

Les régions non traduites des ARN messagers et leur rôle dans la synthèse protéique

Coralie Pallier

Le processus cellulaire conduisant à l'expression des ARN messagers comporte plusieurs étapes, contrôlées par les régions non traduites de ces mêmes ARN. Ces mécanismes de régulation post-transcriptionnelle résultent d'interactions entre ces régions non traduites, situées aux extrémités 5' et 3', et diverses protéines agissant en *trans*, qui contrôlent spécifiquement la localisation intracytoplasmique, la traduction et la stabilité ou la dégradation de l'ARNm. Celui-ci porte donc des séquences actives en *cis* qui permettent à la cellule, via des protéines régulatrices à effet en *trans*, d'adapter le taux de traduction et/ou la stabilité d'un ARNm selon les différentes conditions physiopathologiques.

A partir de la séquence d'un gène, la synthèse d'une protéine comporte obligatoirement différentes étapes: (1) la transcription du gène en pré-ARN messager; (2) la maturation du pré-ARN messager, qui comprend l'épissage des introns, l'addition de la coiffe et la polyadénylation; (3) le transport de l'ARN messager mature depuis le noyau jusqu'au cytoplasme, dans lequel sa concentration est contrôlée par des éléments permettant sa stabilisation ou sa dégradation; et simultanément (4) sa traduction plus ou moins efficace. L'expression d'un gène dans une cellule résulte donc de la régulation coordonnée de chacune de ces étapes. En particulier, lors du

transfert de l'ARNm depuis le noyau vers sa destination cytoplasmique, l'ARNm interagit avec des protéines régulatrices qui vont gouverner son devenir (*figure 1A*). Les mécanismes de la régulation post-transcriptionnelle résultent essentiellement d'interactions entre les régions non traduites situées aux extrémités 5' et 3' de l'ARNm, et des protéines à effet en *trans* qui contrôlent spécifiquement: (1) la localisation cytoplasmique; (2) la traduction et/ou (3) la stabilité de l'ARNm (*figure 1B*).

Chez les eucaryotes, les ARN messagers sont synthétisés par l'ARN polymérase II sous forme de précurseurs, appelés pré-ARN messagers. Au cours de leur maturation, une coiffe, constituée d'une guanine méthylée

ADRESSE

C. Pallier : Service de bactériologie-virologie du Pr P. Nordmann, Hôpital de Bicêtre, 78, rue du Général-Leclerc, 94275 Le Kremlin-Bicêtre Cedex, France.

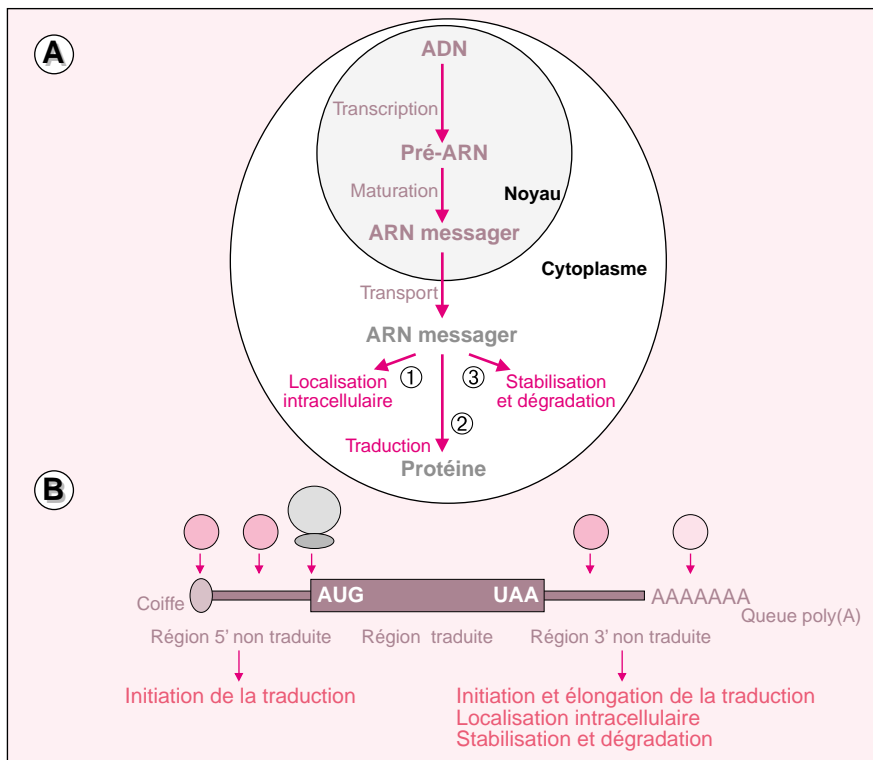


Figure 1. **A.** Les différentes étapes de l'expression des ARNm contrôlées par leurs régions non traduites. Après transcription des gènes, maturation et transport hors du noyau des ARNm, les extrémités 5' et 3' non traduites interviennent et dirigent l'expression des ARNm au cours des trois étapes suivantes: ce sont (1) l'adressage intracytoplasmique des ARNm, (2) le contrôle de leur traduction et (3) leur stabilisation ou leur dégradation. **B.** Mécanismes de la régulation post-transcriptionnelle relayée par les régions non traduites des ARNm. Les interactions entre les régions non traduites en 5' et en 3' des ARNm et des protéines régulatrices à effet en trans contrôlent l'adressage intracytoplasmique des ARNm, leur traduction et leur stabilité ou leur dégradation.

en position N7, est ajoutée à l'extrémité 5' de la plupart des pré-ARNm par une liaison 5'-5' triphosphate [1]. Cette coiffe participe aux réactions d'épissage et de polyadénylation nucléaires, au transport nucléocytoplasmique des ARNm, à l'initiation de leur traduction et à leur protection contre la dégradation par les exoribonucléases dans le sens 5' 3' [2]. En aval de la région 5' non traduite (5'UTR), le cadre ouvert de lecture (ORF) débute avec le codon d'initiation AUG et s'achève avec le codon de terminaison de la traduction. Les pré-ARNm se terminent au niveau de la région 3' non traduite (3'UTR) à laquelle est ajoutée une séquence poly(A) au cours de de leur maturation [3]. Tous les transcrits synthétisés par l'ARN polymérase II

sont polyadénylés, à l'exception de ceux des histones. Dans un premier temps, les pré-ARNm sont clivés au niveau d'une région située de 10 à 30 nucléotides en aval du signal de coupure et de polyadénylation AAUAAA, et à environ 30 nucléotides en amont d'une séquence poly(U) ou riche en doublets GU, appelée DSE (*downstream element*). Une dizaine de résidus adénosine sont alors ajoutés à l'extrémité 3' par la PAP (*poly(A) polymerase*). La distance entre le signal AAUAAA et le DSE détermine donc la région dans laquelle va s'effectuer le clivage. Le site précis de cette coupure est de préférence en 3' d'un résidu A, le plus souvent après le dinucléotide CA. L'efficacité du clivage semble être influencée par une séquence en

