

L'initiation de la synthèse des protéines chez les eucaryotes

**Théophile Ohlmann
Edmund Derrington
Marcelo López-Lastra
Clarence Deffaud
Annabelle Bouchardon
Jean-Luc Darlix**

La traduction des ARN messagers (ARNm), un des mécanismes fondamentaux de la régulation de l'expression des gènes, s'avère très contrôlée, notamment au niveau du déclenchement de la synthèse protéique, qui peut se dérouler selon différents mécanismes. Pour la majorité des ARNm cellulaires, la traduction commence par l'attachement de la petite sous-unité ribosomique (40S) sur la coiffe méthylée en position 5' de l'ARN. La petite sous-unité 40S va ensuite parcourir la région non traduite de l'ARNm de 5' en 3', jusqu'au codon d'initiation, il s'agit du mécanisme de balayage. Cette étape est réglée par les facteurs d'initiation. L'étude d'un groupe de virus à ARN, les picornavirus, a mis en évidence un moyen alternatif d'initiation par lequel la petite sous-unité 40S accède directement à l'ARN en position interne, grâce à un élément génétique la séquence IRES (*internal ribosome entry segment*). Les IRES se sont révélées être à la fois un paradigme pour l'étude de l'initiation de la traduction, et un outil important pour la transgénèse.

Le déclenchement de la synthèse des protéines chez les eucaryotes procède selon trois mécanismes distincts, qui peuvent être complémentaires. Ce sont: (1) le balayage des ribosomes de la coiffe 5' au codon d'initiation, majoritairement un AUG présent dans un contexte de type A/GCCAUGG; (2) le recrutement direct, et en position interne, des ribosomes par un élément génétique appelé IRES (*internal ribosome entry segment*), et (3) le *shunt* (court-circuit) au cours duquel les ribosomes reconnaissent le 5', puis sont transférés au niveau du codon initiateur. La découverte des IRES dans les

ARNm viraux et cellulaires, grâce aux travaux originaux sur les picornavirus, a permis de mettre au point des vecteurs rétroviraux polycistroniques ayant plusieurs IRES. Ces vecteurs dérivés du virus de la leucémie murine, MuLV, et appelés MLV-IRES, possèdent des propriétés remarquables, à savoir des titres viraux très élevés et la possibilité de transférer plusieurs gènes dans un grand nombre de cellules différentes et d'assurer leur expression coordonnée. Ces vecteurs rétroviraux MLV-IRES pourraient donc être utiles pour un transfert de gènes d'intérêt thérapeutique, particulièrement en cancérologie.

ADRESSES

T. Ohlmann, E. Derrington, M. López-Lastra, C. Deffaud, A. Bouchardon, J.L. Darlix : LaboRetro, Unité de virologie humaine, Inserm-ENS U. 412, École normale supérieure, 46, allée d'Italie, 69364 Lyon Cedex, France.

Le mécanisme de balayage et les facteurs d'initiation de la traduction

Le mécanisme de balayage est le processus par lequel le ribosome s'attache à l'ARNm et parcourt la région 5' non traduite (ou 5' *leader*) pour se positionner au niveau du codon d'initiation. Chez les eucaryotes, cette étape est réglée par au moins 9 polypeptides distincts qui sont appelés facteurs d'initiation (eIF pour *eukaryotic initiation factors*).

Formation du complexe 43S

L'initiation de la traduction commence par le recrutement de l'ARNt initiateur-méthionine (Met-ARNt) et d'une molécule de GTP par le facteur d'initiation eIF2, formant ainsi un complexe ternaire eIF2-GTP-Met-ARNt [1]. Ce dernier va alors interagir avec eIF5 et eIF3, préalablement fixés à la sous-unité 40S, pour former le complexe 43S (panneau A de la figure 1) [2, 3]. eIF3 stabilise la sous-unité 40S du ribosome et empêche sa réassociation avec la sous-unité 60S car cela aboutirait à la formation de ribosomes 80S monomères inactifs.

Formation du complexe de pré-initiation 48S

Le complexe 43S se fixe alors à l'ARNm et forme le complexe de pré-initiation (ou 48S), grâce à eIF4F qui joue le rôle d'adaptateur entre la coiffe et la petite sous-unité 40S (figure 1B). L'holoenzyme eIF4F est composé des trois facteurs protéiques eIF4E, eIF4A et eIF4G qui ont chacun une fonction propre. eIF4E est une phosphoprotéine qui reconnaît la coiffe méthylée et s'y attache. La structure tridimensionnelle de l'eIF4E murin [4] et de l'eIF4E de levure [5] a été résolue respectivement par cristallographie et par RMN [5]. Ces études ont révélé une conservation remarquable de la structure tertiaire de ce peptide.

eIF4A est une hélicase à ARN, dépendante de l'ATP, de la famille des protéines *DEAD box* [6]. Les membres de cette famille, qui contiennent 9 régions très conservées, sont impliqués dans plusieurs mécanismes bio-

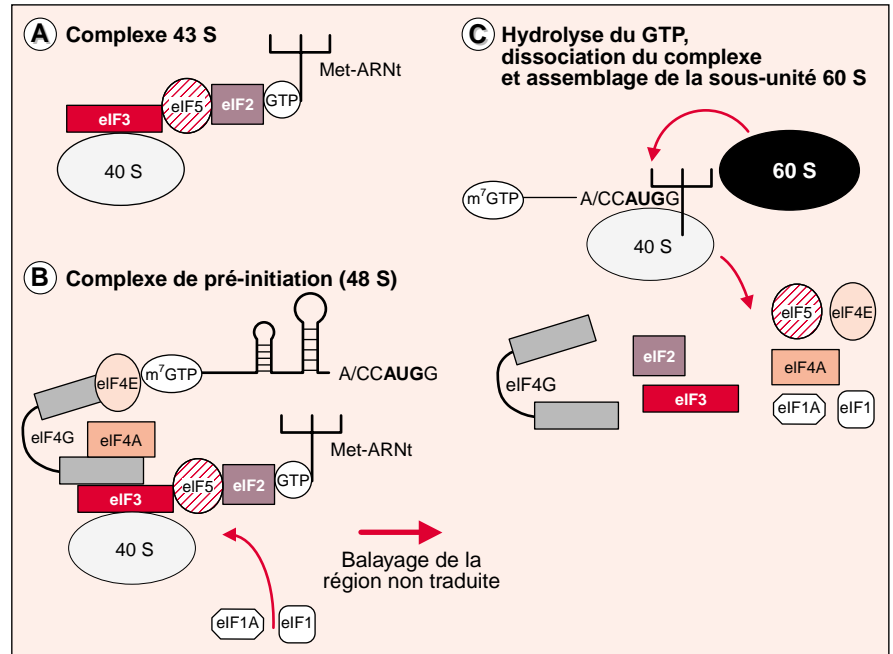


Figure 1. **Assemblage des complexes ribosomiques.** **A.** Le complexe 43S. Le complexe ternaire, formé par le Met-ARNt, eIF2 et une molécule de GTP, est recruté au niveau de la sous-unité 40S. Cela est sous le contrôle du facteur eIF3 qui est fixé à la sous-unité 40S, et par eIF5 qui peut se lier à la fois à eIF3 et à eIF2. **B.** Formation du complexe de pré-initiation 48S. Le facteur multiprotéique eIF4F, qui se compose de 3 protéines distinctes, eIF4E, eIF4A et eIF4G, pilote l'attachement du complexe 43S sur l'ARNm. eIF4E lie spécifiquement la coiffe méthylée en 5' de l'ARNm et eIF4G assure la cohésion de l'ensemble car elle est liée à eIF4E par son extrémité amino-terminale et à eIF3 par son extrémité carboxy-terminale. Le ribosome est alors prêt à balayer la région non traduite. Lors de cette progression, l'activité hélicase du facteur eIF4A déroule les structures secondaires de l'ARN. **C.** Formation du ribosome 80S. Lorsque le complexe de pré-initiation reconnaît le codon AUG d'initiation, eIF5 catalyse l'hydrolyse du GTP, ce qui provoque la dissociation du complexe de pré-initiation. La sous-unité 60S peut alors s'assembler pour former le ribosome 80S.

logiques clés comme l'épissage, la biogenèse du ribosome, et la traduction. Au sein d'eIF4F, le rôle d'eIF4A est de dérouler les structures secondaires d'ARN, proches de la coiffe, qui pourraient empêcher l'attachement du complexe de pré-initiation [7].

