

L'initiation de la traduction chez les eucaryotes, source de diversification et de modulation de l'expression des gènes

Olivier Jean-Jean, Michel Cassan, Jean-Pierre Rousset

Société Française de Génétique

Président

A. Nicolas

Président d'honneur

F. Jacob

Vice-présidents

R. Berger

H. Pinon

C. Stoll

Secrétaire général

M. Solignac

Trésorier

P.-M. Sinet

*Prière d'adresser toute correspondance
au Secrétariat général de la SFG,
Michel Solignac, laboratoire de biologie
et génétique évolutives, bâtiment
13, Cnrs, 91198 Gif-sur-Yvette Cedex,
France.*

Comité de rédaction

A. Bernheim

M. Bolotin-Fukuhara

M. Fellous

J. Générumont

B. Michel

R. Motta

A. Nicolas

S. Sommer

D. de Vienne

Secrétaire

M.-L. Prunier

Les progrès réalisés ces dernières années dans la connaissance du mécanisme de traduction des ARN messagers chez les eucaryotes révèlent l'importance de ce processus dans la modulation de l'expression des gènes et le rôle clé qu'y joue l'étape d'initiation. De ces études, il ressort que l'expression de certains gènes codant pour des protéines de régulation (facteurs de transcription, proto-oncogènes, facteurs de croissance, récepteurs d'hormone, etc.) semble soumise à une régulation traductionnelle.

La traduction de la séquence codante d'un ARN messager est contrôlée par l'énorme complexe macromoléculaire qu'est le ribosome. Bien que les mécanismes de traduction des gènes soient similaires chez les procaryotes et chez les eucaryotes, des différences importantes les séparent. En premier lieu, les ARN messagers procaryotes contiennent souvent plusieurs gènes (ou cistrons), alors que les ARN messagers eucaryotes sont presque toujours monocistroniques; de plus, chez les eucaryotes, les ARN messagers portent à l'extrémité 5' une « coiffe » formée d'un ribonucléotide triphosphate méthylé et, à l'extrémité 3', une « queue » polyadénylée. Une autre différence majeure concerne l'étape d'initiation de la traduction.

Chez les procaryotes, l'initiation de la traduction se fait par fixation de la petite sous-unité du ribosome sur une séquence particulière de l'ARN messager (séquence de Shine-Dalgarno) située à proximité du codon d'initiation, qui est en général le codon méthionine AUG. Le

positionnement de l'ARN de transfert (ARNt) méthionine initiateur puis la fixation de la grosse sous-unité du ribosome achèvent l'étape d'initiation.

Chez les eucaryotes, l'initiation de la traduction commence par la reconnaissance de l'extrémité 5' coiffée de l'ARN messager par la petite sous-unité du ribosome chargée d'un ARNt méthionine initiateur et finit par la reconnaissance du codon d'initiation. La fixation de la grande sous-unité sur la petite sous-unité permet au ribosome ainsi constitué de débiter la synthèse protéique. Contrairement à la relative simplicité du processus d'initiation chez les procaryotes, chez les eucaryotes, un grand nombre de facteurs formés pour la plupart de plusieurs sous-unités participent à l'initiation de la traduction (*figure 1*). Parmi les mécanismes de régulation de l'initiation de la traduction, la phosphorylation de certains facteurs d'initiation (en particulier du facteur eIF-2; voir *figure 1*) joue un rôle prépondérant. Ces phosphorylations en cascade ont été décrites en détail dans deux revues récentes [1, 2]. Nous nous limiterons ici à l'étude des éléments de régulation de la traduction portés par l'ARN messager. Ces éléments qui influencent en *cis* le processus de traduction sont divers. Certains, comme la séquence nucléotidique qui entoure le codon d'initiation, interviennent dans la reconnaissance du site d'initiation par la petite sous-unité du ribosome. D'autres agissent sur le passage du ribosome de l'extrémité 5' coiffée de l'ARN messager au codon d'ini-

tiation et sont soit des structures secondaires de l'ARN messager, soit de petits cadres de lecture placés en amont d'un gène. Ces signaux permettent non seulement la modulation de l'expression mais donnent aussi à certains gènes un caractère plurifonctionnel par le jeu d'initiations multiples.

Il faut rendre hommage au travail de Kozak qui a établi dès 1978 un modèle général d'initiation de la traduction chez les eucaryotes connu sous le nom de *scanning model* [3]. Ce modèle a été confirmé par la plupart des travaux réalisés sur l'initiation de la traduction. Cependant, comme nous le verrons à la fin de cette revue, Pelletier et Sonenberg ont démontré en 1988 [4] qu'il existe une exception à ce modèle appelée « initiation interne de la traduction ». Celle-ci, tout d'abord décrite chez quelques virus, a été récemment étendue à des ARN messagers cellulaires.

Le scanning model

Dans un article de 1989, Kozak écrit : « Le *scanning model* postule que la sous-unité 40S du ribosome (portant l'ARNt méthionine initiateur et des facteurs d'initiation) se lie d'abord à l'extrémité 5' de l'ARN messager et migre ensuite, s'arrêtant au premier codon AUG dans un contexte favorable pour l'initiation de la traduction » [5]. Selon ce modèle, la première étape de l'initiation est l'association de la petite sous-unité du ribosome sous sa forme activée à l'ARN messager. La sous-unité 40S combinée à plusieurs facteurs d'initiation qui maintiennent dissociées les deux sous-unités du ribosome, et à l'ARNt méthionine initiateur associé au facteur eIF-2 activé par le GTP, constitue le complexe de préinitiation 43S (figure 1). L'ARN messager se trouve alors, soit sous sa forme de transport, associé à des protéines nucléaires (particule RNP), soit sous forme de polysome, c'est-à-dire, associé à des ribosomes en cours de traduction. En se fixant à la coiffe du mes-