amont, riche en résidus U, appelée USE (*upstream element*), par l'environnement nucléotidique du signal AAUAAA et probablement par la structure secondaire même de la 3'UTR. Cependant, les modalités du clivage ne sont que partiellement connues: le CPSF (*cleavage and polyadenylation specificity factor*) et le CstF (*cleavage stimulation factor*), qui se fixent respectivement sur le signal AAUAAA et sur le DSE, sont indispensables. De même, la PAP semble intervenir dans le processus, mais la protéine qui porte l'activité endonucléase n'a pas encore été identifiée. Le clivage produit une extrémité 3'-OH sur les pré-ARNm, à laquelle la PAP ajoute une dizaine de résidus A, alors que le fragment présentant l'extrémité 5'-phosphate est rapidement dégradé. Dans un second temps, la séquence poly(A) est allongée, et cette extension dépend uniquement de la présence des 10 premiers résidus précédemment ajoutés et du CPSF. En effet, la séquence de 10 résidus A est nécessaire et suffisante pour que la PABP (*poly(A) binding protein*) s'y fixe et stimule l'activité de la PAP, en coopération avec le CPSF, pour permettre l'addition d'adénosines supplémentaires. La longueur initiale de la séquence poly(A) est d'environ 200 à 250 résidus chez les eucaryotes supérieurs et semble contrôlée par la PABP, mais elle va subir ultérieurement des modifications au niveau du cytoplasme. Cette queue poly(A) intervient dans le transport nucléocytoplasmique des ARNm, dans l'initiation de leur traduction et dans le contrôle de leur stabilité. Les ARNm coiffés, épissés et polyadénylés, sont associés à des protéines sous forme de complexes appelés hnRNP (*heterogeneous nuclear ribonucleoprotein*), avant d'être exportés vers le cytoplasme [4, 5]. Certaines de ces protéines seront relâchées pendant le processus de translocation, tandis que d'autres vont accompagner l'ARNm jusqu'au cytoplasme. Plusieurs protéines sont impliquées dans le transport nucléocytoplasmique des ARNm, et se fixent sur des séquences des ARNm appelées NES, (*nuclear export sequence*). De même, les constituants du pore nucléaire franchi par l'ARNm participent probablement à ce transport actif.

Après son acheminement vers le compartiment cytoplasmique approprié, l'ARNm est soit traduit, soit stocké sous forme de particules ribonucléoprotéiques appelés mRNP (*messenger ribonucleoprotein*) [6]. Ces mRNP renferment les ARNm associés à des protéines, et constituent une sorte de réservoir dans lequel les ARNm sont sous forme inactive mais stable, puisqu'ils sont inaccessibles à l'appareil de traduction et indisponibles pour les réactions de polyadénylation et de désadénylation cytoplasmiques. Les protéines qui interviennent dans l'organisation et la condensation des ARNm sous forme de mRNP sont classées en deux catégories. Les unes constituent la structure même des mRNP, elles sont appelées les *mRNP core protein*, alors que les autres sont responsables du « camouflé » complet des ARNm, il s'agit des *masking proteins*. Les modalités de la mobilisation des ARNm au sein de ces mRNP sont encore mal connues.

UTR et localisation cytoplasmique des ARNm

La distribution intracellulaire des ARNm n'est pas aléatoire. En effet, les ARNm sont répartis dans des régions spécifiques du cytoplasme et l'étude de leur localisation a fait l'objet de plusieurs revues [7-9]. En effet, si certains ARNm uniformément répartis ne sont actifs que dans des compartiments cytoplasmiques précis, d'autres s'accumulent dans des sites spécifiques. Ainsi, la concentration locale de ces ARNm entraîne une localisation ciblée de protéines correspondantes, assurant l'assemblage de complexes macromoléculaires au niveau de sites cellulaires spécifiques et/ou l'établissement d'un gradient de protéines. Ce mécanisme de contrôle de la répartition des ARNm nécessiterait moins d'énergie que celle des protéines correspondantes, puisqu'à partir d'une seule molécule d'ARNm, plusieurs molécules de protéines peuvent être synthétisées.

Les éléments permettant l'adressage des ARNm et leur rétention au niveau des différents compartiments cytoplasmiques sont essentiellement situés dans les 3'UTR [10], sans doute parce qu'à l'inverse des 5'UTR, ces régions ne sont soumises

à aucune contrainte, et fournissent un terrain propice à l'évolution des éléments de régulation. Ainsi, la 3'UTR est nécessaire et suffisante au contrôle de la distribution d'un ARNm réparti de façon non uniforme. Par son interaction avec des protéines régulatrices agissant en *trans*, l'ARNm va être dirigé et retenu vers le site où la protéine correspondante sera ultérieurement nécessaire (*figure 2*). Dans la compréhension de ces mécanismes de régulation, la *Drosophila* représente un modèle de choix. Ainsi, la localisation des ARNm *oskar* et *bicoid*, a été particulièrement bien étudiée. Ces messagers codent respectivement pour une protéine morphogène essentielle à la maturation des cellules germinales de l'embryon et à la segmentation de l'abdomen, et pour une protéine qui régit la formation du thorax et de la tête au niveau du pôle antérieur de l'ovocyte. D'après les travaux rapportés [11], l'adressage de ces ARNm semble pouvoir s'effectuer en plusieurs étapes, qui constituent d'une part le transport des ARNm vers leur destination finale, et d'autre part leur ancrage au niveau du site spéci-

fique. Ces étapes mettent en jeu des interactions entre les ARNm et différents composants du cytosquelette. Ainsi, le transport de l'ARNm de la protéine Vg1 au niveau du pôle végétal de l'ovocyte de xénope requiert la participation des microtubules, tandis que les microfilaments d'actine y assurent ensuite son ancrage [12]. En outre, ces éléments du cytosquelette pourraient également intervenir lors de la traduction des ARNm en favorisant leurs interactions avec les facteurs de la traduction, soit parce que les facteurs de la traduction sont eux-mêmes associés au cytosquelette, soit parce que cette interaction permet le recyclage des ribosomes [13]. En effet, si certains ARNm ne sont associés au cytosquelette que pendant leur traduction, on ne sait pas encore si cette association est nécessaire au recrutement des ARNm par l'appareil de traduction, ou si elle résulte de leur traduction.

UTR et traduction des ARNm

Arrivé à destination, l'ARNm mature sera traduit au cours de trois étapes

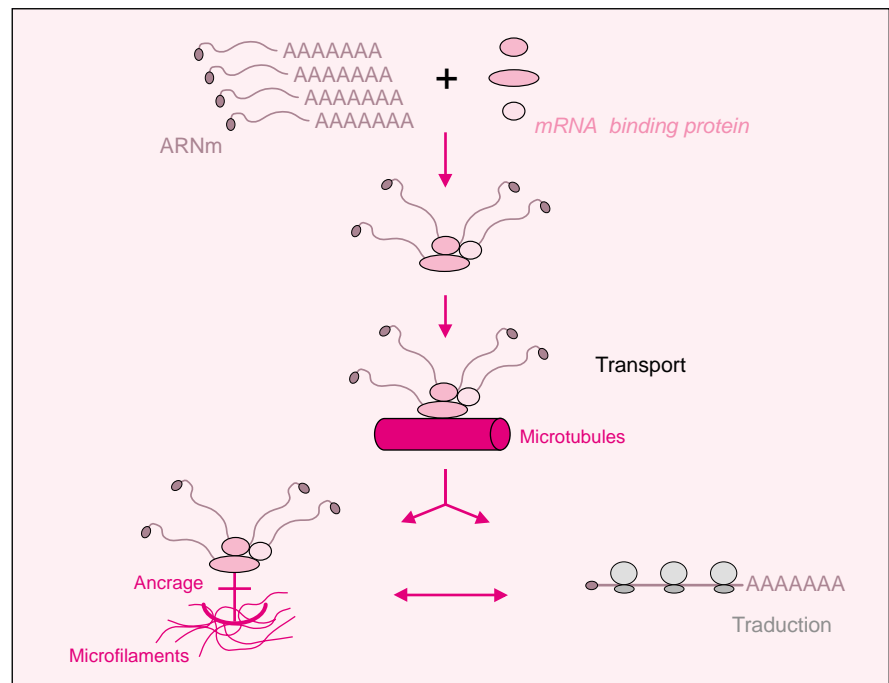


Figure 2. **Modèle de transport des ARNm.** Les ARNm sont pris en charge par des protéines qui se fixent à leur région 3' non traduite et les conduisent le long des microtubules vers leur compartiment cytoplasmique cible. Ils y sont alors maintenus par le biais d'interactions avec les microfilaments d'actine jusqu'à leur traduction (d'après [7]).

successives, l'initiation, l'élongation et la terminaison, décrites de manière détaillée dans la revue de J. Hershey, puis plus récemment par V. Pain [14, 15]. L'initiation peut se dérouler principalement selon deux modèles, celui du balayage et celui de l'entrée interne des ribosomes [16], qui sont résumés dans les figures 3 et 4 et dont la description vient d'être rapportée (*m/s 2000, n° 1, p. 77*). La traduction est un phénomène extrêmement bien contrôlé et sa régulation s'effectue toujours de façon rapide et réversible, quel que soit le mécanisme impliqué. Cette régulation peut être globale et affecter tous les ARNm, ou être sélective et ne toucher qu'une catégorie d'ARNm. Elle peut s'exercer aussi bien lors de l'initiation que lors de l'élongation ou de la terminaison. Cependant, dans la plupart des cas, c'est l'étape de l'initiation qui est contrôlée, soit par l'intermédiaire de l'appareil protéo-synthétique, soit par l'ARNm lui-même.

