

# La structure-coiffe des ARN messagers

Martin Bisailon

Les ARN messagers (ARNm) des cellules eucaryotes subissent d'importantes modifications avant d'être ultérieurement traduits en protéines. Ainsi, l'ajout d'une structure-coiffe à l'extrémité 5' des ARNm augmente considérablement la stabilité de ces molécules et promeut également une initiation efficace de la synthèse protéique. Les enzymes responsables de la synthèse de cette structure ont maintenant été identifiées chez de nombreux organismes. La caractérisation de ces protéines de diverses origines nous permet désormais d'envisager le développement d'inhibiteurs spécifiques ayant des applications thérapeutiques importantes pour empêcher le développement de certains agents pathogènes.

**D**ans les cellules eucaryotes, l'extrémité 5' des ARN messagers (ARNm) possède une structure fondamentale impliquée dans le métabolisme de ces molécules. Cette structure, appelée structure-coiffe, est caractérisée par la présence d'un résidu guanosine méthylé en position N7 et relié par une liaison caractéristique 5'-5' sur le premier nucléotide de l'ARNm ( $m^7GpppN$ ) [1]. Les groupements 2'-hydroxyle des nucléotides adjacents ( $m^7GpppNpN$ ) peuvent également être méthylés, conduisant ainsi à la formation de structures désignées coiffe 0 ( $m^7GpppNpN$ ), coiffe 1 ( $m^7Gppp^{m2}NpN$ ) ou coiffe 2 ( $m^7Gppp^{m2}Np^{m2}N$ ), selon le nombre de méthylations (*figure 1A*).

## Formation de la structure-coiffe

L'analyse d'acides nucléiques viraux et des expériences de transcription *in vitro* ont initialement permis d'élucider les mécanismes impliqués dans la synthèse de la structure-coiffe [2, 3]. Ces travaux ont révélé que trois activités enzymatiques sont requises pour la synthèse de cette structure (*figure 1B*): (1) une ARN 5'-triphosphatase hydrolyse tout d'abord le phosphate  $\gamma$  situé à l'extrémité triphosphorylée

des molécules d'ARNm; (2) une ARN guanylyltransférase catalyse ensuite l'ajout d'un groupement GMP (guanosine monophosphate) à l'extrémité diphosphorylée, induisant ainsi la formation d'une extrémité possédant un résidu guanosine rattaché par une liaison 5'-5' au premier nucléotide de l'ARNm; (3) enfin, une ARN (guanine-7) méthyltransférase assure la méthylation du résidu guanosine en position N7. Dans les cellules eucaryotes, toutes ces étapes se déroulent dans le noyau dès l'amorce de la transcription [4]. Une autre méthylase, l'ARN (nucléoside-2'-O)-méthyltransférase, présente dans le cytoplasme, peut par la suite assurer la méthylation des résidus adjacents lorsque l'ARNm est situé au niveau des polyribosomes [5].

Les ARNm de la plupart des virus à ADN (papovavirus, adénovirus, herpesvirus) sont synthétisés par l'ARN polymérase II cellulaire puisque ceux-ci se répliquent dans le noyau des cellules infectées. De ce fait, l'extrémité 5' de ces ARNm viraux est modifiée par les enzymes cellulaires pour former la structure-coiffe. Le virus de la vaccine, qui possède un cycle de multiplication exclusivement cytoplasmique, possède quant à lui ses propres enzymes responsables de la formation de la structure-coiffe.

## ADRESSE

M. Bisailon: Molecular Biology Program, Sloan-Kettering Institute, 1275 York Avenue, Box 73, New York, NY 10021, États-Unis.

## TIRÉS À PART

M. Bisailon.

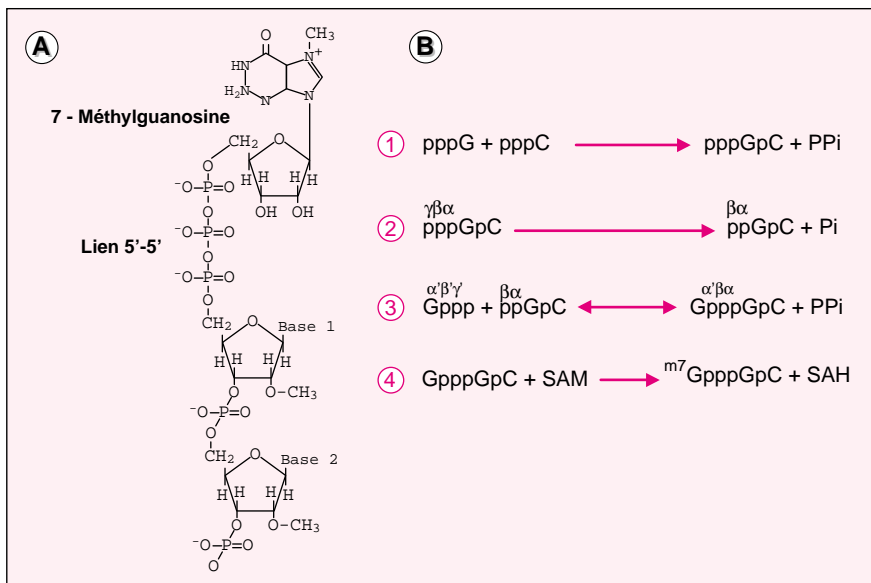


Figure 1. **La structure-coiffe retrouvée à l'extrémité 5' des ARNm eucaryotes.** **A.** Un résidu guanosine méthylé en position N7 est relié par un lien caractéristique 5'-5' au premier nucléotide de l'ARNm. Les groupements 2'-hydroxyle des nucléotides adjacents à la structure-coiffe peuvent également être méthylés. **B.** Activités enzymatiques impliquées dans la formation de la structure-coiffe. À la suite de la polymérisation des deux premiers nucléotides pas l'ARN polymérase II (1), une ARN 5'-triphosphatase hydrolyse le phosphate  $\gamma$  situé à l'extrémité triphosphorylée des molécules d'ARNm (2); une ARN guanylyltransférase catalyse ensuite l'ajout d'un groupement GMP, conduisant à la formation d'une extrémité possédant un résidu guanosine lié par une liaison 5'-5' au premier nucléotide de l'ARNm (3); enfin, une ARN (guanine-7) méthyltransférase assure la méthylation du résidu guanosine en position N7 (4). SAM: S-adenosyl-L-méthionine; SAH: S-adenosyl-L-homocystéine; Pi: phosphate inorganique; Ppi: pyrophosphate.