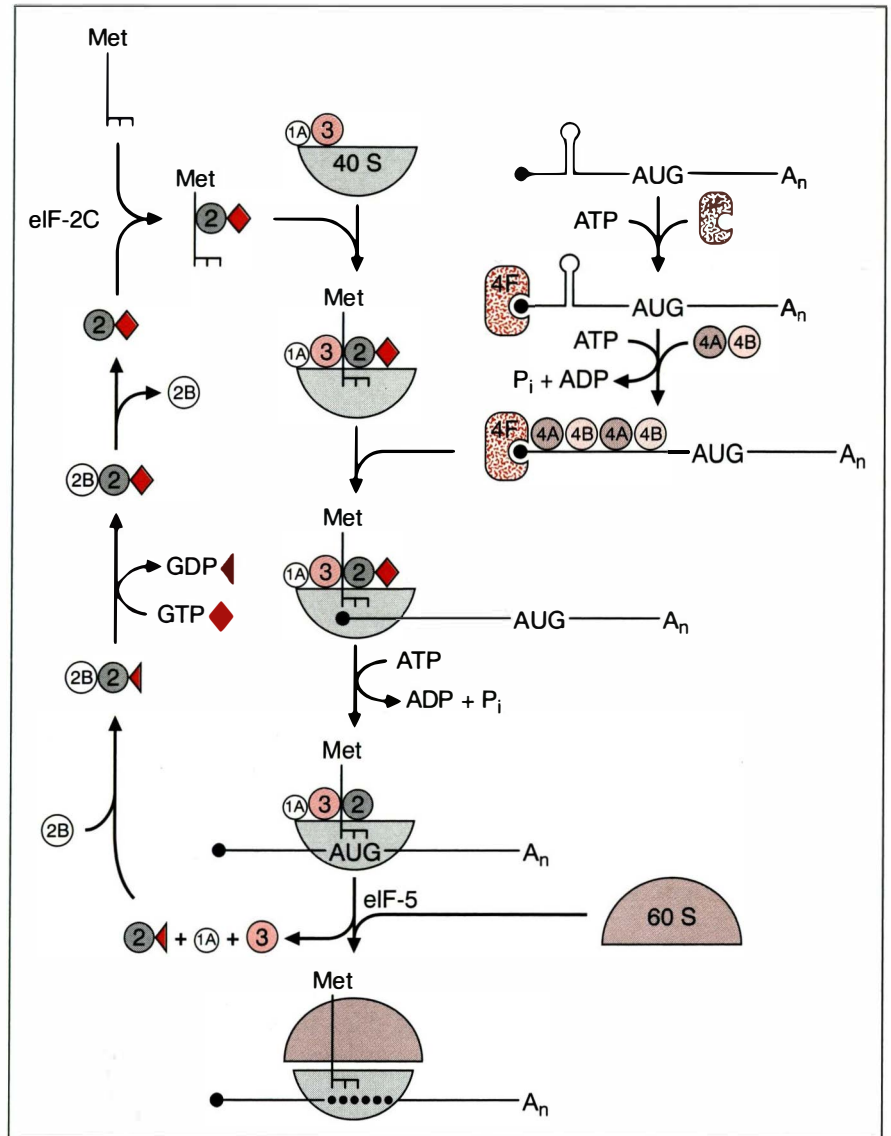


Figure 1. **Schéma de l'initiation de la traduction chez les eucaryotes d'après Hershey [1].** Le complexe de pré-initiation se forme par fixation de l'ARNt méthionine initiateur, associé au facteur eIF-2-GTP, à la sous-unité 40S du ribosome qui porte déjà les facteurs eIF-1A et eIF-3. Le complexe de pré-initiation se fixe ensuite sur l'extrémité 5' coiffée de l'ARN messager. Le complexe d'initiation ainsi formé glisse le long de l'ARN jusqu'au premier codon AUG et l'initiation s'achève par l'éjection des facteurs de préinitiation et fixation de la sous-unité 60S du ribosome (au centre). Cette fixation utilise l'hydrolyse du GTP lié au facteur eIF-2 qui est recyclé (à gauche). La reconnaissance de l'extrémité coiffée de l'ARN messager par le complexe de préinitiation nécessite la fixation préalable du facteur eIF-4F et le déroulement des structures secondaires par les facteurs eIF-4A et eIF-4B (à droite).

sager, le facteur eIF-4F semble jouer un rôle prépondérant dans la reconnaissance de l'extrémité 5' du message par le complexe de préinitiation 43S. Deux autres facteurs se fixent aussi à l'extrémité 5' du message : les facteurs eIF-4A et eIF-4B. Utilisant l'hydrolyse de l'ATP comme source d'énergie, l'activité ARN-hélicase de eIF-4A associé à eIF-4B déroule les structures secondaires de l'ARN message et permet au complexe d'initiation de progresser vers le premier codon d'initiation.

Un des problèmes qui restent à élucider dans la *scanning model* est le mécanisme de progression du complexe d'initiation sur l'ARN message. On ne sait pas si cette progression est passive, suivant simplement le déroulement des structures secondaires de l'ARN, ou si le complexe d'initiation, lié aux facteurs eIF-4A et eIF-4B, utilise pour sa progression une partie de l'énergie fournie par l'hydrolyse de l'ATP. L'hypothèse de la progression passive implique qu'en absence de structure secondaire, le complexe d'initiation peut diffuser librement sur l'ARN message et peut donc revenir vers l'extrémité 5', ou poursuivre son mouvement vers l'extrémité 3' de l'ARN message. La fixation d'autres complexes à l'extrémité 5' de l'ARN message et leur migration seraient alors les seules contraintes entraînant les complexes d'initiation dans la direction 3'. Au contraire, l'hypothèse d'une progression active liée à la progression des facteurs eIF-4A et eIF-4B donnerait au mécanisme de *scanning* son caractère orienté. On peut imaginer, pour expliquer cette progression, un processus analogue à celui qui permet la progression de l'ADN-polymérase III d'*Escherichia coli* sur l'ADN [6]. Dans ce cas, l'hydrolyse de l'ATP permet la fixation de la sous-unité β de l'ADN-polymérase III sur l'ADN. Cette sous-unité β a une structure annulaire qui lui permet de glisser dans les deux directions le long de l'ADN. Après fixation des sous-unités catalytiques de l'ADN-polymérase, la sous-unité β entraîne le complexe

enzymatique dans le sens de la synthèse de l'ADN. Un des facteurs d'initiation, eIF-4A ou eIF-4B, pourrait jouer un rôle analogue à celui de la sous-unité β et, suivant le déroulement des structures secondaires du message, tirer le complexe d'initiation vers le premier codon d'initiation. L'accrochage du complexe d'initiation sur l'ARN message pourrait se faire sur le modèle d'un anneau sur un fil (traduction de la belle expression *a ring on a string*). L'anneau serait constitué d'un sillon de la sous-unité 40S fermé par des facteurs d'initiation dont le facteur eIF-2-GTP. On peut aussi supposer qu'un des facteurs d'initiation (peut-être eIF-4B à cause de sa structure dimérique) forme, comme la sous-unité β , un anneau autour de l'ARN message et joue pour le complexe d'initiation le rôle de la sous-unité β pour l'ADN-polymérase III d'*E. coli*.

Le complexe d'initiation poursuit sa migration jusqu'à ce qu'il rencontre un codon d'initiation, le plus souvent un codon AUG, dans un contexte favorable. Quels sont les éléments qui permettent la reconnaissance du codon AUG et du contexte favorable ? L'appariement de l'anticodon de l'ARNt méthionine initiateur au codon AUG entre sûrement en jeu. En effet, chez la levure *Saccharomyces cerevisiae*, la transformation de l'anticodon 5'CAU3' d'un des ARNt méthionine initiateur en 5'CCU3', dirige l'initiation sur le codon AGG [7]. Le facteur eIF-2 semble aussi impliqué dans ce processus de reconnaissance du codon AUG. La mutation d'une des deux sous-unités du facteur eIF-2 de la levure aboutit à l'initiation aberrante à un codon UUG [8]. En revanche, le mystère reste entier sur le mécanisme de reconnaissance de la séquence favorable qui entoure le codon AUG. Chez *E. coli*, la complémentarité entre l'extrémité 3' de l'ARN ribosomique 16S et la séquence de Shine-Dalgarno, située quelques nucléotides en amont du codon d'initiation, joue un rôle clef dans le processus de reconnaissance. De plus, comme l'a démontré Drey-

fus en 1988 [9], l'information portée par une séquence plus large que la séquence de Shine-Dalgarno (de -20 à +15 autour du codon d'initiation) est indispensable à la constitution du site de fixation du ribosome d'*E. coli*. Une telle complémentarité entre la région d'initiation de la traduction et l'ARN ribosomique n'existe pas chez les eucaryotes.

L'arrêt du complexe de préinitiation sur l'AUG dans un bon contexte provoque la fixation de la sous-unité 60S du ribosome. Cette fixation nécessite la présence du facteur eIF-5 et l'hydrolyse de la molécule de GTP liée au facteur eIF-2. Cette réaction entraîne l'éjection des facteurs d'initiation dont le facteur eIF-2-GDP. Une fois le ribosome constitué, la phase d'élongation commence.