L'efficacité de l'initiation de la traduction est influencée par différents éléments situés essentiellement dans les régions non traduites de l'ARNm.

La 5'UTR

Le rôle de la 5'UTR dans la régulation de la traduction des ARNm se conçoit aisément puisque la sous-unité 40S est en étroite relation avec cette région pendant l'initiation.

La coiffe

La présence et l'accessibilité de la coiffe semblent être deux éléments déterminants dans l'efficacité de l'initiation de la traduction des ARNm. *In vivo*, un ARNm coiffé est toujours traduit de façon plus efficace que le même ARNm non coiffé, et cet effet est également observé *in vitro* mais de façon moins nette [17]. Seule la traduction des ARNm par initiation interne n'implique pas la coiffe.

Les structures secondaires

La présence de tiges et de boucles stables dans la 5'UTR d'un ARNm, ainsi que d'autres structures secondaires intramoléculaires, diminue l'efficacité de l'initiation de la traduction, les facteurs généraux d'initiation à activité hélicase ayant des difficultés à lisser de telles structures [18]. Cet effet dépend de la position de ces

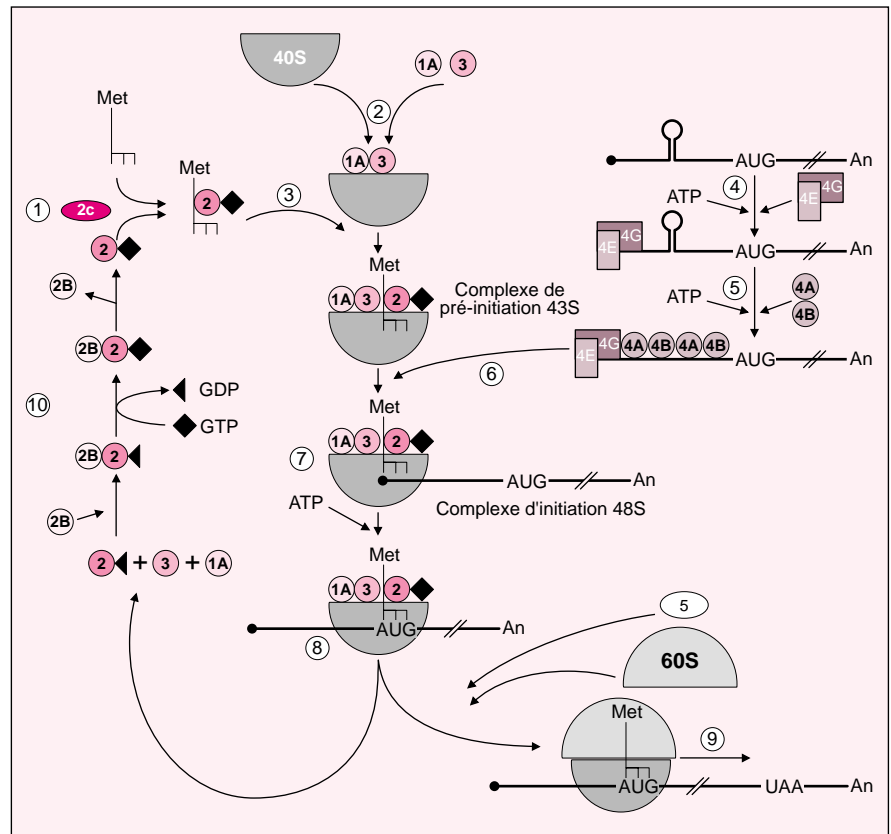


Figure 3. **Les différentes étapes de l'initiation de la traduction selon le modèle par balayage.** (1) mobilisation de l'ARNm^{met-i} par le facteur eIF2-GTP, (2) recrutement de la sous-unité 40S par les facteurs eIF1A et eIF3, (3) formation du complexe de pré-initiation 43S, (4) reconnaissance de la coiffe de l'ARNm par le facteur eIF4E, (5) élimination des structures secondaires par les facteurs eIF4A et eIF4B, (6) formation du complexe d'initiation 48S à l'aide du facteur eIF4G, (7) balayage de l'ARNm jusqu'à ce que le complexe d'initiation 48S reconnaisse un codon AUG d'initiation, (8) hydrolyse du GTP lié au facteur eIF2 en GDP par le facteur eIF5, induisant le relargage des autres facteurs d'initiation et l'association de la sous-unité 60S, (9) l'ARNm^{met-i} est positionné sur le codon AUG de l'ARNm et au site P du ribosome 80S et la traduction de la séquence codante peut alors commencer, (10) recyclage du facteur eIF2-GDP en eIF2-GTP par échange du GDP contre du GTP (eIF: eucaryotic initiation factor) (d'après [14]).

structures, car il est d'autant plus prononcé qu'elles sont situées à proximité de la coiffe, site d'entrée du complexe de pré-initiation 43S sur l'ARNm. L'assemblage du complexe de pré-initiation 43S au niveau de la coiffe semble donc plus gêné par la présence de telles structures secondaires, que ne l'est la progression de la sous-unité 40S le long de l'ARNm. De plus, la traduction de nombreux ARNm, et en particulier des ARNm codant pour des facteurs de croissance ou de différenciation, qui ont des 5'UTR très structurées, semble contrôlée par l'environnement cellulaire. Ainsi, l'efficacité de la traduc-

tion de transcrits contenant la 5'UTR de l'ARNm de l'ODC (*ornithine decarboxylase*) varie selon le type cellulaire dans lequel s'effectue cette traduction [19]. Cette efficacité est corrélée à l'état de phosphorylation du facteur eIF4E, mais ces résultats suggèrent également que des interactions spécifiques entre des protéines autres que les facteurs généraux d'initiation peuvent avoir un rôle essentiel. Ces interactions peuvent concerner les produits de traduction de l'ARNm, et la 5'UTR qui modulerait la formation de ces structures secondaires et régulerait ainsi l'efficacité de l'initiation de la traduction [20].

