




Un inventaire traductionnel du VIH-1 révèle une centaine de cadres de lecture alternative et de nouveaux antigènes

Arnaud Moris¹ , Emiliano Ricci² , Olivier Namy¹ 

¹ Université Paris-Saclay, CEA, CNRS, Institut de biologie intégrative de la cellule (I2BC), Gif-sur-Yvette, France.

² Laboratoire de biologie et modélisation de la cellule, École normale supérieure de Lyon, CNRS UMR 5239, Inserm U1293, université Claude Bernard Lyon 1, Lyon, France.

arnaud.moris@i2bc.paris-saclay.fr

emiliano.ricci@ens-lyon.fr

olivier.namy@i2bc.paris-saclay.fr

► La régulation de la traduction d'un ARN messager (ARNm) détermine non seulement la quantité de protéine produite, mais aussi la nature de cette protéine. Le « profilage ribosomique » (technique Ribo-Seq) et l'immunoprotéomique ont révolutionné notre capacité à étudier, à l'échelle du génome, les mécanismes complexes de la traduction, ainsi que ceux de la présence des antigènes viraux aux lymphocytes [1] (→).

(→ Voir *m/s* n° 10, 2016, page 849

L'application de ces techniques à la virologie a fait progresser non seulement notre compréhension du contrôle de l'expression des protéines virales, mais aussi nos connaissances sur la complexité des génomes viraux grâce à l'identification de nouvelles protéines virales et de nouveaux antigènes.

La technique Ribo-Seq à haut débit permet de répertorier, à l'échelle du transcriptome, les produits de la traduction des ARNm d'une cellule. Cette méthode consiste à isoler et à séquencer les fragments d'ARNm protégés de l'action des nucléases par les ribosomes en cours de traduction (Figure 1A). En cartographiant ces fragments (« empreintes ribosomiques ») sur un génome de référence, il est possible d'identifier avec précision les régions d'ARNm activement traduites, de localiser les sites d'initiation et de terminaison de la traduction, et de quantifier son efficacité pour les différents transcrits (Figure 1B, 1C). Parce que les empreintes ribosomiques présentent une périodicité de

trois nucléotides et que leur positionnement peut être déterminé à la résolution du nucléotide, le profilage ribosomique permet une approche non seulement quantitative mais aussi qualitative de la traduction des ARNm. Cette résolution fine permet de définir le cadre de lecture actif (*translon*, [2]), de détecter d'éventuels glissements de ce cadre (*frameshift*), et d'identifier des cadres de lecture ouverts (*open reading frames* [ORF]) non identifiés précédemment. À ce titre, la technique Ribo-Seq fournit une mesure directe de la traduction *in vivo* et permet de révéler de nombreux événements « non canoniques », telles que l'initiation de la traduction à partir de codons différents du codon AUG, d'un décalage du cadre de lecture (*frameshifting*) programmé, ou la traduction à partir d'un cadre de lecture « alternative » (*alternative reading frame* [ARF]) situé en amont du cadre de référence (*upstream* ORF [uORF]), interne à celui-ci, ou encore le chevauchant mais dans une autre phase de lecture. Ces analyses ont élargi le paysage traductionnel du transcriptome, en montrant que de nombreux petits ORF (*small ORF* [sORF]), qui codent des polypeptides ou miniprotéines de moins de 100 acides aminés, sont activement traduits dans différents contextes cellulaires. Ces sORF peuvent se situer dans toutes les régions du transcrit, que ce soit dans sa région 5' non traduite (*5'untranslated region* [5'UTR]), dans sa région 3'UTR, ou chevauchant les ORF de référence dans des cadres de lecture