Définition d'un contexte fort

On entend généralement par contexte fort ou contexte favorable la séquence qui induit l'initiation de la traduction par tous les ribosomes qui arrivent au codon AUG qu'elle entoure. Cette définition s'appuie sur deux éléments : l'identification d'une séquence consensus provenant de l'analyse d'une compilation des séquences situées en amont du premier codon AUG (*Tableau I*; [10]), et les expériences de mutagenèse dirigée sur la séquence consensus tirée de la compilation [11,12]. Il ressort de ces études que, pour une grande majorité de gènes, le codon d'initiation de la traduction est le premier codon AUG rencontré à l'extrémité 5' de l'ARN message. Le *tableau I* montre que la séquence consensus varie suivant l'embranchement phylogénétique considéré. Seule la prépondérance d'un A en position -3 est constante (par convention, le A du codon AUG est numéroté +1). Peut-être encore plus significatif, le calcul des fréquences du triplet qui précède le codon AUG montre que les triplets ACC et GCC sont utilisés de façon préférentielle chez les vertébrés alors que les non-vertébrés utilisent plus souvent le triplet AAA.

La détermination de l'importance relative de chaque nucléotide de la séquence consensus des vertébrés réalisée par Kozak [11, 12], complète cette analyse. Ces expériences confirment l'importance des nucléotides -1 à -6 de la séquence entourant le codon AUG et en particulier celle de la purine en -3. Le maintien d'un contexte fort en l'absence de purine en -3 requiert une homologie plus stricte à la séquence consensus et la présence d'un G à la position +4. Il est fréquent cependant que l'on considère à tort comme contexte fort la seule présence d'une purine en -3 associée à la présence d'un G en +4. Pour avoir une initiation de la traduction exclusive à un codon AUG, il semble qu'en plus de la purine en -3, les nucléotides en -1 et -2 et, dans une moindre mesure, la séquence -4 à -6 jouent un rôle important. Citons le cas d'une α -thalassémie où la délétion de deux nucléotides, qui transforme le contexte fort entourant le codon AUG du gène de l' α -globine en contexte faible, est à l'origine de la maladie [13].

Le meilleur moyen pour définir un contexte fort est de déterminer s'il permet ou non d'initier la traduction à un codon AUG situé en aval. L'un de nous (O.J.J.) a montré que

dans un ARN messager du virus de l'hépatite B, la séquence AGCAC-CAUGC est, malgré l'absence de G en +4, un contexte fort qui n'autorise pas l'initiation de la traduction à un codon AUG en aval et empêche la synthèse d'une forme courte de la même protéine. Deux espèces d'ARN messagers sont alors nécessaires au virus pour exprimer les deux formes de la protéine [14]. A l'opposé, chez le même virus, la séquence UGGGGCAUGG, qui contient une purine en -3 et un G en +4, n'est pas un contexte fort puisqu'elle laisse passer environ 5 % des complexes d'initiation, ce qui permet la traduction d'un cadre de lecture situé en aval [15].

Un contexte fort autour du premier codon AUG d'un gène n'autorise la synthèse que d'une seule forme de la protéine; cependant, dans un nombre croissant d'exemples, plusieurs formes de la même protéine sont synthétisées sur un même ARN messager par initiation de la traduction à plusieurs codons dans la même phase. Ces initiations multiples requièrent un contexte faible (ou défavorable) autour du premier codon AUG mais peuvent aussi se faire à des codons différents du codon AUG.

Initiations dans un contexte faible

Le contexte qui entoure un codon d'initiation AUG peut être défini comme faible lorsqu'une partie des complexes d'initiation glissant depuis l'extrémité 5' de l'ARN messager ne s'arrêtent pas à ce codon et continuent leur mouvement vers l'extrémité 3'. Ce phénomène de « fuite d'initiation », appelé *leaky scanning* [16], ne répond pas à une loi du tout ou rien. Comme pour le contexte fort, les expériences de mutagenèse dirigée réalisées par Kozak [11, 12] ont permis de définir certains éléments du contexte faible. Ce sont, non seulement l'absence de purine en -3 et de G en +4, mais aussi l'absence de C en -1 et -2. Malheureusement, il est impossible à l'heure actuelle d'estimer par la seule connaissance de la séquence entourant le codon d'initiation, la proportion des ribosomes qui vont y amorcer la synthèse protéique; seule l'approche expérimentale permet de déterminer le degré de faiblesse d'un codon d'initiation. La présence d'un contexte faible autour du premier codon AUG peut être soit responsable du faible niveau de synthèse protéique, soit dans d'autres cas aboutir à l'initiation à un codon AUG en aval.

Si l'initiation de la traduction à deux codons AUG dans la même phase de lecture a d'abord été mise en évidence chez des virus, le nombre de gènes cellulaires connus contenant deux, trois ou plus de trois codons AUG à l'extrémité 5' du cadre de lecture ne cesse d'augmenter [17]. Chez les virus, une telle multiplicité permet d'exprimer de façon coordonnée plusieurs formes d'une protéine (des protéines de structure en particulier). Pour les gènes cellulaires qui possèdent des codons d'initiation multiples, il est surprenant de constater qu'ils font partie de familles de gènes qui codent pour des protéines de régulation: facteurs de transcription, proto-oncogènes, récepteurs d'hormone, facteurs de croissance, médiateurs de l'inflammation ou du système immunitaire, ou autres pro-

Tableau I

FRÉQUENCE DES NUCLÉOTIDES AUTOUR DU CODON D'INITIATION		-6	-5	-4	-3	-2	-1	+4	+5
Vertébrés n = 2595	A	20	19	25	58	28	17	25	24
	G	42	24	13	33	16	23	46	22
	C	20	37	48	6	45	53	15	37
	U	18	20	9	3	12	7	13	17
Consensus	G	C	C	A/G	C	C	AUG	G	C
Drosophile n = 192	A	32	22	21	65	47	39	26	28
	G	28	19	12	20	12	19	35	21
	C	17	36	53	7	23	34	19	30
	U	23	22	14	8	18	8	20	20
Consensus	A	C	C	A	A	A	AUG	G	C
Levure n = 461	A	35	35	43	66	46	43	26	23
	G	18	12	10	18	9	15	28	12
	C	17	22	26	7	22	20	12	43
	U	30	30	21	8	23	21	34	22
Consensus	A	A	A	A	A	A	AUG	U	C

téines impliquées dans la transmission d'un signal.

L'initiation de la traduction aux codons non AUG

L'initiation de la traduction aux codons non AUG représente un cas particulier de l'initiation multiple. Elle s'effectue le plus souvent aux codons CUG et ACG, plus rarement au codon GUG et exceptionnellement au codon AUU [17]. Dans tous les cas, les codons non AUG sont des sites d'initiation faibles qui permettent également l'initiation de la traduction à un codon AUG en aval.

Quels sont les éléments qui décident de l'initiation à ces codons non AUG? Nous avons vu plus haut que le complexe d'initiation porte l'ARNt méthionine initiateur et que l'appariement entre l'anticodon de cet ARNt et le codon AUG est un des éléments de la reconnaissance du codon d'initiation. Le poids de cette reconnaissance dans le mécanisme d'initiation est certainement diminué dans le cas des codons non AUG. Cependant, d'autres éléments pourraient compenser ce déficit. Le plus important semble être le contexte entourant le codon non AUG. La séquence autour de ces codons est en effet toujours très proche de la séquence consensus décrivant le contexte fort. Le second élément pourrait être la présence d'une région riche en nucléotides G et C située immédiatement après le codon non AUG. Enfin, la présence d'une structure secondaire (même peu stable) de l'ARN messager placée une quinzaine de nucléotides en aval du codon non AUG (ce qui correspond à la taille de la séquence couverte par le complexe d'initiation), augmente l'efficacité de l'initiation [18].

Il semble que l'initiation de la traduction à des codons non AUG ne soit pas un phénomène rare. De plus, les gènes qui utiliseraient ce type d'initiation appartiennent aux mêmes familles, décrites plus haut, de gènes impliqués dans des proces-

sus de régulation (proto-oncogènes, facteurs de transcription, facteurs de croissance, etc.). Signalons l'exemple étonnant du gène codant pour un facteur de croissance des fibroblastes (bFGF) qui contient, en amont d'un codon AUG, trois codons CUG, dont deux au moins semblent servir de codon d'initiation de la traduction [19].