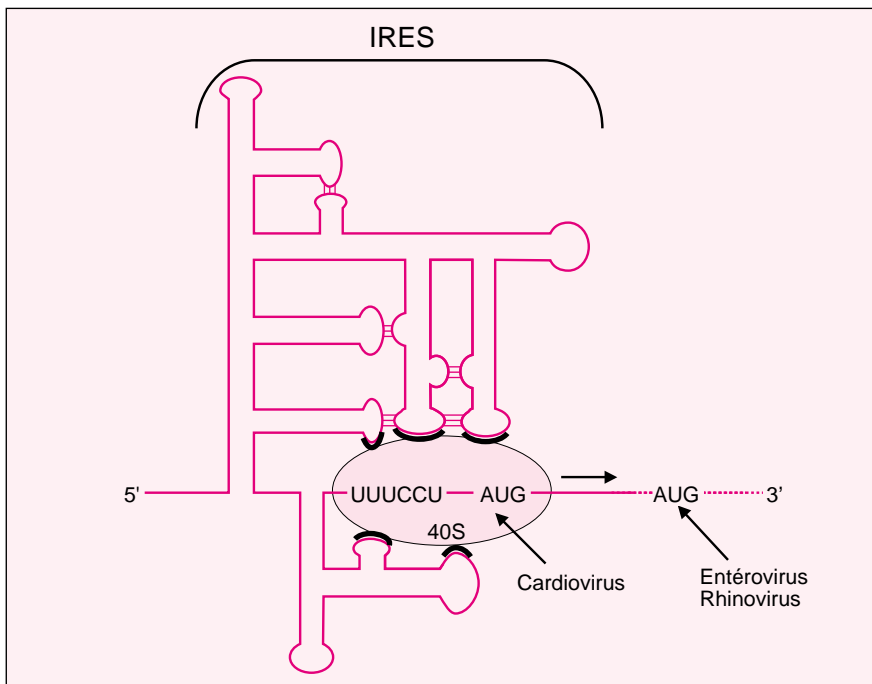


Figure 4. Représentation schématique du mécanisme d'initiation interne de la traduction des ARNm des Picornaviridae. L'IRES (internal ribosome entry segment) est caractérisé par une structure secondaire, composée de nombreuses tiges et boucles, et la formation éventuelle de structures tertiaires par interaction entre les renflements des boucles. La fonction principale d'une telle structure est de réunir des séquences primaires (figurées par un trait épais) et de les maintenir dans une organisation spatiale appropriée. En effet, ces motifs sont probablement impliqués dans la fixation de la sous-unité 40S et dans le recrutement des facteurs d'initiation. La fixation de la sous-unité 40S a lieu en 3' de l'IRES au niveau d'un codon AUG, à proximité d'une séquence riche en bases pyrimidiques. Ce codon est le codon d'initiation authentique pour la synthèse de la polyprotéine dans le cas des cardiovirus. Chez les entérovirus et les rhinovirus, il sert uniquement de site d'entrée pour la sous-unité 40S et n'est pas fonctionnel en tant que codon d'initiation de la traduction; la sous-unité 40S se déplace le long de l'ARNm jusqu'au véritable codon d'initiation où la sous-unité 60S va s'associer (d'après [16]).

Les éléments régulateurs actifs en cis et les protéines à effet en trans

La fixation du complexe de pré-initiation 43S sur l'ARNm et le déplacement de la sous-unité 40S le long de la 5'UTR peuvent être modulés par des interactions entre des séquences de réponse spécifiques et des protéines régulatrices de l'initiation de la traduction, qui sont généralement des facteurs répresseurs, tels que le facteur IRP1 (*iron responsive protein-1*) [21]. En l'absence de fer dans la cellule, le facteur IRP1 bloque l'initiation de la traduction au niveau de la coiffe de l'ARNm de la ferritine en se fixant sur une séquence très structurée, appelée IRE (*iron responsive element*), située à 40 nucléotides de la

coiffe. Par conséquent, la ferritine n'est synthétisée que dans les cellules correctement approvisionnées en fer, et c'est sa concentration qui règle la conversion entre les formes de haute et de faible affinité du facteur IRP1 pour l'IRE. A l'inverse, dans les cellules ravitaillées en fer, le facteur IRP2 est rapidement dégradé par le protéasome. Des études récentes montrent que le facteur répresseur peut être le produit protéique du transcrite lui-même. Ainsi, par exemple, la PABP inhibe sa propre synthèse en se fixant sur une séquence riche en résidus A, qui environne le codon d'initiation de ses propres ARNm, en exerçant ainsi un rétrocontrôle négatif [22]. L'inhi-

bition cesse lorsque le nombre de messagers polyadénylés ou la longueur des queues poly(A) existantes augmente. Des résultats récents indiquent que la traduction des ARNm contenant à l'extrémité 5' une séquence oligopyrimidique, appelée 5'TOP mRNA (*5' terminal oligopyrimidine tract mRNA*), dépend de la croissance cellulaire [23]. Cette famille de 5'TOP mRNA comprend, entre autres, les messagers codant pour des facteurs de l'appareil de traduction, tels que des protéines ribosomiques ou des facteurs d'élongation, mais également le messenger codant pour la PABP.

Le contexte nucléotidique du codon d'initiation

L'initiation de la traduction par balayage se fait dans 95 % des cas au niveau du premier codon AUG rencontré par le complexe d'initiation 48S. Dans un travail de référence, Kozak a étudié l'influence des nucléotides environnants sur le choix du codon AUG d'initiation [24]. Elle a défini une séquence consensus qui révèle le rôle majeur d'une base purique en position -3 et d'une guanine en position +4 dans la séquence 5'-GCCPCCAGG-3' dans lequel la base A du codon AUG est le résidu +1 de traduction.

La position du codon d'initiation

La présence de codons AUG, appelés uAUG (*upstream AUG*), en amont du véritable codon d'initiation, et pouvant être associés à de petits cadres ouverts de lecture, appelés uORF (*upstream ORF*), influence l'efficacité de l'initiation de la traduction [6]. Cette régulation concerne environ 10 % des transcrits eucaryotes. Ainsi, si le complexe d'initiation 48S rencontre un codon AUG situé en amont du véritable codon d'initiation mais dans un environnement peu favorable à l'initiation de la traduction, il ne l'utilisera qu'avec une fréquence très faible et l'on parlera d'un phénomène de « fuite ». En revanche, on observe un phénomène de ré-initiation interne de la traduction lorsque le complexe d'initiation 48S reconnaît, en amont du véritable codon d'initiation, un ou plusieurs codons AUG dans un contexte très favorable à l'initiation de la traduction. Ces codons AUG sont souvent

suivis de très près par un codon de terminaison et on observe alors la synthèse d'un peptide court. Après la traduction de ce minicistron, quelques ribosomes restent associés à l'ARNm, continuent de balayer et ré-initient la traduction au niveau du véritable codon d'initiation.

La longueur de la 5'UTR

La distance entre la coiffe et le véritable codon d'initiation semble jouer un rôle sur l'efficacité de l'initiation de la traduction. En effet, la longueur minimale requise pour permettre la reconnaissance du premier codon AUG, est d'environ 20 nucléotides, mais elle peut être compensée par la présence de structures secondaires en aval du codon d'initiation [25]. Au-delà de 20 nucléotides, l'efficacité de l'initiation de la traduction augmente avec la taille. Cette observation n'est attribuable qu'à la longueur de la 5'UTR et non pas à la présence d'une séquence particulière. La seule caractéristique de cette séquence est la pauvreté en résidus G, qui n'est pas favorable à la formation de structures secondaires. La capacité d'une longue 5'UTR à accroître l'efficacité de l'initiation de la traduction pourrait s'expliquer en partie par l'accumulation de nombreuses sous-unités 40S sur ces ARNm, qui leur confèrent un avantage pour la traduction.

La 3'UTR

Au cours de ces dernières années, le rôle de la 3'UTR dans la régulation de l'initiation de la traduction des ARNm a fait l'objet de nombreux travaux [6, 26]. Plusieurs études ont démontré que cette région est impliquée d'une part dans la fixation des facteurs eIF4F, eIF4A et eIF4B sur l'extrémité 5' [27, 28], et d'autre part dans la ré-initiation de la traduction en permettant le recyclage des sous-unités 60S en direction de l'extrémité 5' du même ARNm (figure 5) [29, 30]. Par conséquent, la 3'UTR participe à l'optimisation de l'initiation de la traduction. Ces travaux ont également montré que ces effets dépendaient de la taille de la queue poly(A) et de la présence de la coiffe. L'action potentiellement synergique de ces deux éléments impliquerait leur reconnaissance mutuelle [31].

