporation in *Saccharomyces cerevisiae*." Genetics **120**(4): 923-934.
- 155 Schirmaier, F. & P. Philippsen (1984). "Identification of two genes coding for the translation elongation factor EF-1 alpha of *S. cerevisiae*." Embo J **3**(13): 3311-3315.
- 156 Schwartz, R. & J. F. Curran (1997). "Analyses of frameshifting at UUU-pyrimidine sites." Nucleic Acids Res **25**(10): 2005-2011.
- 157 Sekine, Y., H. Nagasawa & E. Ohtsubo (1992a). "Identification of the site of translational frameshifting required for production of the transposase encoded by insertion sequence IS 1." Mol Gen Genet **235**(2-3): 317-324.

- 158 Sekine, Y. & E. Ohtsubo (1992b). "DNA sequences required for translational frameshifting in production of the transposase encoded by IS1." Mol Gen Genet **235**(2-3): 325-332.
- 159 Shen, L. X. & I. Tinoco (1995). "The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus." J Mol Biol **247**(5): 963-978.
- 160 Silar, P. & M. Picard (1994). "Increased longevity of EF-1 alpha high-fidelity mutants in *Podospora anserina*." J Mol Biol **235**(1): 231-6.
- 161 Sinclair, D. A. & L. Guarente (1997). "Extrachromosomal rDNA circles--a cause of aging in yeast." Cell **91**(7): 1033-1042.
- 162 Singh, A. (1977). "Nonsense suppressors of yeast cause osmotic-sensitive growth." Proc Natl Acad Sci USA **74**(1): 305-9.
- 163 Siple, J., J. Dunn & E. Goldman (1991). "Bacteriophage T7 morphogenesis and gene 10 frameshifting in *Escherichia coli* showing different degrees of ribosomal fidelity." Mol Gen Genet **230**(3): 376-384.
- 164 Siple, J. & E. Goldman (1993). "Increased ribosomal accuracy increases a programmed translational frameshift in *Escherichia coli*." Proc Natl Acad Sci USA **90**(6): 2315-2319.
- 165 Skuzeski, J. M., L. M. Nichols & R. F. Gesteland (1990). "Analysis of leaky viral translation termination codons in vivo by transient expression of improved beta-glucuronidase vectors." Plant Mol Biol **15**(1): 65-79.
- 166 Skuzeski, J. M., L. M. Nichols, R. F. Gesteland & J. F. Atkins (1991). "The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons." J Mol Biol **218**(2): 365-73.
- 167 Somogyi, P., A. J. Jenner, I. Brierley & S. C. Inglis (1993). "Ribosomal pausing during translation of an RNA pseudoknot." Mol Cell Biol **13**(11): 6931-6940.
- 168 Song, J. M., S. Picologlou, C. M. Grant, M. Firoozan, M. F. Tuite & S. W. Liebman (1989). "Elongation factor EF-1 alpha gene dosage alters translational fidelity in *Saccharomyces cerevisiae*." Mol Cell Biol **9**(10): 4571-5.

- 169 Stahl, G., L. Bidou, J. P. Rousset & M. Cassan (1995). "Versatile vectors to study recoding: conservation of rules between yeast and mammalian cells." Nucleic Acids Res **23**(9): 1557-1560.
- 170 Stapulionis, R., S. Kolli & M. P. Deutscher (1997). "Efficient mammalian protein synthesis requires an intact F-actin system." J Biol Chem **272**(40): 24980-24986.
- 171 Steneberg, P., C. Englund, J. Kronhamn, T. A. Weaver & C. Samakovlis (1998). "Translational readthrough in the hdc mRNA generates a novel branching inhibitor in the drosophila trachea." Genes Dev **12**(7): 956-967.
- 172 Sung, D. & H. Kang (1998). "Mutational analysis of the RNA pseudoknot involved in efficient ribosomal frameshifting in simian retrovirus-1." Nucleic Acids Res **26**(6): 1369-1372.
- 173 Tabor, C. W. & H. Tabor (1984). "Polyamines." Annu Rev Biochem **53**(749): 749-790.
- 174 Tang, C. K. & D. E. Draper (1989). "Unusual mRNA pseudoknot structure is recognized by a protein translational repressor." Cell **57**(4): 531-536.
- 175 Tate, W. P. & C. M. Brown (1992). "Translational termination: "stop" for protein synthesis or "pause" for regulation of gene expression." Biochemistry **31**(9): 2443-50.
- 176 Tate, W. P., E. S. Poole & S. A. Mannering (1996). "Hidden infidelities of the translational stop signal." Prog Nucleic Acid Res Mol Biol **52**(293): 293-335.
- 177 ten Dam, E., C. W. Pleij & D. Draper (1992). "Structural and functional aspects of RNA pseudoknots." Biochemistry **31**(47): 11665-11676.
- 178 ten Dam, E. B., I. Brierley, S. C. Inglis & C. W. Pleij (1994). "Identification and analysis of the pseudoknot-containing gag-pro ribosomal frameshift signal of simian retrovirus-1." Nucleic Acids Res **22**(12): 2304-2310.
- 179 ten Dam, E. B., C. W. Pleij & L. Bosch (1990). "RNA pseudoknots: translational frameshifting and readthrough on viral RNAs." Virus Genes **4**(2): 121-136.