Initiations multiples et régulation de l'expression des gènes

Dans de nombreux cas d'initiations multiples, les formes longues du produit d'un gène sont des espèces mineures difficilement détectables *in vivo*. Pour contourner cette difficulté et augmenter le niveau de synthèse des différentes formes de la protéine, la plupart des études ont été réalisées dans des cellules en culture au moyen de vecteurs d'expression très puissants. Il ressort de ces études que les différentes formes d'une protéine, synthétisées à partir de plusieurs sites d'initiation, peuvent avoir des localisations cellulaires différentes, ou même des fonctions complémentaires ou opposées. Deux exemples illustrent bien les différences de localisation cellulaire. Dans le cas du gène *gag* du virus de la leucémie murine décrit par Prats [20], la forme courte de la protéine (Pr65^{gag}) forme les capsides virales alors que la forme longue (Pr85^{gag}), dont l'initiation se fait à un codon CUG situé 264 nucléotides en amont du codon AUG de Pr65^{gag}, est glycosylée et localisée sur la membrane des cellules infectées par le virus. Cette forme longue faciliterait la dissémination du virus. Dans cet exemple, la région comprise entre les deux sites d'initiation de la traduction contient une séquence qui code pour un peptide signal conférant à la protéine non seulement une localisation différente mais aussi une fonction différente. Dans le cas du bFGF, il a été démontré [21] que les formes longues synthétisées à partir des codons CUG portent à leur extrémité amino-terminale un signal de loca-

lisation nucléaire, alors que la forme courte, synthétisée à partir du codon AUG en aval et dépourvue de ce signal, reste dans le cytoplasme des cellules. Il a été montré par la même équipe toulousaine [22] que l'effet oncogène provoqué par la surexpression du bFGF nécessite la coopération des formes longues nucléaires et de la forme courte cytoplasmique.

Dans d'autres cas, les deux formes de la protéine ont la même localisation cellulaire mais exercent une action opposée. Pour le facteur de transcription CREM (pour *cAMP responsive element modulator*), qui se fixe à une séquence de l'ADN répondant à une stimulation par l'AMP cyclique, deux des multiples formes de la protéine sont synthétisées à partir de l'ARN messager CREMt par le jeu d'initiations à deux codons AUG dans la même phase : la fixation à l'élément régulateur de la forme courte (S-CREM) a un effet inhibiteur sur la transcription induite par l'AMP cyclique alors que celle de la forme longue (CREMt) a un effet activateur [23].

Le proto-oncogène *c-myc* code pour deux protéines, une forme longue initiée à un codon CUG et une forme courte initiée à un codon AUG situé en aval. La présence de la seule forme courte dans plusieurs lymphomes de Burkitt pourrait être responsable de l'effet oncogène [24]. Il a été montré que la faible efficacité d'initiation au codon CUG du gène *c-myc* peut être modifiée dans des conditions particulières de culture de cellules [25]. Lorsque les cellules sont cultivées à haute densité, la synthèse de la forme longue de la protéine *c-myc* qui commence au codon CUG est augmentée de 5 à 10 fois alors que l'efficacité de traduction de la forme débutant au codon AUG n'est pas modifiée. Cette augmentation de la synthèse de la forme longue de la protéine est due à une déplétion en acides aminés du milieu de culture ; l'effet est reproduit lorsque les cellules sont cultivées dans un milieu carencé en méthionine. Les modèles proposés pour expliquer ce phé-

nomène mettent en jeu, d'une part, la modification des composants du complexe de préinitiation et en particulier du facteur eIF-2, dont nous avons vu plus haut qu'il pouvait être directement impliqué dans la sélection du codon d'initiation, et d'autre part, la désacylation de l'ARNt méthionine initiateur ou l'augmentation de l'activité méthionine-ARNt-synthétase de la cellule. Bien que cet exemple de régulation de l'efficacité d'initiation à deux codons en phase soit pour l'instant unique, bien d'autres mécanismes, physiologiques ou non, pourraient moduler l'abondance relative des différentes formes d'une protéine produites par le jeu d'initiations multiples.

Le cas particulier des ARN messagers bicistroniques

Le mécanisme de *leaky scanning* ou de fuite d'initiation permet, comme nous venons de le voir, l'initiation de la traduction à plusieurs codons dans la même phase de lecture et entraîne la synthèse de plusieurs formes de la même protéine à partir d'un seul ARN messager. Cependant, dans quelques cas, mis en évidence exclusivement chez des virus, ce mécanisme de fuite d'initiation permet aussi à une partie des complexes d'initiation d'atteindre un codon d'initiation dans une phase différente. Deux cadres de lecture sont alors traduits et deux protéines qui ont parfois des fonctions très différentes (par exemple, une protéine de structure et une protéine enzymatique), sont synthétisées sur le même ARN messager qui est dit bicistronique.

Généralement, l'extrémité 5' du second cadre de lecture d'un ARN messager bicistronique chevauche, sur quelques dizaines à quelques centaines de nucléotides, l'extrémité 3' du premier cadre de lecture. Dans le cas des ARN messagers polycistroniques procaryotes, le chevauchement des cadres de lecture est fréquent et le cistron situé en aval est traduit par réinitiation après ter-

mination de la traduction du cistron précédent. Cette réinitiation se fait par diffusion passive de la petite sous-unité du ribosome qui reste accrochée à l'ARN messager après dissociation du ribosome à la fin de la traduction du gène précédent [26]. Ce mécanisme de diffusion passive vers l'aval ou vers l'amont (*backward scanning*) permet d'expliquer, dans la majorité des cas, la traduction des différents gènes, chevauchant ou non, des ARN polycistroniques procaryotes.

Un tel mécanisme de *backward scanning* et de réinitiation pourrait aussi expliquer la traduction de deux gènes chevauchant sur un ARN messager bicistronique eucaryote. Cependant, toutes les études sur la traduction des ARN messagers bicistroniques montrent que le chevauchement de deux cadres de lecture empêche la traduction du deuxième gène par *backward scanning* et réinitiation, ce qui est un des arguments les plus forts en faveur du modèle de progression active et orientée des complexes d'initiation des eucaryotes. Il est clair maintenant que seule la fuite d'une partie des complexes d'initiation au codon AUG du premier cistron dans un contexte faible permet l'initiation de la traduction du second cistron des ARN messagers bicistroniques eucaryotes.

Nous allons maintenant voir quelles sont les structures de l'ARN messager qui influent en *cis* sur le mécanisme d'initiation de la traduction. Le premier de ces éléments régulateurs utilise le processus de réinitiation.

La réinitiation

Le processus de réinitiation suppose qu'à la fin de la traduction d'un gène, les deux sous-unités du ribosome se séparent et que la petite sous-unité 40S, fixée sur l'ARN messager, reprenne son mouvement vers l'extrémité 3' et réinitie la traduction au premier codon d'initiation rencontré. Théoriquement, le mécanisme d'initiation de la traduction des eucaryotes schématisé sur la

figure 1 ne devrait pas permettre la réinitiation de la traduction. En effet, la sous-unité 40S ne porte plus l'ARNt méthionine initiateur activé par le facteur eIF-2-GTP. La fixation des facteurs d'initiation sur la sous-unité 40S et la reconstitution du complexe d'initiation pourrait donc être l'étape limitante du processus de réinitiation chez les eucaryotes. Dans les faits, les tentatives d'expression d'ARN messagers bicistroniques artificiels ont montré que la traduction du deuxième cistron était très faible. Il a cependant été établi que l'efficacité d'expression du second cistron augmentait avec la distance intercistronique puis diminuait au-delà d'une distance d'environ 200 nucléotides [27]. Cette distance optimale correspondrait à la distance nécessaire à la reconstitution du complexe d'initiation.