alternative. Les uORF en 5'UTR peuvent augmenter ou réduire l'efficacité de la traduction de l'ORF principal en influant sur son initiation ou en provoquant la ré-initiation. Cependant, ils ne représentent qu'une fraction du paysage élargi de la traduction des ARNm mis en lumière par la technique Ribo-Seq. Cette technique a aussi permis d'identifier de nouveaux ORF dans les transcrits de divers virus infectant l'Homme, y compris le coronavirus SARS-CoV-2 (*severe acute respiratory syndrome coronavirus 2*) [3] et le cytomégalo virus (CMV) [4]. Ces ORF peuvent coder des protéines virales qui jouent un rôle dans la répllication du virus, son échappement au système immunitaire, ou la pathogénèse de l'infection. Pour le SARS-CoV-2, l'analyse par la technique Ribo-Seq a révélé l'existence de 23 ORF viraux non repérés précédemment, y compris des uORF et des ARF [3]. Plus récemment, trois équipes de chercheurs ont utilisé cette technique pour caractériser le transcriptome du virus de l'immunodéficience humaine de type 1 (VIH-1), c'est-à-dire l'ensemble des cadres de lecture effectivement traduits en peptides par les ribosomes [5-7]. En utilisant des lymphocytes T CD4⁺ infectés par le virus, ces chercheurs ont montré que le génome du VIH-1 est traduit au-delà des 12 ORF canoniques codant pour les protéines virales déjà caractérisées. Étonnamment, la traduction est particulièrement active dans la région 5'UTR des transcrits du virus, où des uORF chevauchant les trois cadres de lectures possibles



ont été mis en évidence. Les chercheurs ont constaté que la suppression des codons d'initiation des uORF entraîne une augmentation de la traduction de l'ORF principal, ce qui suggère que ces uORF exercent un effet inhibiteur sur la synthèse des protéines virales cano- niques [6]. Néanmoins, ils montrent que cet effet est partiellement supprimé par l'ARN hélicase cellulaire DDX3 (*DEAD box-containing RNA helicase 3*), qui est requise pour la traduction efficace des transcrits du VIH-1. Les chercheurs ont montré, en utilisant la technique Ribo- Seq, que la traduction à partir des uORF est conservée dans d'autres rétrovi- rus : VIH-2, HTLV-1 (virus du lymphome T humain adulte de type 1). Par analogie avec les uORF des ARNm cellulaires, ces résultats suggèrent que les uORF sont des éléments « cis-régulateurs » de la synthèse des protéines virales [6].

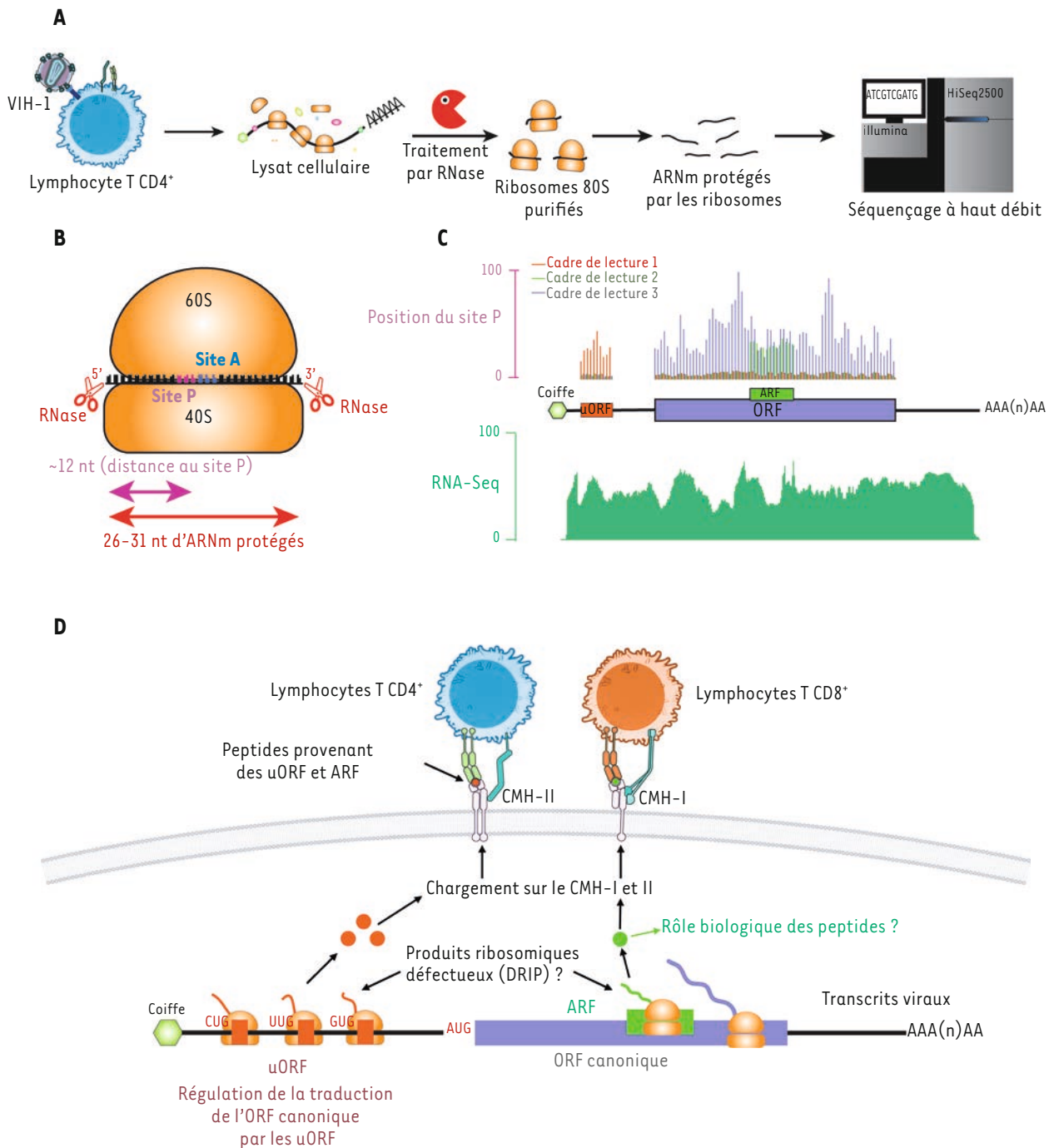
L'analyse du traductome du VIH-1 par la technique Ribo-Seq a aussi révélé que le génome viral héberge au moins 89 ARF répartis sur l'ensemble de celui-ci et chevauchant les ORF « canoniques » [5]. En combinant la technique Ribo-Seq et l'utilisation de drogues bloquant les ribosomes sur les codons d'initiation de la traduction, les chercheurs ont tenté de définir les sites d'initiation de la traduction des ARF nouvellement décrits. Pour 26 ARF, un unique site d'initiation a été identifié, tandis que pour les autres, plusieurs sites d'initiation ont été déli- mités par des zones denses d'empreintes ribosomiques [5]. Ces résultats sug- gèrent que la traduction des ARF en pep- tides pourrait impliquer des mécanismes divers. En effet, pour une grande partie des ORF ou ARF eucaryotes et viraux, la traduction est initiée par des méca- nismes « non canoniques », tels que la *scanning* perméable ou « leaky », la « translecture » du codon stop, le *shun- ting* ribosomique, la ré-initiation, l'IRES (*internal ribosome entry site*), et le décalage du cadre de lecture. En raison de la diversité des sites potentiels d'ini- tiation de la traduction pour la plupart des ARF, les séquences en acide aminés