Enfin, bien que les travaux concernant les mécanismes utilisés par les virus à ARN pour assurer la synthèse de leurs structures-coiffe soient relativement récents, les données actuelles semblent toutefois indiquer que ces virus synthétisent leurs propres protéines possédant toutes les activités enzymatiques requises pour la synthèse de la structure-coiffe [6].

### Importance de la structure-coiffe

La présence de la structure-coiffe constitue un facteur critique pour l'initiation de la synthèse protéique. Des essais de traduction *in vitro* ont en premier lieu démontré que les ARNm non coiffés ne sont pas traduits efficacement par la machinerie traductionnelle cellulaire [7]. Ces études ont aussi clairement indiqué que la présence du groupement  $m^7G$  à l'extrémité 5' des ARNm augmente l'efficacité de la liaison des ARNm à

la sous-unité ribosomique 40S lors de l'initiation de la traduction [8]. Certaines protéines cellulaires jouent toutefois un rôle critique au niveau de l'initiation de la traduction en reconnaissant de manière spécifique la structure-coiffe des ARNm. En effet, la liaison des molécules d'ARNm à la petite sous-unité ribosomique 40S requiert l'action concertée de nombreux facteurs d'initiation de la traduction, dont le complexe de liaison à la structure-coiffe (eIF-4F). La formation de ce complexe est un processus coopératif qui nécessite initialement la liaison de la protéine eIF4E à la structure-coiffe des ARNm. Cette liaison favorise l'adjonction séquentielle des protéines p220 (anciennement eIF4 $\gamma$ ) et eIF4A (ARN hélicase). La liaison subséquente de la protéine eIF4B a pour effet de stimuler l'activité hélicase de eIF4A, qui est responsable du déroulement des régions appariées de la portion non traduite des ARNm, favorisant ainsi

l'accès du codon d'initiation AUG à la machinerie traductionnelle [9].

La structure-coiffe possède également un rôle stabilisateur important pour les ARNm. Il a en effet été démontré, à l'aide d'ARNm viraux injectés dans des ovocytes de *Xenopus laevis*, que les molécules possédant une structure  $m^7GpppGm$  ou  $GpppG$  à leur extrémité 5' sont dégradées plus lentement que les ARNm possédant des structures de type  $pppGm$ ,  $pppG$  ou  $ppG$  [10]. De plus, des analyses cinétiques ont révélé des niveaux de dégradation similaires pour des molécules possédant des structures de type  $m^7GpppGm$  et  $GpppG$ . La présence du groupement guanosine en position 5' (sans méthylation) est donc apparemment suffisante pour la protection des ARNm contre les nucléases cellulaires. On a d'ailleurs remarqué que les ARNm non coiffés sont dégradés dans la direction 5'-3' par les nucléases cellulaires [11]. Cette dégradation exonucléolytique engendre essentiellement des mononucléotides alors que des produits de dégradation intermédiaires (oligonucléotides) ne sont en général pas retrouvés.

De nombreux gènes possèdent des séquences introniques non codantes qui doivent être éliminées lors de la maturation des ARNm. Des travaux utilisant des analogues de la structure-coiffe ont clairement démontré le rôle essentiel de la structure-coiffe dans l'épissage des ARNm, relayé par un complexe hétérodimérique composé des protéines CBP20 (*cap binding protein 20*) et CBP80 (*cap binding protein 80*) [12]. Chacune de ces deux protéines est requise pour la liaison de la structure-coiffe. Aucune des sous-unités ne peut se lier à la structure coiffe, ou à l'ARNm, par elle-même. Dans les cellules eucaryotes, l'élimination des introns se déroule au niveau d'un complexe nucléoprotéique (splicéosome) constitué de 5 petites ribonucléoprotéines nucléaires (snRNP) (U1, U2, U4/U6, U5) et de nombreux autres facteurs d'épissage peu caractérisés. Or, le complexe CBP20-CBP80 facilite l'association de U1 au site d'épissage et participe ainsi à la formation du complexe nucléoprotéique requis pour l'épissage des ARNm [13]. L'utilisation d'anticorps dirigés spécifiquement contre CBP20-CBP80 a par

ailleurs permis de confirmer l'importance de ces protéines dans le processus d'épissage *in vivo* et *in vitro*.

Il est intéressant de remarquer que le complexe formé des protéines CBP20 et CBP80 est également requis pour assurer le passage des ARNm mûrs vers le cytoplasme [12]. En effet, il a été démontré que le complexe CBP20-CBP80 demeure associé aux ARNm lors du passage vers le cytoplasme. Notons cependant que l'utilisation d'inhibiteurs dirigés contre CBP20 et CBP80 n'abolit pas entièrement le transport des ARNm vers le cytoplasme, d'autres facteurs contribuant probablement à l'efficacité de ce processus.

Notons, enfin, qu'à la suite de la synthèse des petits ARN nucléaires (ARNsn), ceux-ci sont transportés dans le cytoplasme pour y subir une modification additionnelle (<sup>m2,2,7</sup>Gp-ppN) [14]. Cette nouvelle structure triméthylée, de concert avec la liaison de certaines ribonucléoprotéines (snRNP), constitue un signal de localisation nucléaire et favorise le transport des molécules d'ARN du cytoplasme vers le noyau. De plus, il a été suggéré que cette modification additionnelle empêche le recrutement de ce type d'ARN à la machinerie traductionnelle [12].