Si, pour un grand nombre d'ARN messagers, le premier codon AUG rencontré est le codon d'initiation du gène, on trouve de plus en plus d'exemples où la séquence 5' de l'ARN messager contient, en amont du gène, un ou plusieurs petits cadres de lecture ouverts, nommés « minicistrons », dont la taille peut aller de quelques dizaines à quelques centaines de nucléotides. Ces minicistrons peuvent être traduits et jouent un rôle important dans la régulation de l'expression du gène situé en aval. Ainsi, la destruction par mutagenèse du minicistron placé en amont du gène du FGF-5 provoque une augmentation de la synthèse de ce facteur de croissance et la transformation des cellules en culture [28] ; dans le cas d'un virus de plante, la destruction du minicistron par mutation de son codon AUG augmente l'expression de la réplicase du virus et transforme une infection localisée de la plante en une infection systémique [29].

Cependant, l'exemple le plus étudié de régulation de la traduction par réinitiation est celui du gène *GCN4* de *S. cerevisiae* qui code pour un facteur qui active la transcription d'une trentaine de gènes impliqués dans la biosynthèse des acides aminés. Alors que l'ARN messager du gène *GCN4*

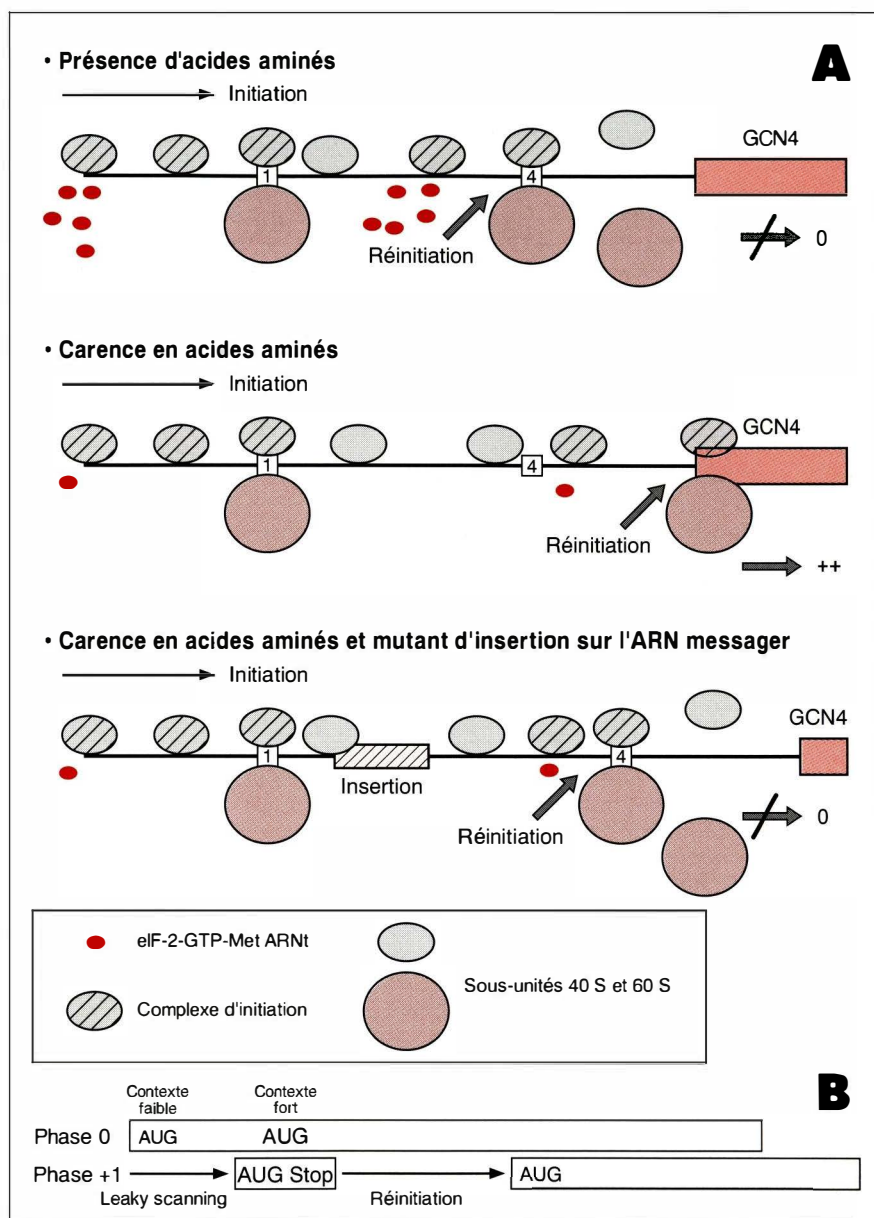


Figure 2. Influence des minicistrons sur l'expression des gènes. (A) Régulation de l'initiation du gène *GCN4* de *S. cerevisiae*. En présence d'acides aminés, l'abondance de facteur eIF-2 actif permet la réinitiation de la traduction au minicistron 4 dont la traduction inhibe la réinitiation au codon AUG du gène *GCN4*. En carence en acides aminés, l'inactivation partielle du facteur eIF-2 augmente le temps nécessaire à la reconstitution des complexes d'initiation et permet la réinitiation au codon AUG de *GCN4*. L'augmentation de la distance séparant les minicistrons 1 et 4 rétablit la réinitiation au minicistron 4 et l'inhibition de la traduction de *GCN4*. (B) Traduction d'un ARN messager bicistronique chez le virus de l'hépatite B. La présence d'un contexte faible autour du codon AUG du premier cadre de lecture provoque le leaky scanning d'une partie des complexes d'initiation. La traduction d'un minicistron de 7 codons permet à ces complexes d'initiation d'éviter un codon AUG dans un contexte fort et de réinitier la traduction au codon AUG du deuxième cadre de lecture.

est produit de façon constitutive, la protéine n'est synthétisée que lorsque les cellules sont privées d'acides aminés. L'extrémité 5' de l'ARN messager contient 4 minicistrons en amont du codon AUG du gène *GCN4* (figure 2A), dont seuls les minicistrons 1 et 4, distants de quelque 200 nucléotides, semblent avoir un rôle régulateur [30]. En présence d'acides aminés dans le milieu de culture, le complexe d'initiation se reforme rapidement après la traduction du minicistron 1 et la traduction est réinitiée au codon AUG du minicistron 4. La séquence qui entoure le codon stop du minicistron 4 empêche la réinitiation au codon AUG du gène *GCN4* et inhibe ainsi sa traduction [31]. Au contraire, dans des conditions de carence en acides aminés, l'inactivation par phosphorylation d'une partie des facteurs eIF-2 entraîne une augmentation du temps requis pour la reconstitution du complexe d'initiation. Le minicistron 4 n'est alors plus traduit et la traduction est réinitiée plus de 350 nucléotides après le minicistron 1, au codon AUG de *GCN4*. Toujours en condition de privation, l'augmentation artificielle de la distance séparant les minicistrons 1 et 4 de 200 à 356 nucléotides restaure l'effet inhibiteur du minicistron 4.

Dans l'exemple du gène *GCN4*, l'effet régulateur des minicistrons tient à la fois à la distance qui sépare les deux minicistrons et au contexte qui entoure la fin du minicistron 4. Hormis l'effet inhibiteur, d'autres fonctions peuvent être attribuées à un minicistron. Ainsi, pour un ARN messager bicistronique du virus de l'hépatite B (figure 2B), la traduction d'un minicistron permet aux complexes d'initiation d'éviter un codon AUG dans un contexte fort et de réinitier la traduction au codon AUG du second gène. En raison du chevauchement des cadres de lecture, la destruction du minicistron empêche la traduction du second gène par réinitiation [15].