Enfin, eIF4G est la protéine « centrale » du complexe eIF4F. Elle contient plusieurs sites de liaison à des eIF, un à eIF4E dans sa partie amino-terminale, un à eIF3 dans sa partie centrale, et deux à eIF4A dans sa partie carboxy-terminale. eIF4G est la protéine qui assure la cohésion du complexe eIF4F, car elle forme un « pont » entre l'ARNm et la sous-unité ribosomique 40S (pour revues, voir [8, 9]).

Balayage de la région 5' non traduite

Le complexe de pré-initiation va migrer de 5' en 3' jusqu'à la rencontre d'un codon d'initiation dans un bon contexte nucléotidique : c'est l'étape de balayage. Lors de la progression du complexe de pré-initiation, l'activité hélicase de eIF4A déroule les structures d'ARN, qui se trouvent sur le trajet du ribosome, en consommant de l'ATP [7]. Cette étape de balayage est sous le contrôle de eIF1 et de eIF1A. En effet, ces deux protéines, qui rejoignent le complexe de pré-initiation au niveau de la coiffe, vont surveiller la progression du ribosome et empêcher qu'il initie à un site aberrant. En cas de mésappariement codon-antico-

don, eIF1 et eIF1A peuvent dissocier le complexe, assurant ainsi la reconnaissance fidèle du codon d'initiation [10] (figure 1B).

Assemblage du ribosome 80S

Au niveau du site d'initiation, le facteur eIF5 hydrolyse le GTP, provoquant la dissociation du complexe de pré-initiation [3]. La sous-unité ribosomique 60S rejoint le complexe de pré-initiation, et le Met-ARNⁱ se place dans le site P du ribosome nouvellement formé. La synthèse peptidique peut alors commencer (figure 1C).

Les contrôles

Le déclenchement de la traduction est contrôlé par des structures de l'ARNm eucaryote (en *cis*) et par l'activité des facteurs d'initiation (en *trans*).

Contrôles en *cis* par des structures de l'ARNm

La majorité des ARNm eucaryotes commence par une guanosine modifiée (méthylée en position 7 : m⁷GTP) – appelée coiffe – qui a plusieurs rôles-clés dans le métabolisme de l'ARN : elle permet l'export nucléocytoplasmique, facilite l'épissage, augmente la stabilité et améliore l'efficacité de la traduction en interagissant spécifiquement avec certains facteurs d'initiation [11].

La longueur et les structures de la région de l'ARNm comprise entre la coiffe et le codon AUG sont des déterminants majeurs de l'efficacité de la traduction. Les ARNm ayant une région 5' non traduite longue, avec un ou plusieurs cadres de lecture ouverts (ORF, *open reading frame*) précédant la région codante, sont mal traduits. Ces caractéristiques se retrouvent généralement dans les ARNm qui codent pour des protéines impliquées dans la prolifération cellulaire, et dont la traduction doit être étroitement réglée (proto-oncogènes, facteurs de croissance, récepteurs hormonaux ou facteurs de transcription) [12]. A l'inverse, si la longueur de la région entre la coiffe et le codon AUG est trop courte (inférieure à 15 nucléotides), la sous-

unité 40S ne reconnaît pas le codon d'initiation, et continue à balayer plus en aval [13]. La propension de la région 5' non traduite à former des appariements d'ARN double-brin de type tige-boucle influence le niveau de « traductibilité » de l'ARN, et la position de ces structures par rapport à la coiffe est déterminante. En effet, une structure secondaire d'ARN, même de stabilité modeste, située dans les 20 premiers nucléotides de l'extrémité 5' du messager peut empêcher l'attachement du ribosome [13]. En revanche, une tige-boucle située 10-15 nucléotides en aval d'un codon initiateur peut exercer un effet positif sur la traduction en ralentissant la progression du ribosome, augmentant la durée du contact codon-anticodon entre le Met-ARNⁱ et l'AUG, ce qui favorise les chances d'initiation à ce site [13]. Cependant, une tige-boucle de stabilité supérieure à -50 kcal/mol est généralement un obstacle infranchissable lors du balayage des ribosomes, et cela quelle que soit sa position sur la région 5' non codante.

L'analyse des séquences nucléotidiques de vertébrés de plus de 699 ARNm, ainsi que des expériences de mutagenèse dirigée ont montré que la séquence consensus GCCA/GCCAUGG était optimale pour la sélection du codon d'initiation [14]. Il est très rare de trouver la totalité de cette séquence, et la présence d'une purine (A ou G) en position -3 et d'un G en position +4 (le A du codon AUG étant +1) suffisent généralement à assurer la reconnaissance fidèle et efficace du site d'initiation. Certains ARNm viraux, et quelques ARNm cellulaires, peuvent également initier la traduction à partir d'autres codons que l'AUG, comme les CUG, GUG, ACG et AUU [15]. Il est à noter que le contexte nucléotidique entourant ces codons « non-AUG » joue également un rôle prépondérant dans la reconnaissance du site d'initiation.

De nombreux travaux ont montré l'importance de la région 3' non traduite de l'ARNm dans la régulation de la traduction, notamment aux stades précoces de l'embryogenèse et du développement. La queue poly (A) et les protéines associées assurent le contrôle de la stabilité ainsi que la

mobilité et la localisation du messager. Récemment, il a été mis en évidence que la région 3' non traduite interagissait directement avec l'extrémité 5' pour le recrutement du complexe de pré-initiation [16]. La *poly(A) binding protein 1* (Pabp 1) se lie au facteur d'initiation eIF4G, ce qui crée un pont 5'-3' et une pseudo-circularisation de l'ARNm [17-19].