## Études biochimiques

Bien que chacune des étapes nécessaires à la synthèse de la structure-coiffe ait été démontrée il y a déjà plus de 20 ans, l'identification des gènes impliqués dans la synthèse de la structure-coiffe n'est survenue que beaucoup plus récemment. L'expression, la purification et la caractérisation des diverses protéines a permis de démontrer des différences significatives quant à l'organisation physique et fonctionnelle des enzymes nécessaires à la synthèse de la structure-coiffe des ARNm chez divers organismes (métazoaires, levures, virus). Différentes stratégies sont donc utilisées par ces organismes pour établir un lien étroit entre les activités enzymatiques (figure 2). Par exemple, les sites actifs de l'ARN 5'-triphosphatase, de l'ARN guanylyltransférase et de l'ARN (guanine-7) méthyltransférase sont positionnés de façon modulaire au sein d'un unique peptide chez les virus à ADN se répliquant dans le cytoplasme des cellules infectées (virus de la vaccine, virus du fibrome de Shope, virus africain de la fièvre porcine) [6]. Les métazoaires (homme, souris, *Caenorhabditis elegans*) semblent posséder,

quant à eux, une protéine bifonctionnelle possédant les activités ARN 5'-triphosphatase et ARN guanylyltransférase. Chez ces métazoaires, une autre protéine est responsable de l'activité ARN (guanine-7) méthyltransférase [15]. Enfin, la formation de la structure-coiffe chez les levures (*Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Candida albicans*) implique des protéines dans lesquelles les activités ARN 5'-triphosphatase, ARN guanylyltransférase et ARN (guanine-7) méthyltransférase sont catalysées par des protéines distinctes [16-18]. Notons toutefois que ces protéines de levures forment un complexe hétéromérique permettant ainsi un couplage étroit entre les différentes activités enzymatiques.

Contrairement aux protéines eucaryotes et à celles des virus à ADN, les protéines impliquées dans la synthèse de la coiffe des virus à ARN n'ont été que partiellement identifiées et caractérisées (Tableau I). Cependant, des études effectuées chez trois groupes de virus révèlent des différences significatives quant à l'organisation physique des protéines. Ainsi, les réovirus semblent assurer leurs activités ARN guanylyltransférase et ARN (guanine-7) méthyltransférase par l'intermé-



Figure 2. **Organisation fonctionnelle des enzymes impliquées dans la synthèse de la structure-coiffe de divers organismes.** Les sites actifs de l'ARN 5'-triphosphatase, de l'ARN guanylyltransférase et de l'ARN (guanine-7) méthyltransférase sont positionnés de façon modulaire au sein d'un unique peptide chez les virus à ADN se répliquant dans le cytoplasme des cellules infectées. Les métazoaires possèdent une protéine bifonctionnelle possédant les activités ARN 5'-triphosphatase et ARN guanylyltransférase et une autre protéine est responsable de l'activité ARN (guanine-7) méthyltransférase. La formation de la structure-coiffe chez les levures implique plutôt des protéines dans lesquelles chacune des activités est catalysée de façon distincte. Enfin, différentes stratégies sont utilisées par les virus à ARN afin d'assurer la synthèse de la structure-coiffe.

diaire d'une seule protéine [19]. Une situation analogue est également rencontrée chez les alphavirus [20]. Dans chacun des cas, l'activité ARN 5'-triphosphatase est catalysée par une autre protéine qui a toutefois la capacité d'interagir avec le peptide possédant les activités ARN triphosphatase et ARN guanylyltransférase [20, 21]. En revanche, chez les orbivirus, une protéine semble posséder toutes les activités nécessaires à la synthèse de la structure-coiffe (ARN 5'-triphosphatase, ARN guanylyltransférase et ARN (guanine-7) méthyltransférase) [22]. En somme, différentes stratégies semblent avoir été développées au cours de l'évolution afin d'assurer la synthèse de la structure-coiffe des ARNm.

Les protéines possédant une activité ARN 5'-triphosphatase divergent quant à leurs structures et aux mécanismes utilisés pour assurer leur acti-

vité enzymatique. Deux classes distinctes de protéines possédant ce type d'activité ont été décrites: la famille des ARN 5'-triphosphatases dépendantes ou indépendantes de la présence de cations divalents [23]. Ainsi, l'activité ARN 5'-triphosphatase des levures, de baculovirus, du virus de la vaccine, des réovirus et des alphavirus est dépendante de la présence de cations divalents (magnésium, manganèse, cobalt) alors que l'activité des protéines de métazoaires (homme, souris, *C. elegans*) semble indépendante de la présence de ces cations divalents. Par ailleurs, l'analyse de la structure primaire des diverses protéines possédant une activité ARN 5'-triphosphatase révèle la présence de certains motifs caractéristiques. Ainsi, le domaine responsable de l'activité ARN 5'-triphosphatase des protéines de métazoaires est caractérisé par la présence d'un motif

consensus (I/V)HcxAGxGR(ST)G initialement observé chez les membres de la famille des protéines tyrosine-phosphatases [24]. Des travaux de mutagenèse ont d'ailleurs clairement révélé l'importance de ce motif pour l'activité de ces protéines. L'analyse des ARN 5'-triphosphatases de levures et de virus à ADN a également révélé la présence de certains motifs caractéristiques (désignés motifs A, B et C) distincts de ceux retrouvés chez les métazoaires [23]. Ces motifs sont caractérisés par la présence de résidus chargés qui semblent impliqués dans la liaison du substrat et/ou des cations divalents nécessaires pour l'activité enzymatique des protéines. Notons toutefois que des études récentes indiquent que de nombreux autres résidus situés hors de ces motifs consensus sont également importants pour l'activité des protéines de levures [25, 26].