Le mécanisme d'inhibition par traduction d'un minicistron n'est pas encore complètement élucidé mais

le polypeptide codé par le minicistron ne semble avoir aucune action en *trans* sur le processus de traduction.

Autres éléments de modulation

Les autres éléments modulateurs de l'initiation de la traduction sont de courtes séquences ou des structures secondaires de l'ARN messager qui influent de façon positive ou le plus souvent négative sur l'initiation de la traduction. Le mécanisme fréquemment invoqué pour expliquer ces effets est la fixation d'un facteur qui modulerait de façon spécifique la traduction d'un gène ou d'un ensemble de gènes.

Cependant, en elle-même, la présence d'une structure secondaire peut inhiber l'initiation de la traduction. Elle doit pour cela, soit être placée très près (à moins de quinze nucléotides) de l'extrémité 5' de l'ARN messager pour empêcher la reconnaissance de la structure en coiffe de l'ARN messager par le facteur eIF-4F, soit avoir une très grande stabilité pour bloquer l'action de l'ARN-hélicase et arrêter la progression du complexe d'initiation [32].

Le seul exemple bien documenté de répression de la traduction par une protéine se fixant sur une structure secondaire de l'ARN messager est celui des gènes qui contrôlent le métabolisme du fer dans la cellule (revues dans [1 et 2]). En cas de déplétion en fer (*figure 3*), l'inhibition de la traduction est provoquée par la fixation d'une protéine spécifique sur une structure secondaire appelée IRE (pour *iron responsive element*) et située dans la région 5' non traduite des ARN messagers de la ferritine (protéine de stockage du fer) et de la 5'-aminolévulinat-synthétase (première enzyme de la voie métabolique de l'hème). De plus, l'ARN messager du récepteur de la transferrine qui permet le captage du fer par la cellule, est stabilisé par la fixation de cette même protéine sur plusieurs structures IRE

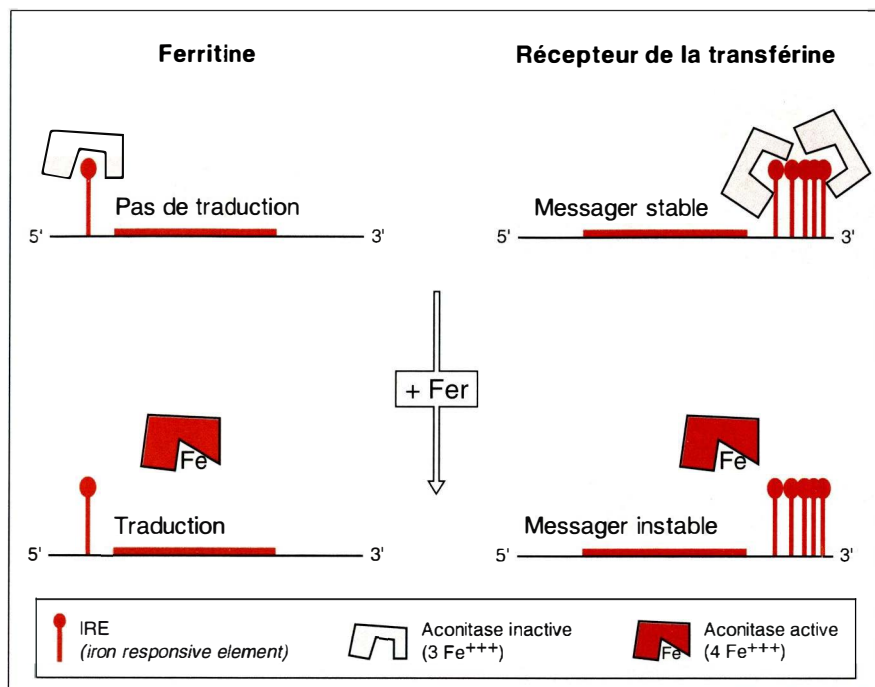


Figure 3. **Régulation traductionnelle par fixation d'une protéine spécifique sur une structure secondaire de l'ARN messager.** En absence de fer, l'aconitase peut se fixer sur une structure secondaire appelée IRE (iron responsive element) située dans la région 5' de l'ARN messager de la ferritine et bloque les complexes d'initiation. De plus, la fixation de l'aconitase sur plusieurs structures IRE situées en 3' du gène du récepteur de la transferrine stabilise le message. En présence de fer, l'aconitase fixe un atome de fer supplémentaire et perd son affinité pour la structure IRE.

situées à son extrémité 3'. La protéine spécifique de la structure IRE est la forme réduite (contenant trois atomes de fer) et inactive de l'aconitase mitochondriale. Le passage de la forme réduite à la forme active oxydée (par fixation d'un quatrième atome de fer) entraîne une perte de son affinité pour la structure IRE. Existe-t-il des éléments activateurs de l'initiation de la traduction ? Jusqu'à présent aucun exemple d'activation directe de la traduction par une protéine se fixant sur une séquence particulière de l'ARN messager n'a été publié. Dans la plupart des cas d'activation décrits, la fixation d'une protéine sur la séquence 3' non codante entraînerait la stabilisation

de l'ARN messager. Dans d'autres cas, la protéine activatrice pourrait modifier la localisation cytoplasmique de certains ARN messagers (voir [2] pour revue). Cependant, lors du développement embryonnaire, certains ARN messagers présents dans la cellule ne sont traduits qu'à un stade précis du processus de maturation. L'activation de la traduction de ces messagers se fait soit par addition d'une queue polyadénylée à leur extrémité 3', soit en déplaçant une protéine inhibitrice [33]. Ce mécanisme de démasquage d'un sous-ensemble d'ARN messager permet de coordonner la synthèse de protéines qui interviennent au même stade du développement.

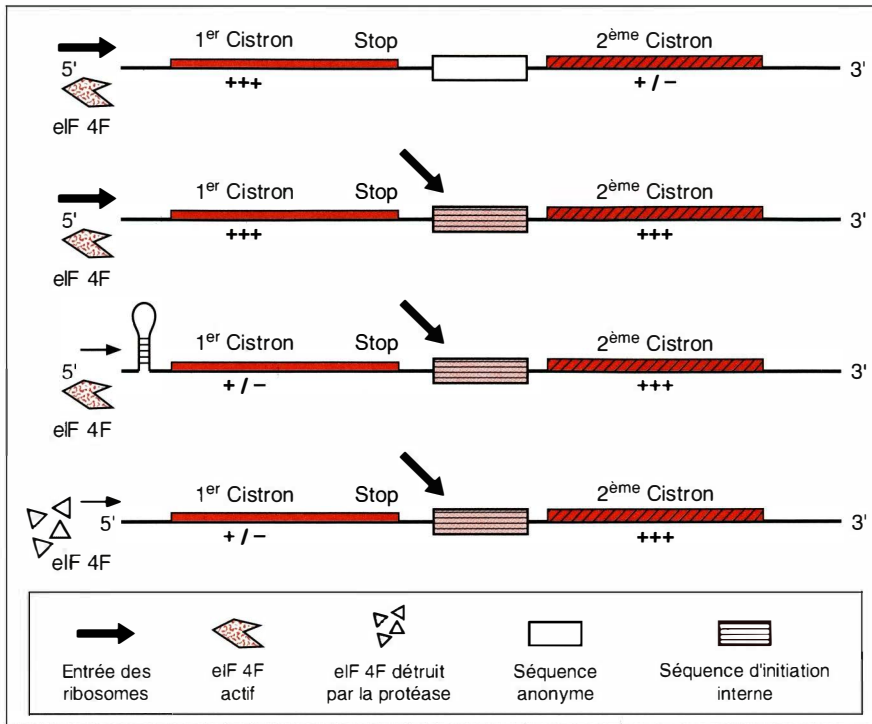


Figure 4. **Utilisation d'ARN messagers bicistroniques artificiels pour mettre en évidence l'initiation interne de la traduction.** En l'absence de site d'initiation interne, seul le premier cistron d'un ARN messager bicistronique est traduit efficacement. L'introduction d'une séquence d'initiation interne dans la région intercistronique permet la traduction efficace des deux cistrons. L'addition d'une structure secondaire stable à l'extrémité 5' de l'ARN messager ou la destruction du facteur eIF-4F empêchent l'initiation de la traduction du premier cistron sans affecter celle du second cistron par initiation interne.