des peptides synthétisés à partir des ARF du VIH-1 identifiés ont été définies en fonction des codons stop flanquants. Avec ces critères, la longueur de ces peptides variait entre 10 et 101 acides aminés, suggérant que les ARF du VIH-1 correspondent à des sORF. Les séquences d'acides aminés correspondant à ces ARF sont remarquablement conservées parmi les isolats cliniques du VIH-1 : de 48 % à 98 % d'identité parmi les virus de clade B, par exemple. Bien que la plupart des ARF ne soient probablement pas soumis à une pression de sélection évolutive, il convient de noter que les séquences des codons de huit d'entre eux sont plus conservées que celles des ORF qu'ils chevauchent, ce qui pourrait indiquer un rôle biologique pour les pep- tides qu'ils codent [5]. Toutefois, cette conservation des ARF dans les isolats cliniques du VIH-1 pourrait également être imposée par la structure tertiaire de l'ARN viral ou bien par une pression de sélection exercée sur l'ORF codant une protéine virale « vitale ».

Afin de déterminer si les ARF identi- fiés codent effectivement des peptides viraux, les chercheurs ont utilisé deux approches indépendantes et complé- mentaires : d'une part, la détection, parmi les cellules du sang d'individus vivant avec le VIH-1, de lymphocytes T spécifiques des peptides dérivés d'ARF, et d'autre part, l'isolement des peptides dérivés d'ARF liés aux molécules du com- plexe majeur d'histocompatibilité (CMH) par la technique d'immunopeptidomique, qui permet l'identification directe des peptides présentés par les molécules du CMH de classe I et de classe II [5]. Cette technique consiste à isoler, par immu- nopréciptation, les complexes CMH des cellules, à éluer les peptides qui leur sont liés, et à caractériser ces peptides par spectrométrie de masse de haute résolu- tion [8]. Il convient de rappeler que dans les cellules infectées, les peptides pré- sentés aux lymphocytes T par les molé- cules du CMH ne proviennent pas uni- quement de la dégradation de protéines « canoniques » virales, mais proviennent