La formation d'intermédiaires enzyme-guanylate a été identifiée chez de nombreuses guanylyltransférases. Dans la majorité des cas, la protéine catalyse le transfert d'un groupement GMP, à partir d'une molécule de GTP, à l'extrémité diphosphorylée des ARNm [27]. Le groupement GMP est lié à l'ARN guanylyltransférase par un lien phosphamide au groupement ε-amino d'un résidu lysine [28]. Le site actif des protéines possédant une activité ARN guanylyltransférase semble être conservé chez toutes les ARN guanylyltransférases des eucaryotes et des virus à ADN. Un motif de type KxDG (Lys-X-Asp-Gly) est en effet retrouvé chez tous ces organismes étudiés jusqu'à maintenant. Une exception existe toutefois chez les virus membres de la famille des *Reoviridae* qui possèdent une version quelque peu différente de ce motif. Notons par ailleurs que le motif KxDG est également observé chez de nombreuses protéines possédant une activité de type polynucléotide ligase (ADN ligases, ARN ligases, ARNt ligases) [29]. Tout comme les ARN guanylyltransférases, ces protéines catalysent une réaction enzymatique suite à la formation d'un lien enzyme-GMP. L'analyse de la structure tridimensionnelle de certaines ligases a d'ailleurs révélé l'adoption d'une structure tridimensionnelle très similaire à celle retrouvée chez

**Tableau 1.** Protéines de diverses origines impliquées dans la synthèse de la structure-coiffe.

Origine	Protéine	GTase	RTase	MTase
Humain	Hce (hCAP)	X	X	
	Hcm (hCMT)			X
Souris	Mce	X	X	
<i>Caenorhabditis elegans</i>	Cel1	X	X	
<i>Xenopus laevis</i>	xCAP	X	X	
	xCMT			X
<i>Saccharomyces cerevisiae</i>	Ceg1	X		
	Cet1		X	
	Abd1			X
<i>Candida albicans</i>	Cgt1	X		
	CaCet1		X	
	Ccm1			X
<i>Schizosaccharomyces pombe</i>	Pce1	X		
	Pcm1			X
Virus de la vaccine	D1	X	X	X
Virus du fibrome de Shope	D3R	X	X	X
Virus de la fièvre porcine africaine	NP868R	X	X	X
Réovirus	λ2	X		X
	λ1		X	
Virus bluetongue (orbivirus)	VP4	X	X	X
Rotavirus	VP3	X		X
Baculovirus	LEF-4	X	X	
Virus PBCV-1 de <i>Chlorella</i>	A103R	X		
Virus de la forêt de Semliki	Nsp1	X		X
	Nsp2		X	
Virus de la mosaïque du tabac	p126	X		X
Virus de la mosaïque de Brome	1a	X		X

Les protéines possédant des activités ARN 5'-triphosphatase (RTase), ARN guanylyltransférase (GTase) et ARN (guanine-7) méthyltransférase (MTase) sont indiquées.

les ARN guanylyltransférases. L'existence d'un ancêtre commun pour ces deux familles de protéines (ARN guanylyltransférase et polynucléotide ligase) a d'ailleurs été suggéré. Outre le motif KxDG, notons que d'autres motifs impliqués dans l'activité ARN guanylyltransférase ont été identifiés. L'importance de ces motifs additionnels (I, III, IIIa, IV, V, VI) a été clairement établie suite à des études de mutagenèse dirigée [30].

Les gènes responsables de la synthèse des protéines possédant une activité ARN (guanine-7) méthyltransférase de divers organismes n'ont été que récemment identifiés. Le site actif de ces protéines, constitué d'un résidu aspartate ou glutamate, semble être conservé chez les métazoaires, les levures et les virus à ADN. Notons également que la structure primaire des ARN (guanine-7) méthyltransférases cellulaires (*S. cerevisiae*, *S. pombe*, *C. elegans*, *C. albicans*, homme) présente un haut degré de similitude. La présence de résidus identiques ou structurellement conservés chez ces protéines suggèrent l'existence d'une structure et d'un mécanisme enzymatique apparenté.

### Spécificité de l'ARN polymérase II

Dans les cellules eucaryotes, les ARNm synthétisés par l'ARN polymérase II sont coiffés de façon co-transcriptionnelle tandis que les ARN synthétisés par l'ARN polymérase I et III ne possèdent pas de structure-coiffe à leur extrémité 5'. La raison de cette spécificité n'a été que récemment élucidée et implique le domaine carboxy-terminal (CTD) de l'ARN polymérase II (figure 3). Ce domaine est constitué d'une série de répétitions de l'heptapeptide YSPTSPS (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) qui ont été conservées au cours de l'évolution et dont le nombre varie de 26 (*S. cerevisiae*) à 52 (souris). Il a également été remarqué que le niveau de phosphorylation de ce domaine carboxy-terminal varie lors des diverses étapes de la transcription. Ainsi, lors de la formation du complexe de pré-initiation, de même que lors des étapes précoces de la transcription, le domaine carboxy-terminal est présent sous une forme hypophosphorylée, appelée ARN polymérase IIa. A

la suite de la synthèse des 20 à 30 premiers nucléotides des ARN pré-messagers, on a remarqué que l'ARN polymérase s'arrête avant de poursuivre l'élongation des molécules d'ARNm néosynthétisées. Il a été démontré que le domaine CTD est phosphorylé par le facteur de transcription TFIIF lors du passage au mode d'élongation, produisant ainsi la forme hyperphosphorylée nommée ARN polymérase IIo. Or, cette réaction survient précisément au moment même où la structure coiffe est synthétisée. L'étude des protéines impliquées dans la synthèse de la structure-coiffe chez les cellules eucaryotes a permis de démontrer que les protéines possédant une activité de type ARN guanylyltransférase et ARN

(guanine-7) méthyltransférase sont recrutées au complexe transcriptionnel d'élongation par leur capacité individuelle de liaison au domaine carboxy-terminal hyperphosphorylé de l'ARN polymérase II [31-33]. Chez les mammifères, l'activité ARN 5' triphosphatase est recrutée à ce complexe en raison de l'association physique entre l'ARN 5' triphosphatase et l'ARN guanylyltransférase présente au niveau du même peptide. Cependant, chez les levures, cette association se produit à la suite de l'hétéromérisation de ces deux sous-unités (ARN 5' triphosphatase et ARN guanylyltransférase). Notons, enfin, que l'analyse de cellules humaines a révélé que la présence d'un domaine carboxy-terminal possédant un