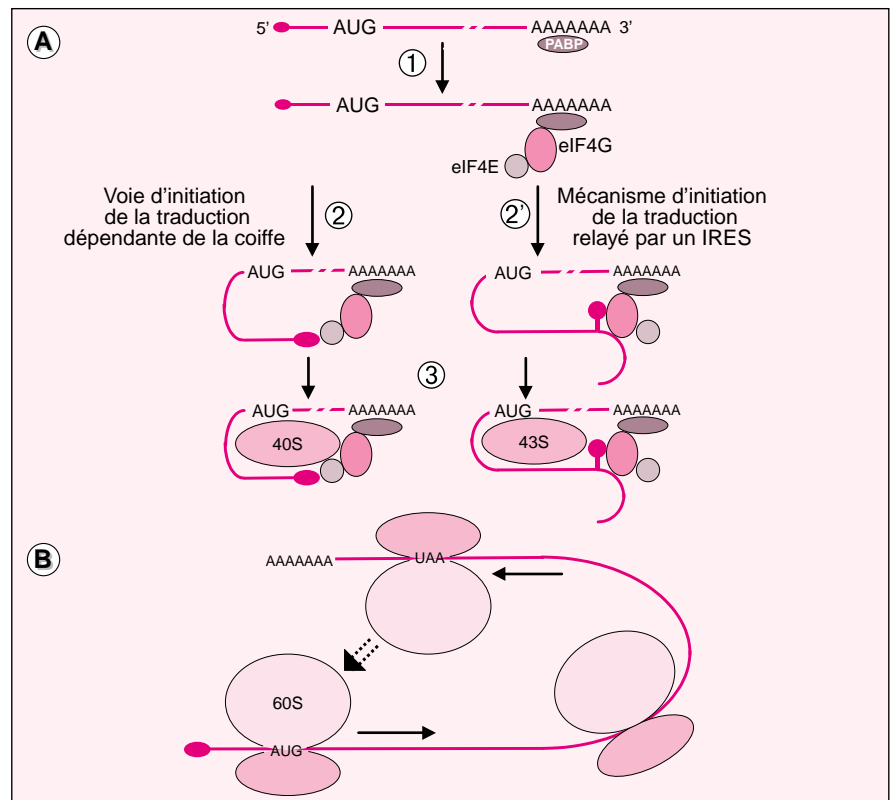


Figure 5. **Représentation schématique de l'interaction entre les extrémités 5' et 3' des ARNm lors de l'initiation de la traduction.** (d'après [26]). La queue poly(A) stimule l'initiation de la traduction (A) d'une part en participant au recrutement de la sous-unité 40S via des interactions entre la PABP et le facteur eIF4G: (1) la PABP liée à la queue poly(A) interagit avec le complexe eIF4G/eIF4E, (2) le facteur eIF4E en se fixant à la coiffe ou (2') le facteur eIF4G en se fixant sur l'IRES circularise alors l'ARNm. (3) La capture du complexe de pré-initiation 43S serait alors facilitée. (B) D'autre part, l'extrémité 3' favoriserait le recyclage des sous-unités 60S en direction de l'extrémité 5'.

Les deux extrémités d'un ARNm, en particulier la queue poly(A) et la coiffe, entreraient donc en contact, probablement de façon indirecte, par l'intermédiaire de la PABP et des facteurs généraux d'initiation [32], mais aussi certainement d'autres protéines non encore identifiées se fixant au niveau des 3' et 5' UTR. Ainsi, la coopération entre les extrémités 3' et 5' d'un ARNm au cours de sa traduction (*cross-talk*) aboutirait à une pseudo-circularisation transitoire de l'ARNm (*closed-loop model*) facilitant la ré-initiation [6]. Une telle structure de l'ARNm suggère donc que la traduction s'effectue de manière continue selon un processus en « cercle roulant ». De plus, les extrémités 3' et 5' se protégeraient mutuellement contre la dégradation par les exoribonucléases. Un exemple illustrant cette dernière

hypothèse est fourni par les ARNm non-sens introduits à la suite de mutations, qui contiennent des codons de terminaison de la traduction prématurés [33, 34]. La traduction de ces ARNm est ré-initiée de manière interne avec participation de la queue poly(A) et de la PABP; la coiffe n'étant plus protégée, elle est éliminée et les ARNm sont dégradés par les exoribonucléases dans le sens 5' → 3' (*RNA surveillance* ou *non-sens mediated decay*).

Les éléments régulateurs actifs en cis et les protéines à effet en trans

Des facteurs répresseurs se fixant à la 3'UTR peuvent bloquer la traduction de plusieurs ARNm. La mutation de leurs sites de fixation sur la 3'UTR est suffisante pour annuler cet effet inhibiteur. Bien que la plupart des exemples connus portent sur la régu-

lation de l'initiation de la traduction des ARNm au cours du développement [6], quelques-uns concernent les cellules somatiques. C'est notamment le cas de l'ARNm qui code pour la 15-lipoxygénase, une protéine synthétisée par les érythrocytes [35, 36]. Cet ARNm est synthétisé dans les précurseurs érythroïdes de la moelle osseuse, mais il reste stocké sous une forme non traduisible jusqu'à la maturation des réticulocytes en érythrocytes. L'inhibition de l'initiation de la traduction de cet ARNm pendant les stades précoces de l'érythropoïèse est due à une protéine de 48 kDa, la LOXBP (*lipoxygenase binding protein*), qui se fixe sur une séquence de la 3'UTR, dans laquelle se trouve un motif de 19 nucléotides, riche en bases pyrimidiques, et répliqué 10 fois. Ce blocage de l'initiation de la traduction s'effectue indépendamment de la taille de la queue poly(A) et de la présence de la coiffe. On ignore encore comment est assurée la levée de l'inhibition. Il existe d'autres exemples du rôle de motifs riches en résidus A et U, impliqués dans la répression de l'initiation de la traduction des ARNm, en particulier dans le cas des cytokines inflammatoires, telles que l'interféron β et le TNF α [37, 38]. Nous reparlerons ultérieurement de ces motifs qui sont généralement considérés comme des éléments d'instabilité susceptibles de favoriser la dégradation des ARNm porteurs d'une ou plusieurs copies de cette séquence dans leur 3'UTR. La virologie nous fournit également un exemple avec le parvovirus B19 [39]. Ce virus présente un tropisme limité aux précurseurs érythroïdes, ou BFU-E (*burst-forming unit of erythroid cells*) et aux CFU-E (*colony-forming unit of erythroid cells*), principales cellules permettant la propagation virale. Cependant, étant donné la distribution tissulaire assez large de son récepteur, constitué de l'antigène de groupe sanguin P ou globoside, le virus B19 est capable de pénétrer dans de nombreuses cellules différentes, sans pouvoir nécessairement s'y répliquer. Dans ces cellules semi-permissives, seule la protéine non structurale 1 est synthétisée, et la lyse cellulaire survient probablement en raison de l'accumulation de cette protéine cytotoxique. En revanche,

les deux protéines de capsid VP1 et VP2 (*viral protein 1* et *2*) ne sont pas détectables alors que leurs ARNm sont présents. Il semblerait que les ARNm viraux qui codent pour ces protéines de capsid ne sont pas traduits, car la 3'UTR de ces ARNm réprime la traduction des ARNm de VP1 et VP2 en inhibant le chargement en ribosomes. La ou les séquences nucléotidiques impliquées dans l'inhibition de la traduction de ces ARNm n'ont pas encore été localisées, pas plus que les protéines cellulaires susceptibles d'interagir avec ces régions identifiées. Il est possible que des facteurs spécifiques des précurseurs érythroïdes soient indispensables à la traduction des ARNm de VP1 et VP2, et ne permettent la synthèse des protéines de capsid que dans les BFU-E et les CFU-E. On constate donc que les 3'UTR jouent un rôle particulièrement important dans la régulation post-transcriptionnelle de l'expression génique. Leur effet n'est pas limité à l'initiation de la traduction d'un ARNm, puisqu'ils semblent également intervenir lors de l'étape de l'élongation. En effet, il a été montré que la fixation de la SBP (*secis binding protein*) sur la SECIS (*selenocysteine insertion sequence*) de la 3'UTR de l'ARNm de la glutathion peroxydase, permet l'incorporation de la sélénocystéine, un acide aminé modifié, au cours de la synthèse de la chaîne polypeptidique, en assurant la reconnaissance du codon opale UGA par le sélénocystéyl ARNt [40].