- 180 ten Dam, E. B., P. W. Verlaan & C. W. Pleij (1995). "Analysis of the role of the pseudoknot component in the SRV-1 gag-pro ribosomal frameshift signal: loop lengths and stability of the stem regions." RNA **1**(2): 146-154.
- 181 Thompson, R. C. (1988). "EFTu provides an internal kinetic standard for translational accuracy." Trends Biochem Sci **13**(3): 91-93.
- 182 Tsuchihashi, Z. (1991). "Translational frameshifting in the Escherichia coli dnaX gene in vitro." Nucleic Acids Res **19**(9): 2457-2462.
- 183 Tsuchihashi, Z. & P. O. Brown (1992). "Sequence requirements for efficient translational frameshifting in the Escherichia coli dnaX gene and the role of an unstable interaction between tRNA(Lys) and an AAG lysine codon." Genes Dev **6**(3): 511-519.
- 184 Tsuchihashi, Z. & A. Kornberg (1990). "Translational frameshift generates the  $\beta$  subunit of DNA polymerase III holoenzyme." Proc Natl Acad Sci USA **87**: 2516-2520.
- 185 Tu, C., T. H. Tzeng & J. A. Bruenn (1992). "Ribosomal movement impeded at a pseudoknot required for frameshifting." Proc Natl Acad Sci USA **89**(18): 8636-8640.
- 186 Tumer, N. E., B. A. Parikh, P. Li & J. D. Dinman (1998). "The pokeweed antiviral protein specifically inhibits Ty1-directed +1 ribosomal frameshifting and retrotransposition in Saccharomyces cerevisiae." J Virol **72**(2): 1036-1042.
- 187 Valle, R. P., G. Drugeon, M. D. Devignes Morch, A. B. Legocki & A. L. Haenni (1992). "Codon context effect in virus translational readthrough. A study in vitro of the determinants of TMV and Mo-MuLV amber suppression." Febs Lett **306**(2-3): 133-139.
- 188 Vendeland, S. C., M. A. Beilstein, J. Y. Yeh, W. Ream & P. D. Whanger (1995). "Rat skeletal muscle selenoprotein W: cDNA clone and mRNA modulation by dietary selenium." Proc Natl Acad Sci USA **92**(19): 8749-8753.
- 189 Vickers, T. A. & D. J. Ecker (1992). "Enhancement of ribosomal frameshifting by oligonucleotides targeted to the HIV *gag-pol* region." Nucleic Acids Res **20**(15): 3945-3953.

- 190 Vimaladithan, A. & P. J. Farabaugh (1994). "Special peptidyl-tRNA molecules can promote translational frameshifting without slippage." Mol Cell Biol **14**(12): 8107-8116.
- 191 von Ahsen, U. (1998). "Translational fidelity: error-prone versus hyper-accurate ribosomes." Chem Biol **5**(1): R3-R6.
- 192 Wakem, L. P. & F. Sherman (1990). "Isolation and characterization of omnipotent suppressors in the yeast *Saccharomyces cerevisiae*." Genetics **124**(3): 515-22.
- 193 Wandelt, C. & G. Feix (1989). "Sequence of a 21 kd zein gene from maize containing an in-frame stop codon." Nucleic Acids Res **17**(6): 2354.
- 195 Weiss, R. B., D. M. Dunn, J. F. Atkins & R. F. Gesteland (1987a). "Slippery runs, shifty stops, backward steps, and forward hops: -2, -1, +1, +2, +5, and +6 ribosomal frameshifting." Cold Spring Harb Symp Quant Biol **52**(687): 687-693.
- 196 Weiss, R. B., D. M. Dunn, J. F. Atkins & R. F. Gesteland (1990a). "Ribosomal frameshifting from -2 to +50 nucleotides." Prog Nucleic Acid Res Mol Biol **39**(159): 159-183.
- 197 Weiss, R. B., D. M. Dunn, A. E. Dahlberg, J. F. Atkins & R. F. Gesteland (1988). "Reading frame switch caused by base-pair formation between the 3' end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli*." Embo J **7**(5): 1503-1507.
- 198 Weiss, R. B., D. M. Dunn, M. Shuh, J. F. Atkins & R. F. Gesteland (1989). "*E. coli* ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent." New Biol **1**: 159-169.
- 199 Weiss, R. B., W. M. Huang & D. M. Dunn (1990b). "A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60." Cell **62**(1): 117-126.
- 200 Weiss, W. A., I. Edelman, M. R. Culbertson & E. C. Friedberg (1987b). "Physiological levels of normal tRNA(CAGGln) can effect partial suppression of amber mutations in the yeast *Saccharomyces cerevisiae*." Proc Natl Acad Sci USA **84**(22): 8031-8034.