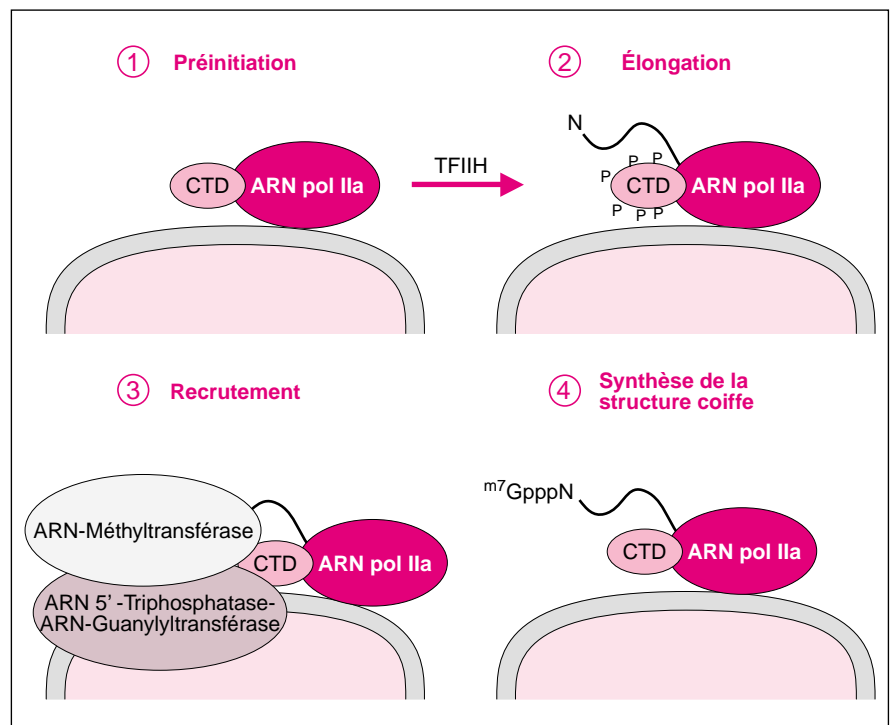


Figure 3. **Rôle du domaine carboxy-terminal de l'ARN polymérase II dans la synthèse de la structure-coiffe.** Lors de la formation du complexe de pré-initiation, de même que lors des étapes précoces de la transcription, le domaine carboxy-terminal (CTD) de l'ARN polymérase II (ARN pol IIa) est présent sous une forme hypophosphorylée, nommée ARN polymérase IIa. À la suite de la synthèse des 20 à 30 premiers nucléotides des ARN pré-messagers, le domaine carboxy-terminal est phosphorylé par le facteur de transcription TFIIF lors du passage au mode d'élongation, produisant ainsi la forme hyperphosphorylée désignée ARN polymérase IIo. Les protéines possédant une activité de type ARN guanylyltransférase et ARN (guanine-7) méthyltransférase sont recrutées dans le complexe transcriptionnel d'élongation par leur capacité de se lier individuellement au domaine carboxy-terminal hyperphosphorylé de l'ARN polymérase II. Chez les mammifères, l'activité ARN 5' triphosphatase est recrutée dans ce complexe en raison de l'association physique entre l'ARN 5' triphosphatase et l'ARN guanylyltransférase présente au niveau du même peptide.

nombre d'heptapeptides YSPTSPS restreint (5 plutôt que 52) entraîne une diminution marquée du nombre de molécules d'ARNm possédant une structure-coiffe.

L'identification des gènes responsables des trois activités enzymatiques nécessaires à la synthèse de la structure-coiffe chez les mammifères et la levure a permis de mieux comprendre le rôle de ces protéines *in vivo*. L'étude de la levure *S. cerevisiae* a largement contribué à mettre en évidence l'importance de la structure coiffe pour la survie cellulaire. Ainsi, l'analyse du gène *CEG1*, qui code pour l'ARN guanylyltransférase de la levure, a inauguré les études génétiques portant sur les enzymes cellulaires requises pour la formation de la structure-coiffe [17]. Des études de mutagenèse dirigée ont initialement permis d'identifier le site actif de la protéine (Ceg1) responsable de l'activité guanylyltransférase *in vitro* [34]. Il a ensuite été clairement démontré que des allèles de ce gène possédant les mutations correspondantes ne permettent pas le maintien de la croissance cellulaire [30]. Ces travaux ont donc constitué les premières preuves directes de la nécessité de la structure-coiffe *in vivo*. L'identification subséquente des gènes de *S. cerevisiae* responsables des activités ARN (guanine-7) méthyltransférase (*ABD1*) et ARN 5'-triphosphatase (*CET1*) a également permis d'établir que les trois activités enzymatiques sont essentielles pour la survie des levures [16, 18]. Il a été suggéré que la létalité observée en l'absence de ces activités enzymatiques est le résultat des effets cumulatifs sur la stabilité, l'épissage, le transport et la traduction des ARNm non coiffés.

### **Cibles potentielles contre les agents pathogènes**

En raison du rôle primordial de la coiffe de l'extrémité 5' des ARNm, les protéines impliquées dans la formation de cette structure constituent une cible potentielle pour le développement de molécules thérapeutiques. Il semble désormais qu'une inhibition sélective de ces protéines pourrait aboutir à d'importantes applications thérapeutiques. La

caractérisation récente de protéines de mammifères et de divers agents pathogènes permet désormais d'envisager l'utilisation d'inhibiteurs exploitant spécifiquement les différences entre ces protéines (mécanismes d'action, structure). En principe, chacune des étapes impliquées dans la synthèse de la coiffe représente une cible potentielle pour le développement d'une thérapie antivirale.