aussi de polypeptides viraux tronqués ou mal repliés, appelés produits ribo- somiques défectueux (*defective ribo- some products* [DRiP]) [9]. Ce sont des produits de traduction instables, rapi- dement dégradés après leur synthèse, ce qui favorise leur chargement rapide sur des molécules du CMH de classe I. L'analyse des peptides présentés par les molécules du CMH permet donc d'accé- der aux protéines et polypeptides viraux issus des ORF ou des ARF, indépendam- ment de leur demi-vie. En analysant des échantillons de sang provenant de sept personnes vivant avec le VIH-1, les chercheurs ont constaté que les lympho- cytes T ciblent des peptides dérivés de 42 ARF différents, et ont ainsi confirmé que les ARF identifiés par la technique Ribo- Seq codent des polypeptides viraux [5]. Les chercheurs ont aussi identifié, par immunopeptidomique, un peptide dérivé d'ARF du VIH-1 naturellement présenté sur les lymphocytes T CD4⁺ infectés par le virus : ce peptide est traduit à partir d'un ARF, très conservé, qui chevauche la séquence codant le site actif de la trans- criptase inverse du virus [5].

La qualité de l'activation des lympho- cytes T spécifiques du VIH, définie par leur capacité à produire plusieurs cyto- kines ou chimiokines antivirales plutôt que par l'ampleur de leur réponse, a été associée à l'évolution de la maladie [10]. Les chercheurs ont donc étudié *ex vivo* le profil cytokinique des lymphocytes T spécifiques de peptides viraux dérivés d'ARF, et ont montré que ces peptides induisent une réponse polyfonctionnelle des lymphocytes T mémoires qui rappelle les réponses ciblant les épitopes « immu- nodominants » dérivés des ORF cano- niques du virus. L'ARF codant le peptide viral identifié par immunoprotéomique induit également une réponse forte et polyfonctionnelle des lymphocytes T [5]. En raison du petit nombre d'individus étudiés, les chercheurs n'ont pas pu tirer de conclusions quant au rôle de ces lymphocytes dans le contrôle de l'infec- tion. Néanmoins, ils ont montré que les peptides viraux issus de la traduction



des ARF induisent des réponses fonctionnelles efficaces des lymphocytes *T in vivo* au cours d'une infection naturelle par le VIH-1 (Figure 1D). Les ARF du VIH-1 pourraient donc constituer des cibles pour la conception de vaccins et d'immunothérapies contre le virus.

Dans le contexte de l'infection par le VIH-1, la technique Ribo-Seq et l'im-

munoprotéomique ont donc permis d'importantes découvertes concernant la régulation traductionnelle de l'expression des gènes viraux et la réponse immunitaire des lymphocytes T. L'analyse récente du paysage épitranscriptomique [11] (→) viral a par ailleurs révélé l'existence de nombreux sites de

méthylation en position 6 de l'adénosine (m^6A) spécifiques de différentes isoformes des transcrits du VIH-1 [12], ce qui laisse entrevoir de nouvelles intrications dans le contrôle de l'expression du génome de ce virus. ♦

HIV-1 translational inventory reveals about a hundred of alternative ORFs and new antigens