Régulations en *trans* par la modulation de l'activité de certains facteurs d'initiation

L'activité de eIF2 est réglée par la phosphorylation de sa sous-unité α qui fait intervenir au moins 3 protéine-kinases distinctes (HCR : *heme control repressor*; GCN2 chez la levure, et PKR, protéine-kinase activée par des régions ARN double-brins). Pour ne citer qu'un seul exemple, lors de l'infection virale, la production d'interférons α et β active la PKR qui phosphoryle eIF2. Cette phosphorylation règle négativement eIF2 qui ne peut plus recycler de molécules de

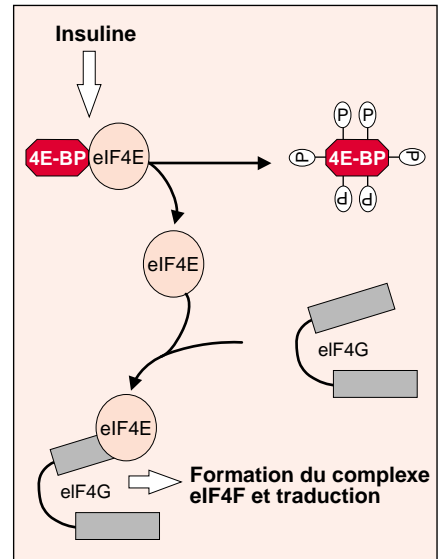


Figure 2. **Modulation de l'activité de eIF4E par le facteur 4E-BP.** Sous sa forme inactive, eIF4E est lié au facteur 4E-BP1 (4E binding protein 1). En réponse à des stimuli extracellulaires, comme la production d'insuline, 4E-BP est phosphorylée et se détache d'eIF4E. Ce dernier peut alors se lier à eIF4G et former le complexe eIF4F nécessaire à la traduction.

Tableau I		
COMPILATION DE SÉQUENCES IRES VIRALES ET CELLULAIRES : ORIGINES ET CARACTÉRISTIQUES PRINCIPALES		
IRES	Origine	Caractéristiques
PV, HRV, FMDV, EMCV, HAV	Virale (picornavirus)	L'extrémité 5' de l'ARNm viral n'est pas coiffée Ce sont les premières séquences IRES identifiées
HCV	Virale (flavivirus)	L'IRES s'étend sur les 10 premiers nucléotides de la région codante
BVDV et CSFV	Virale (pestivirus)	Ces deux IRES permettent un mécanisme d'initiation semblable à celui utilisé par les procaryotes
MLV, HaMSV, HTLV, REV-A, VL30, SIV	Virale (rétrovirus)	Ces IRES ont été identifiées dans l'ARN génomique codant pour gag ou dans l'ARN subgénomique codant pour env (MLV)
Bip	Cellulaire	Premier ARN cellulaire contenant un IRES
<i>Antennapedia</i> et <i>Ultrabithorax</i>	Cellulaire (drosophile)	Activité IRES réglée lors du développement de la drosophile
FGF-2	Cellulaire	Initiation alternative par un mécanisme IRES dépendant de la coiffe
IGF2	Cellulaire	Activité IRES réglée lors du développement
eIF4G	Cellulaire	eIF4G est clivé lors de l'infection par les picornavirus
c-myc, PDGF2/c-sis	Cellulaire	L'IRES pourrait contrôler rigoureusement l'expression de ces deux proto-oncogènes
VEGF	Cellulaire	L'ARNm contient deux IRES dans la région non traduite

Les abréviations utilisées sont les suivantes: poliovirus (PV); foot-and-mouth disease virus (FMDV); encephalomyocarditis virus (EMCV); hepatitis A virus (HAV); hepatitis C virus (HCV); bovine viral diarrhoea virus (BVDV); classical swine fever virus (CSFV); murine leukemia virus (MLV); Harvey murine sarcoma virus (HaMSV); Human T-cell leukaemia virus (HTLV); reticuloendotheliosis virus type A (REV-A); 30S virus-like element (VL30); simian immunodeficiency virus (SIV); immunoglobulin heavy-chain binding protein (Bip); fibroblast growth factor 2 (FGF-2); insulin-like growth factor II (IGF2); eucaryotic initiation factor 4G (eIF4G); platelet-derived growth factor 2 (PDGF2); insulin-like growth factor II (IGF2); Vascular endothelial growth factor (VEGF). À noter que cette compilation n'est pas exhaustive.
Les références des travaux cités peuvent être obtenues auprès des auteurs.

GTP. Ce processus, se faisant au détriment de la cellule, empêche néanmoins l'utilisation de la machinerie traductionnelle par le virus, et constitue une des premières défenses antivirales [20].

L'activité de eIF4E est modulée à deux niveaux: (1) par phosphorylation sur un site unique (Ser 209); (2) par l'activité d'une famille de protéines qui le lient spécifiquement à eIF4E (4E-BP pour *eIF4E binding proteins*). Sous sa forme phosphorylée, eIF4E est co-purifiée préférentiellement avec le complexe eIF4F et présente une affinité plus forte pour la coiffe [21, 22]. La phosphorylation de eIF4E met en jeu deux kinases, la protéine-kinase C (PKC) et la protéine-kinase MNK1 (*MAP*

kinase-interacting protein 1), en réponse à de nombreux stimulus comme des hormones, des facteurs de croissance et des mitogènes [11]. Le niveau de phosphorylation de eIF4E varie également au cours du cycle cellulaire. Pendant les phases G1 et S, le niveau de phosphorylation d'eIF4E est beaucoup plus élevé que lors des phases G0 et M [11]. Ainsi, on retrouve une corrélation entre la forme phosphorylée d'eIF4E et le taux de croissance global de la cellule.

L'activité de eIF4E est également modulée par une famille de protéines, les *4E binding proteins* (4E-BP). Chez les mammifères, cette famille est composée de trois peptides différents (4E-BP1, 4E-BP2,

4E-BP3) dont 4E-BP1 est l'exemple le mieux caractérisé [23, 24]. 4E-BP1 et eIF4G partagent le même site de liaison sur l'extrémité amino-terminale d'eIF4E [25, 26]. De ce fait, la liaison de 4E-BP et d'eIF4G sur eIF4E est mutuellement exclusive. Ainsi, 4E-BP1 agit comme un inhibiteur compétitif de la formation du complexe eIF4F dont l'activité est elle-même réglée par un jeu subtil de phosphorylations [25]. En réponse à la stimulation par l'insuline dans des adipocytes de rat, 4E-BP1 est phosphorylée, ce qui entraîne sa dissociation d'eIF4E. Le site de liaison est alors accessible pour eIF4G, le complexe eIF4F peut se former et la traduction peut commencer (*figure 2*).

Les mécanismes alternatifs de l'initiation de la traduction

Les picornavirus sont des virus à ARN simple brin positif dont le génome code pour une polyprotéine qui est mûrie après synthèse. Au contraire des ARN cellulaires, les ARN des picornavirus ne possèdent pas de coiffe en 5', ont une longue région 5' non traduite, riche en structures secondaires et qui contient de nombreux codons AUG en amont du site d'initiation [27]. L'insertion de la région 5' non traduite des picornavirus entre deux gènes permet l'expression du second gène indépendamment du premier. Ce segment d'ARN d'environ 450 nucléotides offre aux ribosomes un accès à l'ARN de façon directe et interne, et a été appelé IRES (*internal ribosome entry segment*). Le mécanisme par lequel les ribosomes s'attachent à la molécule d'ARN viral est encore mal défini mais les facteurs canoniques de l'initiation de la traduction (à l'exception de eIF4E) et des protéines cellulaires participent à ce processus. Parmi celles-ci, le facteur d'épissage PTB (*polypyrimidine tract binding protein*), l'auto-antigène La, la protéine PCBP2 (*poly(rC) binding protein 2*), et la protéine *unr* (*upstream of N-Ras*) interagissent avec l'IRES, et facilitent l'attachement des ribosomes sur l'ARN [28].