La queue poly(A)

La présence et la taille de la queue poly(A) contribuent à l'efficacité de l'initiation de la traduction des ARNm. La longueur de la séquence poly(A) des ARNm cytoplasmiques varie, et elle est régie par un phénomène dynamique qui dépend de séquences spécifiques de l'ARNm ainsi que du type cellulaire, ceci à l'inverse des événements nucléaires. La taille de la queue poly(A) de l'ARNm est progressivement réduite au cours du temps, mais à une vitesse qui varie selon l'ARNm et le type cellulaire considérés, et l'efficacité d'initiation de la traduction de ces «vieux» ARNm diminue. Ces réactions de désadénylation, sur lesquelles nous reviendrons ultérieurement, précèdent la dégradation des

ARNm. Dans certaines circonstances, des réactions de polyadénylation peuvent également se produire dans le cytoplasme, et il s'agit probablement là du mécanisme le plus important de la régulation de l'expression des gènes au cours des premiers stades du développement [6]. Dans l'ovocyte de xenope, la plupart des ARNm ont une queue poly(A) de taille normale, qu'ils perdent après la traduction. Cependant, d'autres ARNm, stockés sous une forme non traduisible, ont exceptionnellement une petite queue poly(A). Cette queue oligo(A) provient d'une queue poly(A) initialement synthétisée au niveau du noyau, puis raccourcie ultérieurement. Au cours de la maturation de l'ovocyte, cette queue est rallongée dans le cytoplasme au moment où les ARNm deviennent traduisibles et sont recrutés par l'appareil de traduction [41]. Ces réactions de ré-adénylation cytoplasmiques nécessitent elles aussi le signal AAUAAA, le CPSF et la PAP, comme cela est le cas dans le noyau. En outre, elles nécessitent une séquence en amont riche en résidus U, appelée CPE (*cytoplasmic polyadenylation element*), sur laquelle se fixe la CPEB (*CPE binding protein*). Le mécanisme de polyadénylation cytoplasmique n'est cependant pas encore clairement établi. Il a également été étudié dans les cellules somatiques dans lesquelles la LOXBP, déjà impliquée dans l'inhibition de l'initiation de la traduction de l'ARNm de la 15-lipoxygénase, est capable d'inhiber la polyadénylation cytoplasmique en se fixant soit sur le CPE, soit sur le signal AAUAAA [35, 36].

UTR et stabilité des ARNm

L'abondance des ARNm cytoplasmiques est déterminée par le taux de transcription et la vitesse de dégradation des ARNm, cette dernière étant le principal mécanisme de régulation dans tous les organismes vivants, des bactéries aux mammifères. Chez les eucaryotes supérieurs, le contrôle de la stabilité des ARNm permet à la cellule d'ajuster très rapidement et à tout moment l'expression de ses gènes lors de la croissance, de la différenciation ou des modifications de l'environnement (*pour revue, voir* [42]).

On sait depuis longtemps que la présence de la coiffe s'oppose à la digestion des ARNm par les exoribonucléases dans le sens 5' → 3' [43, 44]. Deux mécanismes décrits plus récemment indiquent que la queue poly(A) protège également les ARNm contre la dégradation. Le premier est la désadénylation, étape première et limitante qui précède la destruction des ARNm, et le second est l'action de la PABP, qui, en interagissant avec la séquence poly(A), forme un complexe qui protège les ARNm polyadénylés de la cellule contre leur digestion par les exoribonucléases, dans le sens 3' → 5'. En outre, chez la levure, la désadénylation conduit à l'élimination de la coiffe et aboutit également à la dégradation par les

exoribonucléases dans le sens 5' → 3' [45, 46]. Ce mécanisme pourrait également s'appliquer chez les eucaryotes supérieurs.

La plupart des éléments qui contrôlent la stabilité des ARNm ont été localisés dans les 3'UTR (figure 6) [6, 47]. C'est le cas des séquences qui stimulent une réaction rapide de désadénylation et favorisent donc la dégradation des ARNm par les exoribonucléases. Ces séquences riches en résidus adénylate et uridylate (ARE ou *AU rich element*), ont été retrouvées dans les 3'UTR de nombreux ARNm labiles qui codent pour des proto-oncoprotéines (*c-myc*, *c-fos*, *c-jun*) et des cytokines (GM-CSF, IL-3). Devant la diversité des séquences impliquées, aucune séquence

consensus n'a pu être rigoureusement définie, mais une classification a été proposée, divisant les ARE en deux groupes [48, 49]. Le premier groupe est constitué par les séquences renfermant une ou plusieurs copies du motif pentamérique AUUUA (classe I) ou tétramérique AUUU (classe II), entouré de diverses répétitions de A et de U, et couplées à une longue région continue riche en U. Le second groupe consiste en séquences divisées en trois domaines dont un est formé par environ 60 % de A ou de U et un autre par trois copies du motif pentamérique GUUUG. De nombreux complexes protéiques (AUBP ou *AU binding protein*), possédant la capacité de se fixer sur ces ARE ont été mis en évidence. Cependant, la seule protéine isolée à ce jour est l'AUF1 (*AU factor 1*), dont la fixation induit *in vitro* la déstabilisation des ARNm de *c-myc* et du GM-CSF [50]. La fixation d'autres protéines non encore identifiées semble plutôt être corrélée à une stabilisation de ces ARNm.

Les 3'UTR qui contiennent des sites spécifiques de clivage par des endoribonucléases peuvent affecter la demi-vie des ARNm. Ces sites peuvent être représentés soit par des structures secondaires en tige et boucle, comme cela est le cas pour l'ARNm du récepteur de la transferrine [51], soit par des séquences nucléotidiques particulières comme cela est le cas pour l'ARNm de la sous-unité R2 de la ribonucléotide réductase [52]. Dans ces différents cas, la stabilité des ARNm semble alors être assurée par un facteur protéique qui, en se fixant au niveau ou à proximité du site de clivage, masque ce dernier, et protège ainsi l'ARNm de l'action des endonucléases [53, 54].