De nombreuses protéines virales impliquées dans la formation de la structure-coiffe ont été caractérisées au cours des dernières années. Ces protéines représentent des cibles potentielles pour le développement de composés antiviraux. Les rotavirus représentent particulièrement des cibles de choix, les traitements actuels étant d'une faible efficacité. Les protéines responsables de l'activité ARN 5'-triphosphatase constitueront probablement les cibles de choix pour le développement d'agents antiviraux puisque des différences importantes quant à la structure et aux mécanismes d'action sont observées entre les diverses protéines virales et les protéines de mammifères. Le développement d'agents thérapeutiques visant à inhiber spécifiquement les activités ARN guanylyltransférase et ARN méthyltransférase s'avèrera probablement plus difficile puisque les données actuelles indiquent que la structure tridimensionnelle et les mécanismes d'action de ces protéines semblent conservés chez les mammifères et les divers agents pathogènes étudiés jusqu'à présent.

Nos connaissances quant aux mécanismes mis en jeu par les protéines impliquées dans la synthèse de la structure-coiffe ont été nettement améliorées par la mise en évidence des structures tridimensionnelles de l'ARN guanylyltransférase du virus PBCV-1 (virus de l'algue verte *Paramecium bursaria Chlorella*) et de l'ARN 5'-triphosphatase de *S. cerevisiae* [35, 36]. Ces travaux ont permis de mieux comprendre l'importance de certains résidus, préalablement identifiés par des études de mutagenèse dirigée, dans les réactions catalysées par ces protéines. Ces approches ont permis de progresser dans la compréhension du mode d'action de ces protéines et permettent d'envisager éventuelle-

ment la synthèse de composés visant à inhiber spécifiquement l'action de ces protéines. Au cours des dernières années, le développement d'agents thérapeutiques a été grandement amélioré à la suite de la mise en évidence de la structure tridimensionnelle de diverses protéines importantes pour le développement de certains agents pathogènes. Cependant, relativement peu de données structurales sont disponibles quant aux protéines impliquées dans la synthèse de la structure-coiffe d'organismes pathogènes. La découverte de ces structures tridimensionnelles constituera un outil essentiel pour le développement de nouvelles thérapeutiques. Idéalement, l'agent antiviral devra posséder un certain nombre de propriétés. Le composé devra se lier à la protéine cible avec une haute affinité, posséder une haute spécificité (par exemple, avoir un effet minimal sur des protéines similaires de l'hôte), posséder un faible niveau de toxicité, agir sur le plus grand nombre d'isolats viraux possibles, pouvoir agir de façon synergique avec d'autres composés utilisés pour combattre le virus et posséder une demie-vie suffisamment longue pour permettre le maintien de concentrations suffisantes pour combattre l'infection virale. Certains problèmes sont toutefois associés au développement d'inhibiteurs antiviraux; l'apparition de mutants résistants aux inhibiteurs et la grande variabilité génomique au sein d'un même groupe de virus sont en effet fréquemment rencontrées. La découverte des structures cristallographiques des protéines impliquées dans la synthèse de la structure-coiffe de divers agents pathogènes s'avèrera donc un outil important pour le développement d'agents thérapeutiques.

L'ARN 5'-triphosphatase de la levure constitue une cible de choix pour le développement d'inhibiteurs antifongiques puisque: (1) l'activité ARN 5'-triphosphatase est essentielle pour la survie de la levure; (2) la structure primaire ainsi que la structure tridimensionnelle de la protéine semble conservée chez les différentes espèces de levures (*S. cerevisiae*, *C. albicans*, *S. pombe*); (3) la structure et les mécanismes utilisés par la protéine sont différents de ceux retrouvés chez les

mammifères; (4) et, enfin, les cellules de mammifères ne semblent pas posséder de protéines homologues de l'ARN 5'-triphosphatase de la levure. Le traitement d'infections fongiques opportunistes, telles que celles causées par *C. albicans* chez des patients immunodéprimés, pourrait donc éventuellement être envisagé. Diverses stratégies visant à inhiber spécifiquement l'activité de l'ARN 5'-triphosphatase de la levure, tout en évitant d'agir non seulement sur l'ARN 5'-triphosphatase cellulaire mais également sur les autres phosphatases présentes chez l'homme, sont désormais envisageables. Ainsi, l'inhibition de l'activité intrinsèque de la protéine et/ou l'inhibition de l'interaction ARN 5'-triphosphatase-ARN guanylyltransférase comptent parmi les cibles potentielles envisageables. L'analyse de la structure tridimensionnelle de l'ARN 5'-triphosphatase de *S. cerevisiae* a d'ailleurs révélé que la protéine est constituée de huit feuillets  $\beta$  qui contribuent à former un tunnel caractéristique à l'intérieur duquel est retrouvé le phosphate  $\gamma$  et le cation (manganèse) essentiel à l'activité de la protéine. Or, il a récemment été démontré que les résidus importants pour l'activité enzymatique de la protéine, préalablement identifiés par des travaux de mutagenèse dirigée, sont tous dirigés vers l'intérieur de ce tunnel et interagissent avec le phosphate ou l'ion manganèse essentiel à l'activité de la protéine (figure 4) [25, 26]. De plus, l'analyse de diverses mutations de l'ARN 5'-triphosphatase de *C. albicans* suggère fortement que la protéine de cet organisme adopte une structure tridimensionnelle comparable à celle retrouvée chez *S. cerevisiae*. Notons par ailleurs que le domaine d'interaction entre l'ARN 5'-triphosphatase et l'ARN guanylyltransférase de *S. cerevisiae* a été récemment identifié. Ce domaine de 21 acides aminés, situé à la surface de l'ARN 5'-triphosphatase, est également conservé dans l'ARN 5'-triphosphatase de *C. albicans*. L'introduction de mutations dans ce domaine inhibe d'ailleurs spécifiquement la liaison de l'ARN 5'-triphosphatase à l'ARN guanylyltransférase dans des essais d'arrimage *in vitro* [37]. De plus, il a été démontré que la présence de ces mutations inhibe la survie de *S. cerevisiae*. En somme, le