Des IRES ont été identifiés chez tous les membres de la famille des picornavirus et également dans d'autres familles virales, comme les rétrovirus, et dans des ARNm cellulaires (Tableau I), suggérant que l'initiation dépendante d'IRES est un des contrôles de l'expression génique.

L'intérêt d'une séquence IRES

Étant donné la grande efficacité du mécanisme classique d'initiation par balayage, on peut se demander pourquoi les séquences IRES sont apparues. Chez certains picornavirus comme le poliovirus, l'avantage de la séquence IRES apparaît lors de l'infection de la cellule hôte. Le poliovirus code pour une protéase (P2A) qui intervient dans la maturation de la polyprotéine virale, et qui clive également le facteur d'initiation eIF4G en deux parties : amino- et car-

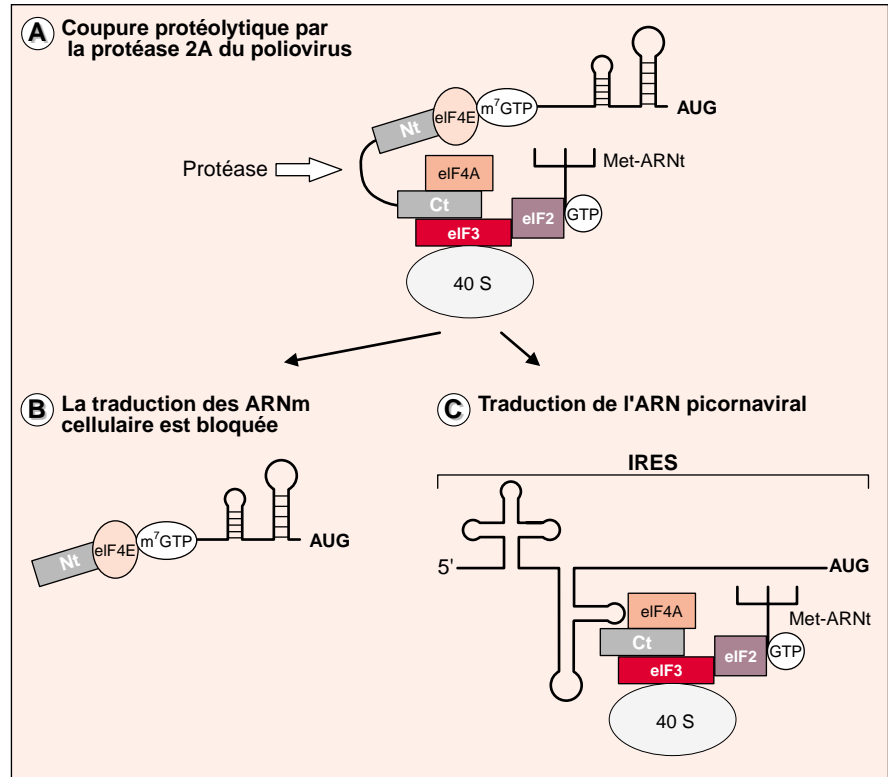


Figure 3. **Coupure du facteur eIF4G par la protéase 2A du poliovirus.** **A.** La protéase 2A clive le facteur d'initiation eIF4G en un site unique. **B.** La partie amino-terminale de eIF4G reste associée à eIF4A, ce qui inhibe la traduction dépendante de la coiffe. **C.** La partie carboxy-terminale de eIF4G reste associée au ribosome et le complexe restant peut être recruté en position interne par une séquence IRES et promouvoir l'expression d'une polyprotéine (cas du poliovirus), ou celle d'un transgène (cas des vecteurs, voir figure 5).

boxy-terminale [29] (figure 3A). Cette coupure protéolytique a plusieurs conséquences. L'ensemble eIF4E-extrémité amino-terminale d'eIF4G reste associé au niveau de la coiffe des messagers cellulaires qui ne peuvent plus lier une sous-unité 40S (figure 3B). En revanche, l'ARN peut recruter le complexe de pré-initiation en position interne grâce à la partie carboxy-terminale d'eIF4G qui reste associée au ribosome [30, 31] (figure 3C). Cela provoque une chute du taux de synthèse protéique cellulaire sans affecter la production des protéines virales qui est sous la dépendance de l'IRES. Un autre picornavirus, l'*Encephalomyocarditis virus* (EMCV), utilise une stratégie visant à détourner la machinerie traductionnelle. L'infection de la cellule hôte par EMCV provoque une hypophosphorylation de 4E-BP qui se lie à eIF4E et empêche la formation d'eIF4F [32] (voir figure 2). En consé-

quence, la traduction dépendante de la coiffe décroît fortement au profit de la traduction déclenchée par l'IRES qui, elle, n'a pas besoin d'eIF4E.

Parmi les ARNm cellulaires, l'exemple de l'ARNm du FGF-2 (*fibroblast growth factor 2*) montre que la synthèse protéique peut être initiée par un mécanisme de balayage classique ou par entrée interne des ribosomes [33]. En réponse à un stress oxydant ou à un choc thermique, l'expression du FGF-2 est assurée de manière prépondérante par entrée interne des ribosomes [34]. En accord avec ces observations, dans des conditions physiologiques particulières comme le choc thermique, ou l'arrêt en métaphase, l'eIF4E est hypophosphorylé et donc l'eIF4F est en partie inactivé [35]. On comprend alors que la présence d'un IRES permet le maintien de l'expression de certaines protéines dont le FGF2 dans des conditions de stress cellulaire.

L'ARNm codant pour le facteur d'initiation eIF4G est lui-même exprimé par entrée interne des ribosomes [36], ce qui assure le maintien de l'expression de eIF4G lors de l'infection par un picornavirus.

Chez des virus à ARN positif, l'extrémité 5' du génome contient des structures secondaires d'ARN très stables qui sont des signaux de réplication. Dans les rétrovirus, par exemple, on trouve une longue région 5' non traduite, structurée et ayant des signaux nécessaires à la synthèse de l'ADN viral. Cette organisation génétique de l'ARN est, de ce fait, incompatible avec un mode d'initiation classique. Il demeure que les protéines et les enzymes virales doivent être synthétisées et la présence d'une séquence IRES dans la région 5' non traduite permet de concilier réplication et traduction de l'ARN génomique.

Le shunt

L'étude de certains virus comme le *cauliflower mosaic virus* [37], le *rice tungro bacilliform virus* [38], le virus Sendaï [39] ainsi que les *late adenovirus mRNAs* [40] a permis de mettre en évidence un troisième mécanisme de déclenchement de la traduction, le *shunt*, qui combine certains aspects du balayage classique et de l'initiation interne. Le complexe de pré-initiation se forme au niveau de la coiffe 5' de l'ARNm puis commence à balayer jusqu'à la rencontre de régions d'ARN structurées qui pilotent son transfert en aval, où il peut atteindre le codon d'initiation par balayage. Comme pour l'initiation interne, ce processus permet de passer outre des structures très stables d'ARN qui ne pourraient pas être déroulées lors du balayage.