Outre ces deux mécanismes principaux de déstabilisation des ARNm par leur 3'UTR, les ARNm codant pour les histones, qui ne sont pas polyadénylés, contiennent dans leur 3'UTR des structures secondaires impliquées dans leur dégradation par des exoribonucléases, qui se produit en fin de phase S du cycle cellulaire [55]. L'élément *cis* responsable de cette régulation est formé d'une tige et d'une boucle, précédée d'une séquence riche en bases puriques. L'hypothèse avancée est que le facteur stabilisateur SLBP (*stem-loop bin-*

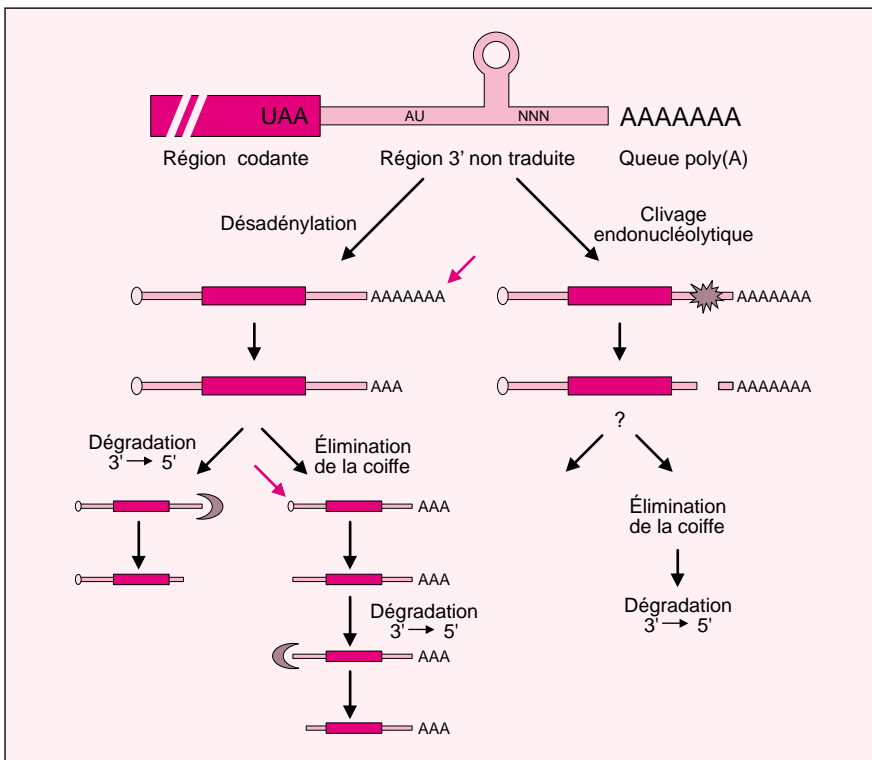


Figure 6. Rôle de la 3'UTR dans la régulation de la stabilité des ARNm. À gauche, la voie de dégradation des ARNm commence par une réaction de désadénylation et se poursuit par la digestion par les exoribonucléases dans le sens 3' → 5'. Le raccourcissement de la queue poly(A) pourrait probablement aussi conduire à l'élimination de la coiffe et aboutirait à la digestion des ARNm par les exoribonucléases dans le sens 5' → 3'. Ce mécanisme est essentiellement relayé par des ARE, situées dans la 3'UTR, et riches en résidus adénylate et uridylate, sur lesquelles des complexes protéiques pourraient se fixer et induire la déstabilisation des ARNm. À droite, la voie de dégradation démarre après clivage par des endoribonucléases dont les sites spécifiques sont représentés soit par des structures secondaires en tige et boucle, soit par des séquences nucléotidiques particulières. Les ARNm seraient ensuite digérés de la même manière que dans la voie précédente (d'après [47]).

ding protein) agissant en *trans* s'opposerait à la dégradation de ces ARNm pendant la division cellulaire, en se fixant sur cette séquence cible [56]. D'autre part il a été récemment rapporté que deux ARNm de 4,2 et 1,6 kb, résultant de la sélection alternative de quatre sites de polyadénylation, sont capables de coder pour le facteur eIF2 [57]. Ces deux ARNm sont traduits avec la même efficacité, mais, dans les lymphocytes T activés, le messenger de 4,2 kb est plus stable que celui de 1,6 kb. La sélection du site de polyadénylation et la stabilité de ces ARNm sont probablement contrôlées par des éléments *cis* contenus dans la 3'UTR, mais aucune séquence n'a encore été identifiée.

Conclusions

Au cours de cette dernière dizaine d'années, la découverte de séquences régulatrices dans les régions 5' et 3' non traduites des ARNm et de facteurs capables d'agir en *trans* a permis de mieux appréhender la régulation de la traduction. Les mécanismes décrits reposent principalement sur des interactions entre des protéines régulatrices, jouant le rôle d'activateurs ou de répresseurs et les régions non traduites souvent très structurées. Les séquences ou les structures secondaires des UTR impliquées interviennent au niveau de la régulation de la traduction, et contrôlent également le transport nucléo-cytoplasmique, la localisation cytoplasmique et la stabilité des ARNm. La détermination de la nature des interactions entre les régions 5' et 3' non traduites, et en particulier entre la coiffe et la queue poly(A), ainsi que les facteurs impliqués dans ces interactions, devrait se révéler essentielle pour comprendre les mécanismes par lesquels les régions non traduites contrôlent l'expression génique ■

* GLOSSAIRE *

UTR : untranslated region.
ORF : open reading frame.
PABP : poly(A) binding protein.
RNP : ribonucleoprotein.
eIF : eucaryotic initiation factor.
ARE : AU rich element.

Remerciements

L'auteur tient à exprimer toute sa gratitude à Monsieur le Professeur Frédéric Morinet pour avoir encadré ce travail réalisé dans le cadre d'une Thèse de Doctorat et à Monsieur le Professeur Jean-Jacques Madjar pour avoir permis de réaliser une expérience décisive de ce travail. Elle remercie également Monsieur le Docteur Ali Saib et Madame le Docteur Anna Greco pour la lecture minutieuse de ce manuscrit. Ce travail a bénéficié du support financier de la FRM, de l'AP-HP/Cnrs et de l'Inserm.

RÉFÉRENCES

1. Varani G. A cap for all occasions. *Structure* 1997; 5: 855-8.
2. Lewis JD, Izaurralde E. The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem* 1997; 247: 461-9.
3. Wahle E, Kuhn U. The mechanism of 3' cleavage and polyadenylation of eukaryotic pre-mRNA. *Prog Nucleic Acids Res Mol Biol* 1997; 57: 41-71.
4. Izaurralde E, Mattaj IW. RNA export. *Cell* 1995; 81: 153-9.
5. Nakielny S, Dreyfuss G. Nuclear export of proteins and RNAs. *Curr Opin Cell Biol* 1997; 9: 420-9.
6. Hershey J, Mathews M, Sonenberg N. *Translational control*. New York: Cold Spring Harbor Laboratory Press, 1996: 794.
7. Wilhelm JE, Vale RD. RNA on the move: the mRNA localization pathway. *J Cell Biol* 1993; 123: 269-74.
8. St Johnston D. The intracellular localization of messenger RNAs. *Cell* 1995; 81: 161-70.
9. King ML. Molecular basis for cytoplasmic localization. *Dev Genet* 1996; 19: 183-9.
10. Singer RH. RNA zipcodes for cytoplasmic addresses. *Curr Biol* 1993; 3: 719-21.
11. Grunert S, St Johnston D. RNA localization and the development of asymmetry during *Drosophila* oogenesis. *Curr Opin Genet Dev* 1996; 6: 395-402.
12. Bassell G, Singer RH. mRNA and cytoskeletal filaments. *Curr Opin Cell Biol* 1997; 9: 109-15.
13. Hesketh J. Translation and the cytoskeleton: a mechanism for targeted protein synthesis. *Mol Biol Rep* 1994; 19: 233-43.
14. Hershey JW. Translational control in mammalian cells. *Annu Rev Biochem* 1991; 60: 717-55.
15. Pain VM. Initiation of protein synthesis in eukaryotic cells. *Eur J Biochem* 1996; 236: 747-71.
16. Jackson RJ, Hunt SL, Gibbs CL, Kaminski A. Internal initiation of translation of picornavirus RNAs. *Mol Biol Rep* 1994; 19: 147-59.
17. Shatkin AJ. mRNA cap binding proteins: essential factors for initiating translation. *Cell* 1985; 40: 223-4.
18. Gray NK, Hentze MW. Regulation of protein synthesis by mRNA structure. *Mol Biol Rep* 1994; 19: 195-200.
19. Manzella JM, Blackshear PJ. Regulation of rat ornithine decarboxylase mRNA translation by its 5'-untranslated region. *J Biol Chem* 1990; 265: 11817-22.
20. Ito K, Kashiwagi K, Watanabe S, Kameji T, Hayashi S, Igarashi K. Influence of the 5'-untranslated region of ornithine decarboxylase mRNA and spermidine on ornithine decarboxylase synthesis. *J Biol Chem* 1990; 265: 13036-41.
21. Rouault T, Klausner R. Regulation of iron metabolism in eukaryotes. *Curr Top Cell Regul* 1997; 35: 1-19.
22. Hornstein E, Harel H, Levy G, Meyuhos O. Overexpression of poly(A)-binding protein down-regulates the translation or the abundance of its own mRNA. *FEBS Lett* 1999; 457: 209-13.
23. Hornstein E, Git A, Braunstein I, Avni D, Meyuhos O. The expression of poly(A)-binding protein gene is translationally regulated in a growth-dependent fashion through a 5'-terminal oligopyrimidine tract motif. *J Biol Chem* 1999; 274: 1708-14.
24. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 1986; 44: 283-92.
25. Kozak M. A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expr* 1991; 1: 111-5.
26. Sachs AB, Sarnow P, Hentze MW. Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* 1997; 89: 831-8.
27. Gallie DR, Tanguay R. Poly(A) binds to initiation factors and increases cap-dependent translation *in vitro*. *J Biol Chem* 1994; 269: 17166-73.
28. Kessler SH, Sachs AB. RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. *Mol Cell Biol* 1998; 18: 51-7.
29. Galili G, Kawata EE, Smith LD, Larkins BA. Role of the 3'-poly(A) sequence in translational regulation of mRNAs in *Xenopus laevis* oocytes. *J Biol Chem* 1988; 263: 5764-70.
30. Munroe D, Jacobson A. mRNA poly(A) tail, a 3' enhancer of translational initiation. *Mol Cell Biol* 1990; 10: 3441-55.
31. Iizuka N, Najita L, Franzusoff A, Sarnow P. Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from *Saccharomyces cerevisiae*. *Mol Cell Biol* 1994; 14: 7322-30.