développement de composés pouvant se loger spécifiquement dans le tunnel caractéristique des ARN 5'-triphosphatases de levures et/ou le développement de composés pouvant inhiber la liaison de l'ARN 5'-triphosphatase à l'ARN guanylyltransférase pourrait permettre de combattre certaines infections fongiques opportunistes. Ce type d'inhibition pourrait constituer une alternative aux traitements antifongiques actuels (azoles, amphotéricine B) qui reposent essentiellement sur des composés affectant la perméabilité cellulaire et possédant de nombreux inconvénients (résistance, toxicité). Tout récemment, des études ont démontré que le complexe enzymatique responsable de la synthèse de la structure-coiffe de la levure peut être substitué en totalité par les gènes correspondants de cellules de mammifères [38]. Cet élégant procédé a d'ailleurs permis d'appuyer le modèle limitant la synthèse exclusive de la structure-coiffe aux ARNm syn-

thétisés par l'ARN polymérase II. L'établissement de ces souches de levures isogéniques possédant des gènes étrangers d'autres organismes (levures, mammifères, virus) nécessaires à la synthèse de la structure coiffe constitue un modèle permettant l'identification de composés capables d'inhiber spécifiquement la formation de la structure-coiffe chez divers organismes ■

## RÉFÉRENCES

1. Reddy R, Singh R, Shimba S. Methylated cap structures in eukaryotic RNAs: structure, synthesis and functions. *Pharmacol Ther*1992; 54: 249-67.
2. Moss B, Gershowitz A, Wei CM, Boone R. Formation of the guanylated and methylated 5'-terminus of vaccinia virus mRNA. *Virology*1976; 72: 341-51.
3. Furuichi Y, Muthukrishnan S, Tomasz J, Shatkin AJ. Mechanism of formation of reovirus mRNA 5'-terminal blocked and methylated sequence m<sup>7</sup>GpppG<sup>m</sup>pC. *J Biol Chem* 1976; 251: 5043-53.

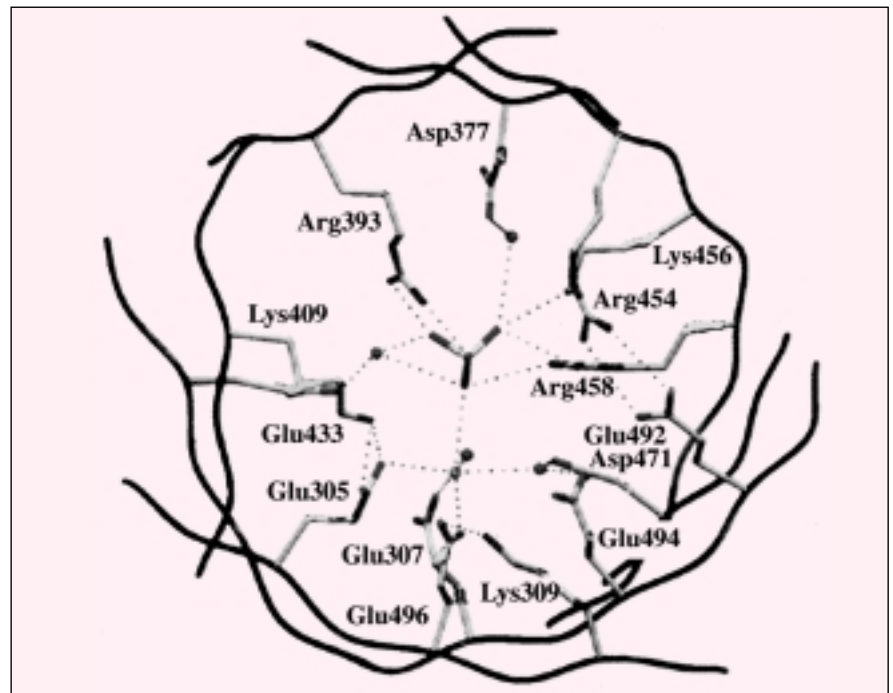


Figure 4. **Structure tridimensionnelle de l'ARN 5'-triphosphatase (Cet1) de *S. cerevisiae*.** L'analyse de la structure de la protéine Cet1 révèle la présence de huit feuillets  $\beta$  qui contribuent à former un tunnel caractéristique à l'intérieur duquel se trouvent le phosphate  $\gamma$  (au centre du tunnel) et le cation (manganèse) essentiel à l'activité de la protéine. Les résidus importants pour l'activité enzymatique de la protéine, préalablement identifiés par des travaux de mutagenèse dirigée, sont dirigés vers l'intérieur de ce tunnel et interagissent avec le phosphate  $\gamma$  ou l'ion manganèse.