L'initiation interne de la traduction

Le mécanisme d'initiation interne de la traduction et encore appelé entrée directe du ribosome, et découvert chez le poliovirus [4], est une exception au *scanning model*. L'infection d'une cellule par le poliovirus arrête la traduction des ARN messagers cellulaires et ne permet que la traduction quasi exclusive de son ARN génomique. Or, l'ARN génomique du poliovirus présente la caractéristique inhabituelle pour un messager d'être dépourvu de coiffe à son extrémité 5'. De plus, l'initiation de la traduction de

l'ARN messager de poliovirus ne requiert pas la présence du facteur eIF-4F et l'on sait depuis peu [34] qu'une des protéases virales nécessaire à la maturation de la polyprotéine induit la protéolyse de la sous-unité p220 du facteur eIF-4F, ce qui pourrait être la cause de l'arrêt des synthèses cellulaires. Ces observations suggéraient l'existence d'un mécanisme d'initiation interne de la traduction. Sa démonstration est venue de trois expériences [4] :

- des ARN messagers hétérologues liés à l'extrémité 5' non codante de l'ARN messager du poliovirus sont traduits dans des cellules infectées par le poliovirus ;

- la séquence 5' du poliovirus introduite dans la région intergénique d'un ARN messager bicistronique chimérique permet la traduction efficace du second cistron ;

- enfin, l'introduction d'une structure secondaire stable proche de l'extrémité 5' coiffée de ce même ARN messager bicistronique, ou l'infection des cellules par le poliovirus entraîne une inhibition de la traduction du premier cistron alors que la traduction du second cistron n'est pas affectée (figure 4).

Ces expériences montrent qu'une séquence de l'extrémité 5' de l'ARN messager de poliovirus permet la traduction d'un gène de façon indépendante de l'extrémité coiffée de l'ARN messager. Le mécanisme d'initiation proposé est l'entrée directe du complexe d'initiation en amont du gène à traduire sans *scanning* depuis l'extrémité 5'. L'utilisation d'ARN messager bicistronique a servi à définir la région d'entrée du complexe d'initiation sur l'ARN messager du poliovirus. La même approche expérimentale a ensuite permis de montrer que le mécanisme d'initiation interne de la traduction était utilisé par d'autres ARN messagers et en particulier par un ARN messager cellulaire [35].

Si toutes ces régions d'entrée du complexe de préinitiation ne présentent entre elles aucune homologie de séquence, certains éléments semblent cependant indispensables. Les ARN messagers traduits par initiation interne contiennent toujours à leur extrémité 5' une longue séquence non codante organisée en de nombreuses structures secondaires. L'autre élément important est une courte séquence de pyrimidines qui précède d'une trentaine de nucléotides un codon AUG qui n'est pas toujours le codon d'initiation [36]. En effet, dans certains cas, les complexes d'initiation glissent après leur entrée vers un codon d'initiation en aval [37].

Comment se fixe le complexe de préinitiation sur le site d'initiation interne en absence de facteur eIF-4F ? Plusieurs protéines cellulaires qui se fixent de façon spécifique

sur les structures secondaires des sites d'initiation interne, ont été mises en évidence [38, 39], mais leur rôle dans l'entrée du complexe de préinitiation est encore incertain. Si le mécanisme d'initiation de la traduction qui commence par la fixation du complexe de préinitiation sur la coiffe des ARN messagers est la règle lors du fonctionnement normal de la cellule, le mécanisme d'initiation interne pourrait permettre, comme le suggère Bensaude [40], de traduire certains gènes dans des conditions de stress, telles que l'infection par certains virus ou le choc thermique. Dans ces conditions, la réduction de l'activité du facteur eIF-4F serait responsable de l'inhibition de l'initiation de la traduction, et seuls les ARN messagers qui utilisent l'initiation interne de la traduction, échapperaient à cette inhibition.

Conclusions

La grande complexité des mécanismes mis en jeu lors de l'initiation de la traduction offre de nombreuses potentialités pour moduler l'expression des gènes. Nous avons montré dans cette revue comment des éléments de la structure primaire ou secondaire de l'ARN messager participent à cette régulation. La plupart de ces éléments agissent sur le complexe d'initiation dont la majorité des composants sont maintenant connus même si leur fonction est encore mal définie. Pour quelques-uns cependant, l'effet modulateur nécessite la présence de protéines cellulaires spécifiques qui, à quelques rares exceptions près, n'ont pas été identifiées.

La modification, par les conditions de culture, de la traduction du gène *GCN4* et de l'abondance relative des deux formes de la protéine codée par le gène *c-myc* illustrent bien le rôle régulateur du mécanisme de traduction et son importance dans le processus global d'expression des gènes. On peut spéculer que, pour tous les gènes qui possèdent des initiations multiples de la traduction,

l'efficacité relative d'initiation aux différents sites est soumise à des variations des conditions physiologiques, voire à l'état de différenciation de la cellule.

La transformation maligne de cellules en culture, provoquée par la surexpression de la petite sous-unité du facteur eIF-4F (eIF-4E), est un autre exemple du rôle régulateur de l'appareil de traduction [41]. Il est tentant de penser que cet effet pourrait être dû à la traduction préférentielle de messagers codants pour des facteurs de croissance [42]. Des découvertes restent à faire qui permettront de comprendre en détail ces mécanismes de régulation ■

O. Jean-Jean : département de biologie, École Normale Supérieure, 46, rue d'Ulm, 75230 Paris Cedex 05, France.

M. Cassan : Institut de Génétique et Microbiologie, université Paris-Sud, 91405 Orsay Cedex, France.

J.-P. Rousset : unité de génétique de la Différenciation, Institut Pasteur, 25, rue du Docteur-Roux, 75724 Paris Cedex 15, France.

TIRÉS A PART

O. Jean-Jean.

Remerciements

Nous remercions Marguerite Picard, Pierre-Olivier Angrand, Marc Dreyfus et Jean-Michel Rossignol d'avoir accepté de critiquer et de corriger ce manuscrit.