◀ **Figure 1. Analyse globale de la traduction des ARN messagers par profilage ribosomique (Ribo-Seq).** **A.** Les lymphocytes T CD4⁺ infectés par le virus VIH-1 sont incubés avec des inhibiteurs de la traduction des ARN messagers (ARNm), comme le cycloheximide, afin de figer les ribosomes sur les ARNm. Les cellules sont ensuite lysées, et la fraction cytoplasmique, correspondant au compartiment cellulaire où a lieu la traduction des ARNm, est récupérée. Les lysats obtenus sont ensuite séparés en deux fractions. La première est utilisée pour extraire les ARNm et mesurer, par « séquençage massif en parallèle » (*RNA-sequencing*, RNA-Seq), le degré d'expression des gènes correspondants. La deuxième est incubée avec des ribonucléases (généralement la RNase I, qui clive l'ARN simple brin quelle que soit sa séquence nucléotidique) qui vont dégrader toutes les régions d'ARN accessibles à l'exception de celles qui sont associées aux protéines se liant à l'ARN et de celles associées aux ribosomes 80S. Les ribosomes étant plus denses et volumineux que les protéines libres et autres complexes ribonucléoprotéiques, ils peuvent être purifiés dans un gradient de sucrose, sur coussin de sucrose ou par chromatographie d'exclusion. Une fois les ribosomes 80S isolés, les fragments d'ARNm protégés sont récupérés, puis utilisés pour préparer des banques d'ADN complémentaire (ADNc) qui seront séquencées à haut débit. **B.** Les fragments d'ARNm protégés par le ribosome ont généralement une longueur comprise entre 26 et 31 nucléotides. La position du site P du ribosome (qui contient la chaîne polypeptidique associée au dernier acide aminé incorporé) peut être extrapolée en ajoutant un décalage d'environ 10 nucléotides en direction 3' à partir de l'extrémité 5' de la séquence d'ARNm protégée par le ribosome. On peut ainsi déterminer, pour chaque ribosome, son cadre de lecture de l'ARNm en cours de traduction. **C.** Une fois la position de chaque site P alignée sur le transcriptome cellulaire et viral, il est possible d'identifier, pour chaque transcrite, la séquence de codons « principale » ainsi que des séquences codantes « alternatives » (uORF, ARF, etc). **D.** En appliquant le profilage ribosomique aux cellules infectées par le VIH-1, les chercheurs ont pu identifier des événements de traduction « non canonique », conduisant à la production de peptides viraux inconnus jusqu'alors. Certains de ces peptides viraux peuvent être chargés sur les molécules du complexe majeur d'histocompatibilité (CMH) pour être présentés, à la surface des cellules infectées ou de cellules présentatrices d'antigènes, aux lymphocytes T CD4⁺ et CD8⁺. Figure réalisée à l'aide de *NIAID Visual & Medical Arts*.

CONFLITS D'INTÉRÊT

Les auteurs déclarent qu'ils n'ont aucun conflit d'intérêt.

RÉFÉRENCES

1. Blin J, Ricci EP. Le profilage ribosomique : une technique de nouvelle génération pour l'étude de la traduction au cours d'une infection virale. *Med Sci (Paris)* 2016 ; 32 : 849-60.
2. Świrski MI, Tierney JAS, Albà MM, et al. Translons: a single term for translated regions. *Nat Methods* 2025 ; 22 : 2002-6.
3. Finkel Y, Mizrahi O, Nachshon A, et al. The coding capacity of SARS-CoV-2. *Nature* 2021 ; 589 : 125-30.
4. Erhard F, Halenius A, Zimmermann C, et al. Improved Ribo-seq enables identification of cryptic translation events. *Nat Methods* 2018 ; 15 : 363-6.

5. Bertrand L, Nelde A, Ramirez BC, et al. Unveiling conserved HIV-1 open reading frames encoding T cell antigens using ribosome profiling. *Nat Commun* 2025 ; 16 : 1707.
6. Labaronne E, Décimo D, Bertrand L, et al. Non-AUG HIV-1 uORF translation elicits specific T cell immune response and regulates viral transcript expression. *Nat Commun* 2025 ; 16 : 1706.
7. Kibe A, Buck S, Gribling-Burrer A-S, et al. The translational landscape of HIV-1 infected cells reveals key gene regulatory principles. *Nat Struct Mol Biol* 2025 ; 32 : 841-52.
8. Yewdell JW. MHC class I immunopeptidome: Past, present, and future. *Mol Cell Proteomics* 2022 ; 21 : 100230.
9. Yang N, Gibbs JS, Hickman HD, et al. Defining viral defective ribosomal products: standard and alternative translation initiation events generate a

common peptide from influenza A virus M2 and M1 mRNAs. *J Immunol* 2016 ; 196 : 3608-17.

10. Almeida JR, Price DA, Papagno L, et al. Superior control of HIV-1 replication by CD8⁺ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *J Exp Med* 2007 ; 204 : 2473-85.
11. Salvetti A. Épitranscriptome. *Med Sci (Paris)* 2024 ; 40 : 287.
12. Naquin D, Blanchet S, van Dijk E, et al. High-resolution HIV-1 m6A epitranscriptome reveals isoform-dependent methylation clusters and unique 2-LTR transcript modifications. *NAR Genom Bioinform* 2025 ; 7 : lqaf200.



Tarifs d'abonnement m/s - 2026

Abonnez-vous
à médecine/sciences

> Grâce à m/s, vivez en direct les progrès
des sciences biologiques et médicales

Abonnez-vous sur
www.medecinesciences.org