- 201 Weiss-brummer, B., A. Zollner, A. Haid & S. Thompson (1995). "Mutation of a highly conserved base in the yeast mitochondrial 21S rRNA restricts ribosomal frameshifting." Mol Gen Genet **248**(2): 207-216.
- 202 Weng, Y., K. Czaplinski & S. W. Peltz (1996a). "Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein." Mol Cell Biol **16**(10): 5477-5490.
- 203 Weng, Y., K. Czaplinski & S. W. Peltz (1996b). "Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover." Mol Cell Biol **16**(10): 5491-5506.
- 204 Wentzel, A. M., M. Stancek & L. A. Isaksson (1998). "Growth phase dependent stop codon readthrough and shift of translation reading frame in Escherichia coli." Febs Lett **421**(3): 237-242.
- 205 Williams, D. L., R. J. Pierce, E. Cookson & A. Capron (1992). "Molecular cloning and sequencing of glutathione peroxidase from Schistosoma mansoni." Mol Biochem Parasitol **52**(1): 127-130.
- 206 Wills, N. M., R. F. Gesteland & J. F. Atkins (1991). "Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus gag stop codon." Proc Natl Acad Sci USA **88**(16): 6991-6995.
- 207 Wills, N. M., R. F. Gesteland & J. F. Atkins (1994). "Pseudoknot-dependent read-through of retroviral gag termination codons: importance of sequences in the spacer and loop 2." Embo J **13**(17): 4137-4144.
- 208 Wilson, P. G. & M. R. Culbertson (1988a). "SUF12 suppressor protein of yeast. A fusion protein related to the EF-1 family of elongation factors." J Mol Biol **199**(4): 559-573.
- 209 Wilson, W., M. Braddock, S. E. Adams, P. D. Rathjen, S. M. Kingsman & A. J. Kingsman (1988b). "HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast system." Cell **55**: 1159-1169.
- 210 Wolin, S. L. & P. Walter (1988). "Ribosome pausing and stacking during translation of a eukaryotic mRNA." Embo J **7**(11): 3559-3569.

- 211 Wong, S. C. & A. T. Abdelal (1990). "Unorthodox expression of an enzyme: evidence for an untranslated region within carA from *Pseudomonas aeruginosa*." J Bacteriol **172**(2): 630-642.
- 212 Xu, H. & J. D. Boeke (1990). "Host genes that influence transposition in yeast: the abundance of a rare tRNA regulates Ty1 transposition frequency." Proc Nat Acad Sci USA **87**: 8360-8364.
- 213 Xue, F. & L. Cooley (1993). "kelch encodes a component of intercellular bridges in *Drosophila* egg chambers." Cell **72**(5): 681-693.
- 214 Yelverton, E., D. Lindsley, P. Yamauchi & J. A. Gallant (1994). "The function of a ribosomal frameshifting signal from human immunodeficiency virus-1 in *Escherichia coli*." Mol Microbiol **11**(2): 303-13.
- 215 Yoshinaka, Y., H. Katoh, T. D. Copeland & S. Oroszlan (1985). "Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon." Proc Natl Acad Sci USA **82**: 1618-1622.
- 216 Zerfass, K. & H. Beier (1992). "Pseudouridine in the anticodon G psi A of plant cytoplasmic tRNA(Tyr) is required for UAG and UAA suppression in the TMV-specific context." Nucleic Acids Res **20**(22): 5911-8.
- 217 Zinoni, F., J. Heider & A. Bock (1990). "Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine." Proc Natl Acad Sci USA **87**(12): 4660-4664.