Les rétrovirus : un modèle d'étude de la traduction et un outil pour le transfert de gènes

Les rétrovirus sont une famille de virus contenant un génome à ARN de sens positif dont la réplication passe par un intermédiaire ADN qui sera ensuite intégré dans le génome de la cellule hôte. Le provirus intégré va détourner la machinerie cellulaire

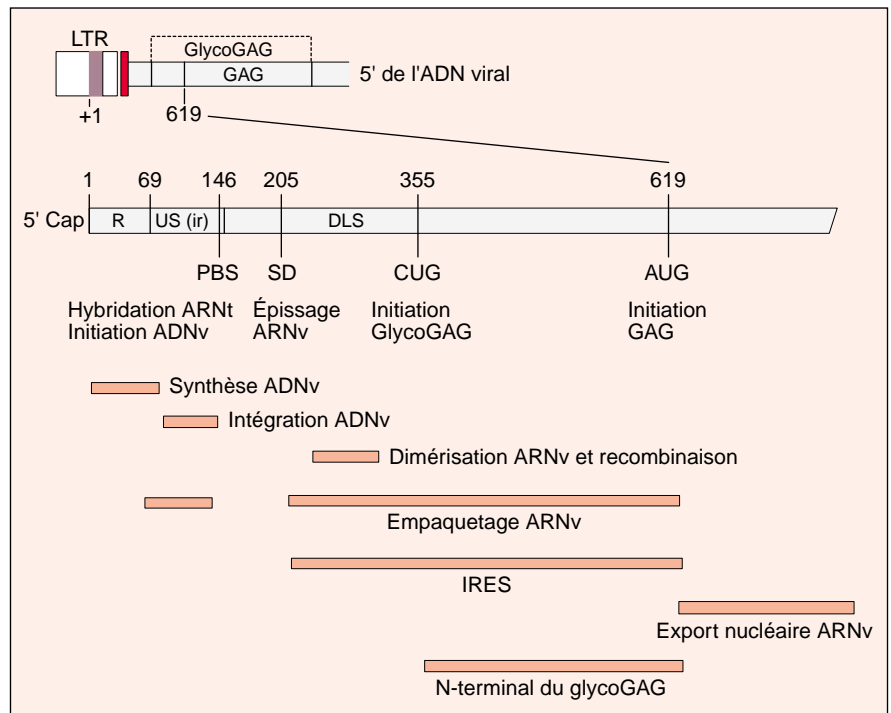


Figure 4. **Structure génétique du 5' leader des rétrovirus de type MuLV.** La région 5' de l'ADN viral est représentée en haut de la figure. Le 5' leader de l'ARN génomique est formé des séquences R (repeat), U5 (région non traduite en 5' avec la séquence inversement répétée ir, nécessaire lors de l'intégration de l'ADN viral), PBS (primer binding site de 18 nt où se fixe l'ARNt, amorce de la transcription inverse; chez les virus MuLV, il s'agit de l'ARNt-Pro), SD (le site d'épissage donneur), DLS (la séquence de dimérisation de l'ARN génomique nécessaire à l'encapsidation du génome dimère dans le virus, et aux événements de recombinaison lors de la synthèse de l'ADN viral). À noter que la séquence DLS est incluse dans le signal d'empaquetage E du génome viral. La séquence IRES recouvre tout ou partie du signal E. La séquence d'export nucléaire est en aval du codon d'initiation de gag. Cette séquence est incluse dans le signal d'encapsidation étendu, appelé E⁺ ou Psir. La région des positions 355 à 619 code pour la partie amino-terminale de la protéine glycoGag, dont l'expression est associée à l'anémie hémolytique et à l'érythroleucémie provoquées par le virus FrMuLV. Les deux codons d'initiation de la synthèse protéique CUG (glycoGag) et AUG (gag) sont dans de très bons contextes.

pour sa transcription, l'épissage et l'export des ARN viraux vers le cytoplasme, et leur traduction. Les protéines virales sont synthétisées dans le cytoplasme de l'hôte pour produire de nouvelles particules virales. L'ARN génomique devra donc être utilisé par les différentes machineries cellulaires au cours de la réplication virale, ce qui se fait grâce à des domaines nucléotidiques, structurés, et localisés en 5' de l'ARN. La région 5' de l'ARN génomique comprend une coiffe, une séquence répétée (R), une région non traduite (U5), le site de fixation de l'ARN de transfert

(PBS), le site d'épissage 5' (SD), la séquence d'empaquetage et de dimérisation (E-DLS) et la séquence d'export nucléaire du génome viral (figure 4). Ce domaine multifonctionnel offre une possibilité de régulation très fine au cours de la traduction et de la réplication du virus. Dans l'hypothèse où le même ARN génomique assume toutes ces fonctions, on peut concevoir que la synthèse des protéines virales a lieu avant la dimérisation de l'ARN, elle-même précédant l'encapsidation. On peut aussi envisager que deux lots d'ARN génomique coexistent dans le

cytoplasme de la cellule infectée. Le premier servirait à la synthèse des protéines virales alors que le second serait destiné à dimériser et à être encapsidé. Quelle que soit l'hypothèse retenue (un ou deux lots d'ARN génomique), les mécanismes régulateurs qui contrôlent le passage d'une fonction à l'autre restent à découvrir.

Les travaux portant sur les oncovirus murins FrMuLV [41] et MoMuLV [42] ont mis en évidence un site d'entrée interne contrôlant la synthèse des précurseurs *gag* et *gag-pol*. D'autres résultats plus récents ont montré l'existence d'un IRES précédant le codon d'initiation de *gag* dans l'oncovirus aviaire REV-A [43], dans l'ARN sous-génomique codant pour la protéine env dans l'oncovirus FrMuLV [44], ainsi que dans le lentivirus SIVmac. Ces données suggèrent que la synthèse des protéines rétrovirales dépend d'IRES et requiert des facteurs cellulaires spécifiques, comme c'est le cas pour les picornavirus. Les identifier et caractériser leur mécanisme d'action devrait permettre de mieux comprendre le fonctionnement des IRES et l'importance de la traduction au cours du cycle réplcatif des rétrovirus.

Le transfert de gènes et les vecteurs rétroviraux

Le transfert de gènes a pour but d'exprimer un ou plusieurs gènes d'intérêt dans une population de cellules afin de corriger un dysfonctionnement cellulaire, ou pour tuer une cellule cancéreuse. Certaines situations nécessitent l'expression de deux protéines distinctes : par exemple, une protéine utilisée comme traceur, et la protéine d'intérêt thérapeutique, ou encore deux sous-unités impliquées dans un mécanisme de cascade enzymatique. Plusieurs stratégies fondées sur une régulation transcriptionnelle ont été développées pour exprimer deux protéines dans un type cellulaire, soit en utilisant deux promoteurs, soit en utilisant un mécanisme d'épissage réglé (figure 5). La présence de deux promoteurs sur la même molécule d'ADN aboutit souvent à une compétition, et à la suppression épigénétique de l'un d'eux. De la même façon, le mécanisme d'épissage réglé permet