RÉFÉRENCES

32. Tarun SZ, Jr., Sachs AB. Association of the yeast poly(A) tail-binding protein with translation initiation factor eIF4G. *EMBO J* 1996; 15: 7168-77.
33. Maquat LE. When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *RNA* 1995; 1: 453-65.
34. Ruiz-Echevarria MJ, Gonzalez CI, Peltz SW. Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA. *EMBO J* 1998; 17: 575-89.
35. Ostareck-Lederer A, Ostareck DH, Standart N, Thiele BJ. Translation of 15-lipoxygenase mRNA is inhibited by a protein that binds to a repeated sequence in the 3'-untranslated region. *EMBO J* 1994; 13: 1476-1481.
36. Ostareck DH, Ostareck-Lederer A, Wilm M, Thiele BJ, Mann M, Hentze MW. mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-lipoxygenase translation from the 3'-end. *Cell* 1997; 89: 597-606.
37. Krays VI, Wathélet MG, Huez GA. Identification of a translation inhibitory element (TIE) in the 3'-untranslated region of the human interferon- β mRNA. *Gene* 1988; 72: 191-200.
38. Krays V, Huez G. Translational control of cytokine expression by 3' UA-rich sequences. *Biochimie* 1994; 76: 862-6.
39. Pallier C, Greco A, Le Junter J, Saib A, Vassias I, Morinet F. The 3'-untranslated region of the B19 parvovirus capsid protein mRNAs inhibits its own mRNA translation in nonpermissive cells. *J Virol* 1997; 71: 9482-9.
40. Shen Q, Wu R, Leonard JL, Newburger PE. Identification and molecular cloning of a human selenocysteine insertion sequence-binding protein. A bifunctional role for DNA-binding protein b. *J Biol Chem* 1998; 273: 5443-6.
41. Meric F, Searfoss AM, Wormington M, Wolffe AP. Masking and unmasking maternal mRNA. The role of polyadenylation, transcription, splicing and nuclear history. *J Biol Chem* 1996; 271: 30804-10.
42. Ross J. Control of messenger RNA stability in higher eukaryotes. *Trends Genet* 1996; 12: 171-5.
43. Furuichi Y, LaFiandra A, Shatkin AJ. 5'-terminal structure and mRNA stability. *Nature* 1977; 266: 235-9.
44. LaGrandeur TE, Parker R. mRNA decapping activities and their biological roles. *Biochimie* 1996; 78: 1049-55.
45. Muhlrud D, Decker CJ, Parker R. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-3' digestion of the transcript. *Genes Dev* 1994; 8: 855-66.
46. Beelman CA, Stevens A, Caponigro G, et al. An essential component of the decapping enzyme required for normal rates of mRNA turnover. *Nature* 1996; 382: 642-6.
47. Beelman CA, Parker R. Degradation of mRNA in eukaryotes. *Cell* 1995; 81: 179-83.
48. Chen CY, Shyu AB. AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem Sci* 1995; 20: 465-70.
49. Peng SS, Chen CY, Shyu AB. Functional characterization of a non-AUUUA AU-rich element from the c-jun proto-oncogene mRNA: evidence for a novel class of AU-rich elements. *Mol Cell Biol* 1996; 16: 1490-9.
50. Zhang W, Wagner BJ, Ehrenman K, et al. Purification, characterization and cDNA cloning of an AU-rich element RNA-binding protein, AUF1. *Mol Cell Biol* 1993; 13: 7652-65.
51. Binder R, Horowitz JA, Basilion JP, Koeller DM, Klausner RD, Harford JB. Evidence that the pathway of transferrin receptor mRNA degradation involves an endonucleolytic cleavage within the 3'UTR and does not involve poly(A) tail shortening. *EMBO J* 1994; 13: 1969-80.
52. Amara FM, Chen FY, Wright JA. Defining a novel cis element in the 3'-untranslated region of mammalian ribonucleotide reductase component R2 mRNA: role in transforming growth factor- β 1 induced mRNA stabilization. *Nucleic Acids Res* 1995; 23: 1461-7.
53. Schlegel J, Gegout V, Schlager B, et al. Probing the structure of the regulatory region of human transferrin receptor messenger RNA and its interaction with iron regulatory protein-1. *RNA* 1997; 3: 1159-72.
54. Amara FM, Smith GM, Kuschak TI, Takeuchi TL, Wright JA. A cis-trans interaction at the 3'-untranslated region of ribonucleotide reductase mRNA is regulated by TGF- β 1, TGF- β 2 and TGF- β 3. *Biochem Biophys Res Commun* 1996; 228: 347-51.
55. Caruccio N, Ross J. Purification of a human polyribosome-associated 3' to 5' exoribonuclease. *J Biol Chem* 1994; 269: 31814-21.
56. Wang ZF, Whitfield ML, Ingledue TC, Dominski Z, Marzluff WF. The protein that binds the 3'-end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev* 1996; 10: 3028-40.
57. Miyamoto S, Chiorini JA, Urcelay E, Safer B. Regulation of gene expression for translation initiation factor eIF-2 alpha: importance of the 3' untranslated region. *Biochem J* 1996; 315: 791-8.

Summary

The role of untranslated regions of eukaryotic transcripts in protein synthesis

Gene expression is a multi-step process, which is mostly under the control of the transcripts untranslated regions. Messenger RNA post-transcriptional regulation results from interactions between the 5' and 3' untranslated regions and specific regulatory proteins directing the intracytoplasmic localization, translation and stability or degradation of the mRNA. Thus, mRNA carry *cis*-acting sequences allowing the cell to respond to *trans*-acting factors adapting the rate of translation and/or the stability of mRNA to different physiological situations.

TIRÉS À PART

C. Pallier.