## RÉFÉRENCES

4. Perry RP. RNA processing comes of age. *J Cell Biol* 1981; 91: 28-38.
5. Mizumoto K, Kaziro Y. Messenger RNA capping enzymes from eukaryotic cells. *Prog Nucleic Acids Res Mol Biol* 1987; 34: 1-28.
6. Mao X, Shuman S. Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia virus capping enzyme D1 subunit is stimulated by the D12 subunit. Identification of amino acid residues in the D1 protein required for subunit association and methyl group transfer. *J Biol Chem* 1994; 269: 24472-9.
7. Banerjee A. 5'-terminal cap structure in eukaryotic messenger ribonucleic acids. *Microbiol Rev* 1980; 44: 175-205.
8. Both GW, Furuichi Y, Muthukrishnan S, Shatkin AJ. Effect of 5'-terminal structure and base composition on polyribonucleotide binding to ribosomes. *J Mol Biol* 1976; 104: 637-58.
9. Rozen F, Ederly K, Meerovitch K, Dever T E, Merrick WC, Sonenberg N. Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol Cell Biol* 1990; 10: 1134-44.
10. Furuichi Y, LaFinandra A, Shatkin AJ. 5'-terminal structure and mRNA stability. *Nature* 1977; 266: 235-9.
11. Beelman CA, Parker R. Degradation of mRNA in eukaryotes. *Cell* 1995; 81: 179-83.
12. Lewis JD, Izaurralde E. The role of the cap structure in RNA processing and export. *Eur J Biochem* 1997; 247: 461-9.
13. Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj JW. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev* 1996; 10: 1683-98.
14. Reddy R, Singh R, Shimba S. Methylated cap structures in eukaryotic RNAs: structure, synthesis and functions. *Pharmacol Ther* 1992; 54: 249-67.
15. Wang SP, Deng L, Ho CK, Shuman S. Phylogeny of mRNA capping enzymes. *Proc Natl Acad Sci USA* 1997; 95: 73-8.
16. Mao X, Schwer B, Shuman S. Yeast mRNA cap methyltransferase is a 50-kDa protein encoded by an essential gene. *Mol Cell Biol* 1995; 15: 4167-74.
17. Shibagaki Y, Itoh N, Yamada H, Nagata S, Mizumoto K. mRNA capping enzyme: isolation and characterization of the gene encoding mRNA guanylyltransferase subunit from *Saccharomyces cerevisiae*. *J Biol Chem* 1992; 267: 9521-8.
18. Tsukamoto T, Shibagaki Y, Imajoh-Ohmi S, et al. Isolation and characterization of the yeast mRNA capping enzyme  $\beta$  subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability. *Biochem Biophys Res Commun* 1997; 239: 116-122.
19. Luongo CL, Contreras CM, Farsetta DL, Nibert ML. Binding site for S-adenosyl-L-methionine in a central region of mammalian reovirus lambda2 protein. Evidence for activities in mRNA cap methylation. *J Biol Chem* 1998; 273: 23773-80.
20. Vasiljjeva L, Merits A, Auvinen P, Kaariainen L. Identification of a novel function of the alphavirus capping apparatus-RNA triphosphatase activity of Nsp2. *J Biol Chem* 2000; 275: 17281-7.
21. Bisailon M, Lemay G. Characterization of the reovirus lambda1 protein RNA 5'-triphosphatase activity. *J Biol Chem* 1997; 272: 29954-7.
22. Ramadevi N, Burroughs NJ, Mertens PP, Jones IM, Roy P. Capping and methylation of mRNA by purified recombinant VP4 protein of bluetongue virus. *Proc Natl Acad Sci USA* 1998; 95: 13537-42.
23. Ho CK, Pei Y, Shuman S. Yest and viral RNA 5' triphosphatase comprise a new nucleoside triphosphatase family. *J Biol Chem* 1998; 273: 34151-6.
24. Takagi T, Moore CR, Diehn F, Buratowski S. An RNA 5'-triphosphatase related to the protein tyrosine phosphatases. *Cell* 1997; 89: 867-73.
25. Pei Y, Lehman K, Tian L, Shuman S. Characterization of *Candida albicans* RNA triphosphatase and mutational analysis of its active site. *Nucleic Acids Res* 2000; 28: 1885-92.
26. Bisailon M, Shuman S. Structural analysis of the active site of the *Saccharomyces cerevisiae* RNA triphosphatase. *J Biol Chem* 2001 (sous presse).
27. Mizumoto K, Kaziro Y. Messenger RNA capping enzymes from eukaryotic cells. *Prog Nucleic Acids Res Mol Biol* 1987; 34: 1-28.
28. Roth MJ, Hurwitz J. RNA capping by the vaccinia virus guanylyltransferase. Structure of enzyme-guanylate intermediate. *J Biol Chem* 1984; 259: 13488-94.
29. Fresco LD, Buratowski S. Active site of the mRNA-capping enzyme guanylyltransferase from *Saccharomyces cerevisiae*: similarity to the nucleotidyl attachment motif of DNA and RNA ligases. *Proc Natl Acad Sci USA* 1994; 91: 6624-8.
30. Shuman S, Liu Y, Schwer B. Covalent catalysis in nucleotidyl transfer reactions: essential motifs in *Saccharomyces cerevisiae* RNA capping enzyme are conserved in *Schizosaccharomyces pombe* and viral capping enzymes and among polynucleotide ligases. *Proc Natl Acad Sci USA* 1994; 91: 12046-50.
31. Yue Z, Maldonado E, Pillutla R, et al. Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc Natl Acad Sci USA* 1997; 94: 12898-903.
32. McCracken S, Fong N, Rosonina E, et al. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* 1997; 11: 3306-18.
33. Cho E, Tagaki T, Moore C, Buratowski S. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* 1997; 11: 3319-26.
34. Schwer B, Shuman S. Mutational analysis of yeast mRNA capping enzyme. *Proc Natl Acad Sci USA* 1994; 91: 4328-32.
35. Hakansson K, Wigley D. Structure of a complex between a cap analogue and mRNA guanylyl transferase demonstrates the structural chemistry of RNA capping. *Proc Natl Acad Sci USA* 1998; 95: 1505-10.
36. Lima CD, Wang LK, Shuman S. Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell* 1999; 99: 533-43.
37. Ho CK, Lehman K, Shuman S. An essential surface motif (WAQWK) of yeast RNA triphosphatase formation of the mRNA capping enzyme complex with RNA guanylyltransferase. *Nucleic Acids Res* 1999; 27: 4671-8.
38. Saha N, Schwer B, Shuman S. Characterization of human, *Schizosaccharomyces pombe*, and *Candida albicans* mRNA cap methyltransferase and complete replacement of the yeast capping apparatus by mammalian enzymes. *J Biol Chem* 1999; 274: 16553-62.

## Summary

### The mRNA cap structure

Eukaryotic mRNAs contain a 5'-terminal cap structure consisting of a 7-methylguanosine residue linked to the end of the transcript via a 5'-5' triphosphate bridge. The cap earmarks the mRNA for subsequent processing and nucleocytoplasmic transport, protects the mRNA from degradation and promotes efficient initiation of protein synthesis. A consilience of biochemistry, molecular genetics, and structural biology has illuminated the mechanism of cap synthesis. The expression, purification, and characterization of the proteins of various origins involved in the synthesis of the cap structure revealed that the physical and functional organizations of the capping apparatus differ in significant respects in fungi, metazoan, and viruses. It now appears that a selective inhibition of the enzymes involved in the synthesis of the cap structure could have important therapeutic applications.