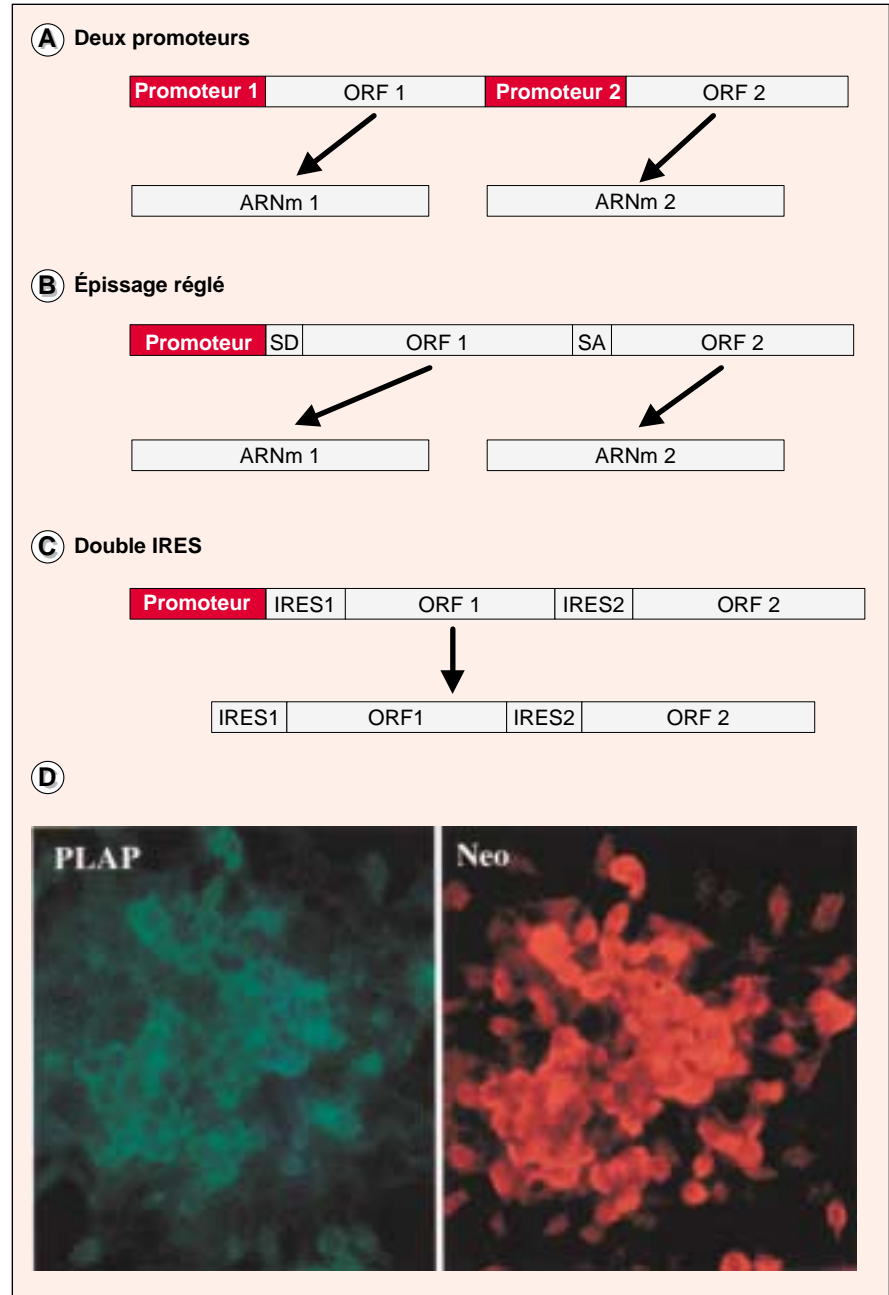


Figure 5. **Stratégies pour exprimer deux protéines à partir d'un vecteur unique.** **A.** Deux promoteurs permettent d'obtenir deux ARN différents après transcription, chacun codant pour la protéine d'intérêt. **B.** La présence des signaux d'épissage SD (splice donor) et SA (splice acceptor) permet l'épissage réglé de l'ARN recombinant et aboutit à la production de deux messagers qui exprimeront les protéines d'intérêt. **C.** L'utilisation d'un vecteur contenant deux IRES permet l'expression de deux protéines à partir du même ARNm. **D.** Exemple d'utilisation d'un vecteur rétroviral bicistronique MLV-REV, à titre élevé. Le gène marqueur codant pour la phosphatase alcaline humaine (Plap), et un gène de sélection codant pour la néomycine (neo), ont été co-exprimés dans des cellules de précurseurs neuronaux humains. Les protéines résultantes sont détectées par immunocytochimie. Il est remarquable de voir que les deux gènes sont exprimés dans ces cellules pluripotentes. Ils le sont aussi chez les descendants différenciés que sont les neurones, les astrocytes et les oligodendrocytes du système nerveux central humain [51].

d'exprimer deux protéines mais le rapport d'expression qui en résulte est déséquilibré, et les éléments de la séquence codante des gènes exprimés pourraient interférer avec le mécanisme d'épissage.

Exploitant les propriétés des IRES, des vecteurs rétroviraux bicistroniques ont été utilisés pour diriger l'expression de deux gènes à partir de la même unité transcriptionnelle [45-47]. La régulation de l'expression génique au niveau traductionnel permet d'éliminer les problèmes liés à la suppression épigénétique. De plus, les deux gènes d'intérêt sont exprimés de façon indépendante à la fois *in vitro* et *ex vivo*. Ainsi, *in vitro*, en quantifiant le nombre de résidus méthionine radiomarqués, on s'aperçoit que l'expression du deuxième gène, sous contrôle d'une séquence IRES de type EMCV (*encephalomyocarditis virus*), est souvent supérieure à celle du gène en amont qui est sous la dépendance de la coiffe. Dans des cultures de cellules, l'utilisation d'inhibiteurs spécifiques de l'initiation dépendante de la coiffe – comme la rapamycine – a montré que l'inhibition de l'expression du gène en amont n'affectait pas l'expression du gène rapporteur en aval [48]. Nous avons construit un vecteur rétroviral dans lequel deux séquences IRES issues de deux virus distincts permettent la co-expression de deux protéines et cela dans des types cellulaires variés [43, 49]. Nous avons également montré que trois IRES d'origine virale (EMCV, REV A et MuLV) peuvent « cohabiter » sur la même unité transcriptionnelle et permettre l'expression de trois protéines distinctes. Après transduction par un vecteur rétroviral de type MLV-IRES – ici le MLV-REV – les protéines sont exprimées dans des cellules humaines multipotentes non différenciées, et également dans les cellules différenciées de phénotype glial et neuronal qui en dérivent [50].

Conclusions et perspectives

En un peu plus de dix ans, nous sommes passés d'une vision assez simpliste du contrôle du déclenchement de la synthèse des protéines chez les eucaryotes, à une perception beau-

coup plus colorée, et riche de mécanismes alternatifs et de régulations potentielles de cette étape-clé de l'expression des gènes. On a au minimum trois mécanismes de base, qui sont le balayage classique des ribosomes de la coiffe 5' au codon d'initiation AUG (mais pas toujours AUG), l'entrée interne des ribosomes par l'intermédiaire d'un IRES (qui peut être éventuellement précédée ou suivie par un balayage) et le *shunt* qui pourrait être un *patchwork* entre le balayage et l'IRES. Une fois de plus, les virus, et plus précisément les picornavirus, ont joué un rôle moteur dans ces découvertes.

On peut aussi souligner que notre vision de l'ARN messager eucaryote tend à profondément se modifier, car l'ARNm aurait une structure pseudo-circulaire au moment de l'initiation de la traduction. En effet, la *poly(A) binding protein* (Pabp) et eIF4G formeraient un pont 5'-3' entre la coiffe et la queue polyA. L'avenir devrait préciser quels éléments de ces deux protéines sont indispensables à ce pont, si d'autres facteurs protéiques et des structures ARN sont nécessaires à la formation de l'ARNm pseudo-circulaire, et quelle est la contribution des interactions Pabp-eIF4G à la régulation de la synthèse des protéines.