Références

- Hershey JWB. Translational control in mammalian cells. *Annu Rev Biochem* 1991; 60: 717-55.
- Kozak M. Regulation of translation in eukaryotic systems. *Annu Rev Cell Biol* 1992; 8: 197-225.
- Kozak M. How do eucaryotic ribosomes select initiation regions in messenger RNA. *Cell* 1978; 15: 1109-23.
- Pelletier J, Sonenberg N. Internal initiation of translation directed by a sequence derived from poliovirus RNA. *Nature* 1988; 334: 320-5.
- Kozak M. The scanning model of translation: an update. *J Cell Biol* 1989; 108: 229-41.
- Stukenberg PT, Studwell-Vaughan PS, O'Donnell M. Mechanism of the sliding β -clamp of DNA polymerase III holoenzyme. *J Biol Chem* 1991; 266: 11328-34.
- Cigan AM, Feng L, Donahue TF. tRNA^{met} functions in directing the scanning ribosome to the start site of translation. *Science* 1988; 242: 93-7.
- Donahue TF, Cigan AM, Pabich EK, Valavicius BC. Mutations at a Zn(II) finger motif in the yeast eIF-2 β gene alter ribosomal start-site selection during the scanning process. *Cell* 1988; 54: 621-32.
- Dreyfus M. What constitutes the signal for the initiation of protein synthesis on *Escherichia coli* mRNAs? *J Mol Biol* 1988; 204: 79-94.
- Cavener DR, Ray SC. Eukaryotic start and stop translation sites *Nucleic Acids Res* 1991; 19: 3185-92.
- Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 1986; 44: 283-92.
- Kozak M. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* 1987; 196: 947-50.
- Morlé F, Lopez B, Henni T, Godet J. α -Thalassaemia associated with the deletion of two nucleotides at position -2 and -3 preceding the AUG codon. *EMBO J* 1985; 4: 1245-50.
- Jean-Jean O, Levrero M, Will H, Pericaudet M, Rossignol JM. Expression mechanism of the hepatitis B virus (HBV) C gene and biosynthesis of HBe antigen. *Virology* 1989; 170: 99-106.
- Fouillot N, Carlier D, Tlouzeau S, Rossignol JM, Jean-Jean O. Translation of the hepatitis B virus P gene by ribosomal scanning as an alternative to internal initiation. *J Virol* 1993; 67: 4886-95.
- Kozak M. Bifunctional messenger RNAs in eukaryotes. *Cell* 1986; 47: 481-3.
- Kozak M. An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol* 1991; 115: 887-903.
- Kozak M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci USA* 1990; 87: 8301-5.
- Prats H, Kaghad M, Prats AC, Klagsbrun M, Lélías JM, Liauzin P, Chalon P, Tauber JP, Amalric F, Caput D. High molecular mass of basic fibroblast growth factor are initiated by alternative CUG codons. *Proc Natl Acad Sci USA* 1989; 86: 1836-40.
- Prats AC, De Billy G, Wang P, Darlix JL. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J Mol Biol* 1989; 205: 363-72.
- Bugler B, Amalric F, Prats H. Alternative initiation of translation determines cytoplasmic or nuclear localization of basic fibroblast growth factor. *Mol Cell Biol* 1991; 11: 573-7.

22. Couderc B, Prats H, Bayard F, Amalric F. Potential oncogenic effects of basic fibroblast growth factor requires cooperation between CUG and AUC-initiated forms. *Cell Reg*; 1991; 2: 709-18.
23. Delmas V, Laoide BM, Masquillier D, de Groot RP, Foulkes NS, Sassone-Corsi P. Alternative usage of initiation codon in mRNA encoding the cAMP-responsive-element modulator generates regulators with opposite functions. *Proc Natl Acad Sci USA* 1992; 89: 4226-30.
24. Hamm SR, King MW, Bentley DL, Anderson CW, Eisenman RN. A non-AUG translational initiation in *c-myc* exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. *Cell* 1988; 52: 185-95.
25. Hamm SR, Sloan-Brown K, Spotts GD. Translational activation of the non-AUG-initiated *c-myc* 1 protein at high cell densities due to methionine deprivation. *Genes Dev* 1992; 6: 1229-40.
26. Adhin MR, van Duin J. Scanning model for translational reinitiation in eubacteria. *J Mol Biol* 1990; 213: 811-8.
27. Kozak M. Effects of intergenic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol Cell Biol* 1987; 7: 3438-45.
28. Bates B, Hardin J, Zhan X, Drickamer K, Goldfarb M. Biosynthesis of human fibroblast growth factor-5. *Mol Cell Biol* 1991; 11: 1840-5.
29. Petty ITD, Edwards MC, Jackson AO. Systemic movement of an RNA plant virus determined by a point substitution in a 5' leader sequence. *Proc Natl Acad Sci USA* 1990; 87: 8894-7.
30. Abastado JP, Miller PF, Jackson BM, Hinnebusch AG. Suppression of ribosomal reinitiation at upstream ORFs in amino acid-starved cells forms the basis for *GCN4* translational control. *Mol Cell Biol* 1991; 11: 486-96.
31. Miller PF, Hinnebusch AG. Sequences that surround the stop codons of upstream open reading frames in *GCN4* mRNA determine their distinct functions in translational control. *Genes Dev* 1989; 3: 1217-25.
32. Kozak M. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol Cell Biol* 1989; 9: 5134-42.
33. Richter JD. Translational control during early development. *Bio Essays* 1991; 13: 179-83.
34. Hellen CTU, Fäcke M, Kräusslich HG, Lee CK, Wimmer E. Characterization of poliovirus 2A proteinase by mutational analysis: residues required for autocatalytic activity are essential for induction of cleavage of eukaryotic initiation factor 4F polypeptide p220. *J Virol* 1991; 65: 4226-31.
35. Macejak DG, Sarnow P. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* 1991; 353: 90-4.
36. Pilipenko EV, Gmyl AP, Maslova SV, Svitkin YV, Sinyakov AN, Agol VI. Prokaryotic-like *ais* elements in the cap-independent internal initiation of translation on picornavirus RNA. *Cell* 1992; 68: 119-31.
37. Belsham GJ. Dual initiation sites of protein synthesis on foot and mouth disease virus RNA are selected following internal entry and scanning of ribosomes. *EMBO J* 1992; 11: 1105-10.
38. Meerovitch K, Pelletier J, Sonenberg N. A cellular protein that binds to the 5'-noncoding region of poliovirus RNA: implication for internal translation initiation. *Genes Dev* 1989; 3: 1026-34.
39. Jang SK, Wimmer E. Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and essential binding of a cellular 57 kDa protein. *Genes Dev* 1990; 4: 1560-72.
40. Bensaude O. Initiation interne de la traduction et stress cellulaire: des infections virales au choc thermique. *médecine/sciences* 1992; 8: 63-4.
41. Latzaris-Karatzas A, Montine KS, Sonenberg N. Malignant transformation by the eukaryotic mRNA 5' cap binding protein (eukaryotic initiation factor-4E). *Nature* 1990; 345: 544-7.
42. Kahn A. Cancer et traduction des messagers (suite). *médecine/sciences* 1992; 8: 1001.

Summary

Initiation of translation in eukaryotes: diversity and modulation of gene expression

Recent progress in understanding the translation process in eukaryotes gives a new insight on the mechanisms by which gene expression is modulated. The initiation step of translation plays a key role in this control. Apart from eukaryotic viruses, a growing number of genes encoding regulatory proteins uses translation initiation to increase gene diversity. The scanning model of Kozak describes the mechanism by which the 40S ribosomal subunit enters at the capped 5' end of the messenger RNA and scans to the first initiator codon. Recognition of the initiation site by the scanning initiation complex is dependent on the context surrounding the initiator AUG codon. In this article, we describe the *ais*-acting elements of mRNA which influence translation initiation. A weak context around the AUG codon interferes with recognition of the initiator codon and allows initiation at downstream AUGs. This results either in the expression of two genes on a bicistronic mRNA, or in the synthesis of several forms of a protein if the initiator codons are in the same reading frame. This leaky scanning mechanism is also observed when non-AUG codons are used for translation initiation. The regulating elements acting on the scanning of the initiation complexes are secondary structures or translated minicistrons which, in most cases, decrease the translation efficiency at a downstream AUG. As an exception to the scanning model, the mechanism of internal initiation of translation seems to be used to translate some particular mRNA during viral infection or in stress conditions. These elements of translational control allow either the expression of the multiple functions of a gene, or the coordinate regulation of genes involved in the same biochemical pathway. What is more, they permit the fine adjustment of protein synthesis to the variations of cell growth conditions.