# Résumé en anglais - Abstract

## Searching Recoding Modulating Factors In *Saccharomyces Cerevisiae*.

Recoding is a translational phenomenon used to express viral and cellular functions *via* ponctual and programmed modification of canonical translational rules. Recoding is needed for example for the expression of retroviral catalytic functions (coded by the *pol* gene). *pol* translation is dependent on frameshift or readthrough events, during upstream *gag* translation. Some essential cellular genes also need a recoding event to be expressed. Factors necessary for recoding are described :

- slippery sequences for -1 frameshifting,
- special tRNAs for +1 frameshifting
- poor stop context allowing readthrough,
- transaltionnal pause necessary to favor recoding *versus* normal coding.

To isolate factors modifying recoding, we used yeast *S. cerevisiae* as a model. I constructed versatile expression vectors containing *lacZ* and *luc* reading frames, separated by a recoding target. In-frame and out-of-frame luciferase activities comparison, normalized by the respective - galactosidase activity, is used to measure recoding efficiency. TMV (tobacco mosaic virus) readthrough and HIV-1 (human immunodeficiency virus) -1 frameshift efficiencies has been measured in mouse cells and in yeast. Similar results in both organisms validate yeast as a good model to study recoding.

HIV frameshift is stimulated by a stem-loop. We asked whether this effect was dependent on the thermodynamic stability of the structure. Frameshifting efficiency measurements for various G-C contents in the stem-loop show a good correlation between recoding efficiency and structure stability.

A genetic approach in yeast (reporters strains construction, srceening procedures elaboration) has been initiated to identify *trans* acting factors. A review of *trans* acting factors already described is presented, and the importance of such studies on translational fidelity control, and translation process in general, is discussed.



## Résumé

Ce manuscrit présente des travaux réalisés dans le domaine de la traduction des ARN messagers en protéines. Le recodage est un phénomène traductionnel qui permet l'expression de fonctions virales ou cellulaires grâce à une modification ponctuelle et programmée des règles canoniques de la traduction. Le recodage permet par exemple l'expression des fonctions catalytiques (codées par le gène *pol*) de la plupart des rétrovirus. La traduction du gène *pol* dépend d'un décalage de cadre de lecture ou de la translecture d'un codon stop après la traduction du gène *gag* en aval. Certains gènes cellulaires essentiels, de la bactérie à l'homme, nécessitent également un événement de recodage pour leur expression. Dans l'introduction les facteurs responsables du recodage sont décrits :

- séquences glissantes pour le décalage de phase en -1,
- ARNt spéciaux provoquant le décalage de phase en +1,
- codons stops favorables à la translecture,
- pause de l'appareil traductionnel permettant de favoriser le recodage *versus* le codage normal.

Afin d'isoler des facteurs modulant le taux de recodage, nous avons utilisé la levure *Saccharomyces cerevisiae* comme modèle d'étude. J'ai construit des vecteurs d'expression versatiles (bactéries, levures, mammifères) contenant les phases de lecture *lacZ* et *luc*, séparées par une séquence cible de recodage. La comparaison des activités luciférase de constructions en phase et hors de phase, normalisées par les activités -galactosidase respectives permet de mesurer le taux de recodage. J'ai quantifié en parallèle dans des cellules d'eucaryotes supérieurs et dans la levure le taux de translecture du TMV (virus de la mosaïque du tabac) ou de décalage de phase du VIH-1 (virus de l'immunodéficience humaine). Des résultats similaires entre les deux types de cellules valide la levure comme système modèle d'étude du recodage.

Le décalage de phase du VIH est stimulé par une tige-boucle en aval de la séquence glissante. Nous avons voulu savoir si l'effet de cette structure était dépendant de sa stabilité thermodynamique. L'étude du taux de décalage de phase en -1 en fonction du contenu en G-C de la tige-boucle a permis de montrer une corrélation entre taux de recodage et stabilité de la structure.

Une approche génétique chez la levure (construction de souches rapporteurs, élaboration d'un crible basé sur le taux d'expression de la  $\beta$ -galactosidase) a été initiée afin d'identifier des facteurs modulant en *trans* le taux de recodage. Après une revue bibliographique concernant de tels facteurs déjà identifiés, l'importance de telles études pour la compréhension du contrôle de la fidélité de la traduction, et de la traduction en général, est discutée.