Selon l'adage « ni trop, ni trop peu », en l'occurrence de protéines, la régulation du déclenchement de la traduction serait un moyen de régler très finement et rapidement le niveau d'une protéine dans une cellule. Cet ajustement très fin peut toucher des facteurs de transcription et des facteurs de croissance (voir Tableau I), et pour cela faire appel à des protéines cellulaires, comme l'auto-antigène La ou la protéine PTB, qui reconnaissent des séquences dans la région 5' non traduite (5' UTR). La réponse à des stimulus, comme le changement de conditions physiologiques (c'est-à-dire un *stress oxydant*) ou une infection virale, doit être rapide et donc pouvoir modifier profondément l'initiation de la traduction (c'est-à-dire eIF2 et PKR; FGF-2 et VEGF). La façon dont s'organise le jeu entre des facteurs cellulaires, les séquences 5' UTR (et 3' UTR), et les facteurs de base de l'initiation de la traduction, pour contrôler la synthèse protéique en réponse à des stimulus, à la diffé-

renciation cellulaire, ou à une infection virale, reste pour l'essentiel à découvrir.

L'utilisation d'IRES viraux a contribué à améliorer les caractéristiques des vecteurs rétroviraux pour le transfert de gènes. Ainsi, les vecteurs rétroviraux polycistroniques à deux transgènes entre les deux LTR. De ce fait, l'expression des transgènes sous le contrôle d'IRES est devenue beaucoup plus stable et équilibrée. Enfin, un IRES comme celui du virus REV-A (voir Tableau I) contient aussi un signal d'export nucléaire et tout ou partie d'un signal d'encapsulation de l'ARN génomique, ce qui contribue aux titres élevés des vecteurs rétroviraux MLV-REV-A (environ 10^8 à 10^9 TU/ml de milieu). La découverte d'un nombre croissant d'IRES dans les virus et les rétro-éléments (rétrovirus et rétrotransposons) devrait permettre de construire des vecteurs viraux polycistroniques, dérivés d'oncorétrovirus (c'est-à-dire MuLV), de lentivirus (c'est-à-dire HIV, VIS) ou d'adénovirus, selon des protocoles variés, et en quelque sorte à la demande. Ces objectifs ne pourront être atteints avec succès que si nos connaissances sur les IRES, leurs spécificités et leurs propriétés vis-à-vis de l'expression de transgènes, continuent de progresser. En fonction des objectifs poursuivis, fondamentaux et/ou thérapeutiques, des collaborations très étroites devront s'établir entre différents laboratoires (publics et/ou privés), et entre laboratoires et services cliniques ■

RÉFÉRENCES

1. Pain VM. Initiation of protein synthesis in eukaryotic cells. *Eur J Biochem* 1996; 236: 747-71.
2. Das S, Maiti T, Das K, Maitra U. Specific interaction of eukaryotic translation initiation factor 5 (eIF5) with the beta-subunit of eIF2. *J Biol Chem* 1997; 272: 31712-8.
3. Asano K, Krishnamoorthy T, Phan L, Pavitt GD, Hinnebusch AG. Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *EMBO J* 1999; 18: 1673-88.

RÉFÉRENCES

4. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 1997; 89: 951-61.
5. Matsuo H, Li H, McGuire AM, *et al.* Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat Struct Biol* 1997; 4: 717-24.
6. Linder P, Lasko PF, Ashburner M, *et al.* Birth of the D-E-A-D box. *Nature* 1989; 337: 121-2.
7. Pause A, Methot N, Svitkin Y, Merrick WC, Sonenberg N. Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. *EMBO J* 1994; 13: 1205-15.
8. Hentze MW. eIF4G: a multipurpose ribosome adapter? *Science* 1997; 275: 500-1.
9. Morley SJ, Curtis PS, Pain VM. eIF4G: translation's mystery factor begins to yield its secrets. *RNA* 1997; 3: 1085-104.
10. Pestova TV, Borukhov SI, Hellen CU. Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* 1998; 394: 854-9.
11. Sonenberg N, Gingras AC. The mRNA 5' cap-binding protein eIF4E and control of cell growth. *Curr Opin Cell Biol* 1998; 10: 268-75.
12. Kozak M. Effects of long 5' leader sequences on initiation by eukaryotic ribosomes *in vitro*. *Gene Expr* 1991; 1: 117-25.
13. Kozak M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem* 1991; 266: 19867-70.
14. Kozak M. The scanning model for translation: an update. *J Cell Biol* 1989; 108: 229-41.
15. Sprengart ML, Porter AG. Functional importance of RNA interactions in selection of translation initiation codons. *Mol Microbiol* 1997; 24: 19-28.
16. Tarun SZ Jr, Sachs AB. A common function for mRNA 5' and 3' ends in translation initiation in yeast. *Genes Dev* 1995; 9: 2997-3007.
17. Tarun SZ Jr, Sachs AB. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J* 1996; 15: 7168-77.
18. Imataka H, Gradi A, Sonenberg N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J* 1998; 17: 7480-9.
19. Fraser CS, Pain VM, Morley SJ. The association of initiation factor 4F with poly(A)-binding protein is enhanced in serum-stimulated *Xenopus* kidney cells. *J Biol Chem* 1999; 274: 196-204.
20. Clemens MJ. PKR-a protein kinase regulated by double-stranded RNA. *Int J Biochem Cell Biol* 1997; 29: 945-9.
21. Lamphear BJ, Panniers R. Cap binding protein complex that restores protein synthesis in heat-shocked Ehrlich cell lysates contains highly phosphorylated eIF-4E. *J Biol Chem* 1990; 265: 5333-6.
22. Minich WB, Balasta ML, Goss DJ, Rhoads RE. Chromatographic resolution of *in vivo* phosphorylated and nonphosphorylated eukaryotic translation initiation factor eIF-4E: increased cap affinity of the phosphorylated form. *Proc Natl Acad Sci USA* 1994; 91: 7668-72.
23. Pause A, Belsham GJ, Gingras AC, *et al.* Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* 1994; 371: 762-7.
24. Lin TA, Kong X, Haystead TA, *et al.* PHAS-I as a link between mitogen-activated protein kinase and translation initiation. *Science* 1994; 266: 653-6.
25. Mader S, Lee H, Pause A, Sonenberg N. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol Cell Biol* 1995; 15: 4990-7.
26. Haghghat A, Mader S, Pause A, Sonenberg N. Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic translation factor-4E. *EMBO J* 1995; 14: 5701-9.
27. Jackson RJ, Kaminski A. Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA* 1995; 1: 985-1000.
28. Jackson RJ, Wickens M. Translational controls impinging on the 5'-untranslated region and initiation factor proteins. *Curr Opin Genet Dev* 1997; 7: 233-41.
29. Etchison D, Milburn SC, Edery I, Sonenberg N, Hershey JW. Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220 000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. *J Biol Chem* 1982; 257: 14806-10.
30. Ohlmann T, Rau M, Pain V, Morley SJ. The C-terminal domain of eukaryotic protein synthesis initiation factor (eIF) 4G is sufficient to support cap-independent translation in the absence of eIF4E. *EMBO J* 1996; 15: 1371-82.
31. Ohlmann T, Pain VM, Wood W, Rau M, Morley SJ. The proteolytic cleavage of eukaryotic initiation factor (eIF) 4G is prevented by eIF4E binding protein (PHAS-I; 4E-BP1) in the reticulocyte lysate. *EMBO J* 1997; 16: 844-55.
32. Gingras AC, Svitkin Y, Belsham GJ, Pause A, Sonenberg N. Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus. *Proc Natl Acad Sci USA* 1996; 93: 5578-83.
33. Vagner S, Gensac MC, Maret A, *et al.* Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. *Mol Cell Biol* 1995; 15: 35-45.
34. Vagner S, Touriol C, Galy B, *et al.* Translation of CUG- but not AUG-initiated forms of human fibroblast growth factor 2 is activated in transformed and stressed cells. *J Cell Biol* 1996; 135: 1391-402.
35. Rhoads RE, Joshi-Barve S, Rinker-Schaeffer C. Mechanism of action and regulation of protein synthesis initiation factor 4E: effects on mRNA discrimination, cellular growth rate, and oncogenesis. *Prog Nucleic Acids Res Mol Biol* 1993; 46: 183-219.
36. Gan W, Celle ML, Rhoads RE. Functional characterization of the internal ribosome entry site of eIF4G mRNA. *J Biol Chem* 1998; 273: 5006-12.
37. Futterer J, Kiss-Laszlo Z, Hohn T. Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* 1993; 73: 789-802.
38. Schmidt-Puchta W, Dominguez D, Lewentag D, Hohn T. Plant ribosome shunting *in vitro*. *Nucleic Acids Res* 1997; 25: 2854-60.
39. Latorre P, Kolakofsky D, Curran J. Sendai virus Y proteins are initiated by a ribosomal shunt. *Mol Cell Biol* 1998; 18: 5021-31.
40. Yueh A, Schneider RJ. Selective translation initiation by ribosome jumping in adenovirus-infected and heat-shocked cells. *Genes Dev* 1996; 10: 1557-67.
41. Berlioz C, Darlix JL. An internal ribosomal entry mechanism promotes translation of murine leukemia virus gag polyprotein precursors. *J Virol* 1995; 69: 2214-22.
42. Vagner S, Waysbort A, Marena M, Gensac MC, Amalric F, Prats AC. Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor. *J Biol Chem* 1995; 270: 20376-83.
43. Lopez-Lastra M, Gabus C, Darlix JL. Characterization of an internal ribosomal entry segment within the 5' leader of avian reticuloendotheliosis virus type A RNA and development of novel MLV-REV-based retroviral vectors. *Hum Gene Ther* 1997; 8: 1855-65.
44. Deffaud C, Darlix JL. Characterization of an internal ribosomal entry segment in the 5' leader of murine leukemia virus Env RNA. *J Virol* 2000; 74 (sous presse).
45. Adam MA, Ramesh N, Miller AD, Osborne WR. Internal initiation of translation in retroviral vectors carrying picornavirus 5' nontranslated regions. *J Virol* 1991; 65: 4985-90.
46. Koo HM, Brown AM, Kaufman RJ, Prockrock CM, Ron Y, Dougherty JP. A spleen necrosis virus-based retroviral vector which expresses two genes from a dicistronic mRNA. *Virology* 1992; 186: 669-75.
47. Morgan RA, Couture L, Elroy-Stein O, Ragheb J, Moss B, Anderson WF. Retroviral vectors containing putative internal ribosome entry sites: development of a polycistronic gene transfer system and applications to human gene therapy. *Nucleic Acids Res* 1992; 20: 1293-9.

RÉFÉRENCES

48. Beretta L, Gingras AC, Svitkin YV, Hall MN, Sonenberg N. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. *EMBO J* 1996; 15: 658-64.

49. Torrent C, Berlioz C, Darlix JL. Stable MLV-VL30 dicistronic retroviral vectors with a VL30 or MoMLV sequence promoting both packaging of genomic RNA and expression of the 3' cistron. *Hum Gene Ther* 1996; 7: 603-12.

50. Derrington EA, Lopez-Lastra M, Chapel-Fernandez S, et al. Retroviral vectors for the expression of two genes in human multipotent neural precursors and their differentiated neuronal and glial progeny. *Hum Gene Ther* 1999; 10: 1129-38.

51. Derrington EA, Dufay N, Rudkin BB, Belin MF. Human primitive neuroectodermal tumour cells behave as multipotent neural precursors in response to FGF2. *Oncogene* 1998; 17: 1663-72.

TIRÉS À PART

J.L. Darlix.

Summary

Translation initiation in eukaryotic cells

Three different mechanisms are reviewed by which protein translation can be initiated. The scanning mechanism, in which the ribosome attaches to the 5' cap structure before migrating along the mRNA in a 5' to 3' direction until an AUG start codon is encountered, is the most thoroughly characterized. The function of initiation factors in the formation of the ribosomal complex and its attachment to the mRNA are considered. Furthermore distinct mechanisms such as internal initiation, by which the ribosome is directly recruited at a site downstream of the 5' cap (IRES), and ribosomal shunting which allows a 5' cap-bound ribosome to bypass stable secondary structures in the untranslated 5' region intervening between the cap and the AUG start codon are also described. The importance of the novel mechanisms for the construction of MLV-IRES vectors in which internal initiation allows co-expression of two distinct gene products from a single mRNA molecule is considered, as is the possibility that internally initiated translation may represent a novel level at which transgene expression may be controlled.

LES JOURNÉES DE LA DERMATITE ATOPIQUE

Vendredi 19 et samedi 20 mai 2000

Organisateurs : AFPADA

Association française des Personnes Atteintes de Dermatite Atopique

Lieu : École Normale Supérieure de Lyon

Vendredi 19 mai

Ouverture du congrès (AFPADA) • Présentation clinique de la DA (Pr Béatrice Crickx) • Physiopathologie, Développement de l'atopie chez le fœtus, le nouveau-né et le nourrisson (Pr Thomas Bieber) • Le syndrome dermo-respiratoire (Dr Marie-Thérèse Guinépain) • Infections et DA (Pr Jean-François Stalder) • Comprendre la DA (Dr Sylvie Consoli) • Place de l'allergie dans la DA (Dr Gisèle Kanny) • Immunité muqueuse dans la DA (Dr Claude André) • Diététique de l'enfant atopique (Mme Patricia Sergeant) • Eczéma de contact et DA (Dr Annick Barbaud) • Quels tests dans la DA (Dr Michel Castelain) • Orientation professionnelle de la dermatite (Pr Christian Gérard)

Samedi 20 mai

Traitements de la DA (Pr Frédéric Cambazard) • Les Immunosuppresseurs (Pr Yves de Prost) • Photothérapie (Dr Pierre Fortier) • Étude ETAC (Pr Alain Taieb) • Perspectives thérapeutiques (Pr Jean-François Nicolas) • Table Ronde (entre les participants et les orateurs) : « Les questions que l'on se pose sur la DA »

Secrétariat Scientifique :

Association française des Personnes Atteintes de Dermatite Atopique (AFPADA)
BP 36, 77982 Saint-Fargeau-Ponthierry Cedex, France – Tél. Fax (+33) (0)1 60 65 79 05

Secrétariat Technique-Inscriptions :

Société Bawan Stratégie – Gilles Barbier – Immeuble le Rive gauche/12, rue de Cavenne/69007
Lyon, France. Tél. : (+33) (0)4 78 61 09 09 – Fax : (+33) (0)4 72 71 81 06 – E-mail : bawan@media-sites.com