RNA METABOLISM

Scaling up dissection of functional RNA elements

Translation and mRNA decay are tightly connected processes governing protein production, and their regulation involves an elaborate network of protein factors and sequence elements. A massively parallel RNA-based reporter system now reveals regulatory pathways triggered by 5′ UTR elements and allows dissection of the interplay between translation and mRNA decay.

Marc Bühler and Alex Charles Tuck

As our knowledge of the mRNA lifecycle grows, so too does our appreciation that gene expression is regulated at many levels beyond transcription. In the cytoplasm, the opposing forces of translation and mRNA decay govern protein output. A better understanding of these processes would reveal their contributions to disease and help us engineer improved protein synthesis systems, among many other applications. Furthermore, our existing knowledge of mRNA biology is primarily derived from studies of a few model transcripts and regulatory elements. We must now progress from this blinkered view to one that encapsulates the full diversity and complexity of the transcriptome. In this issue of *Nature Structural & Molecular Biology*, Jia et al. describe a massively parallel synthetic mRNA reporter system that helps them do just this and tease apart diverse pathways that govern mRNA translation and decay[1]. Focusing on the 5′ UTR, they examine the impact of inserting a random 10-nt sequence. Some 10-nt sequences (most of which contain AUG) act as a translation initiation site for an upstream open reading frame (uORF), while others directly regulate mRNA stability or translation of the main ORF. These diverse activities result in a remarkable spectrum of mRNA stabilities and translation efficiencies and led to the discovery of new regulatory elements.

Jia et al. began by designing a sensitive reporter system to quantify uORF and main ORF translation (Fig. 1a). This comprises a GFP coding sequence downstream of a uORF encoding SIINFEKL, a peptide that can be detected by the antibody 25D1 with exquisite sensitivity. The authors then generated a library by replacing the uORF start codon with a random 10-nt sequence. As expected, when the random 10-nt sequence contains a canonical start codon (AUG), the uORF is translated and 25D1 labelling is detected by

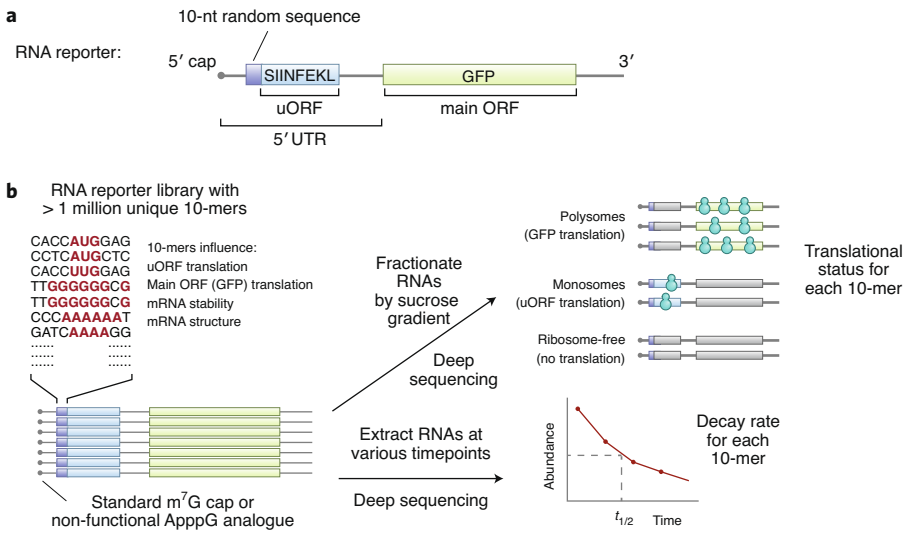

Fig. 1 | Experimental design of the study. a, mRNA reporter testing the effect of a 10-nt 5′ UTR element on translation and decay. **b**, Massively parallel RNA library features and methods used to measure mRNA translation and decay.

fluorescence-activated cell sorting (FACS). At the same time, the GFP signal from the main ORF is reduced, in line with the observation that uORF translation generally represses translation of a downstream ORF (Fig. 2a)[2]. While this concept is well established, the rules governing initiation at non-AUG codons and the influence of the local sequence context are less well defined. To investigate this, Jia et al. refined their system to detect more subtle changes in uORF and main ORF translation (Fig. 1b). Firstly, using in vitro–synthesized reporter RNAs instead of plasmids, they monitored translation shortly after transfecting the library into cells, thus removing several sources of variability. Secondly, as FACS cannot be used to distinguish between different reporter molecules in the same cell, the authors used monosome and polysome extractions to separate mRNAs engaged in uORF and main ORF translation (the uORF can only accommodate one ribosome), respectively. Using this modified approach, it became clear that, while the initiating codon is the main determinant of uORF translation initiation, the sequence context also plays a role.

The real power of this RNA reporter system, however, lies in its ability to dissect the interplay between translation and mRNA decay. Recent studies have found these processes to be intimately coupled, and many cytoplasmic mRNA decay factors bind to ribosomes[3–5]. A consequence of this is that translation regulators often affect mRNA decay and decay factors can affect translation. This makes it hard to identify the primary mechanism of a given regulatory factor. To overcome this challenge, Jia et al. used their RNA library to quantify both translation (by monitoring monosome or polysome association) and mRNA stability (by tracking RNA

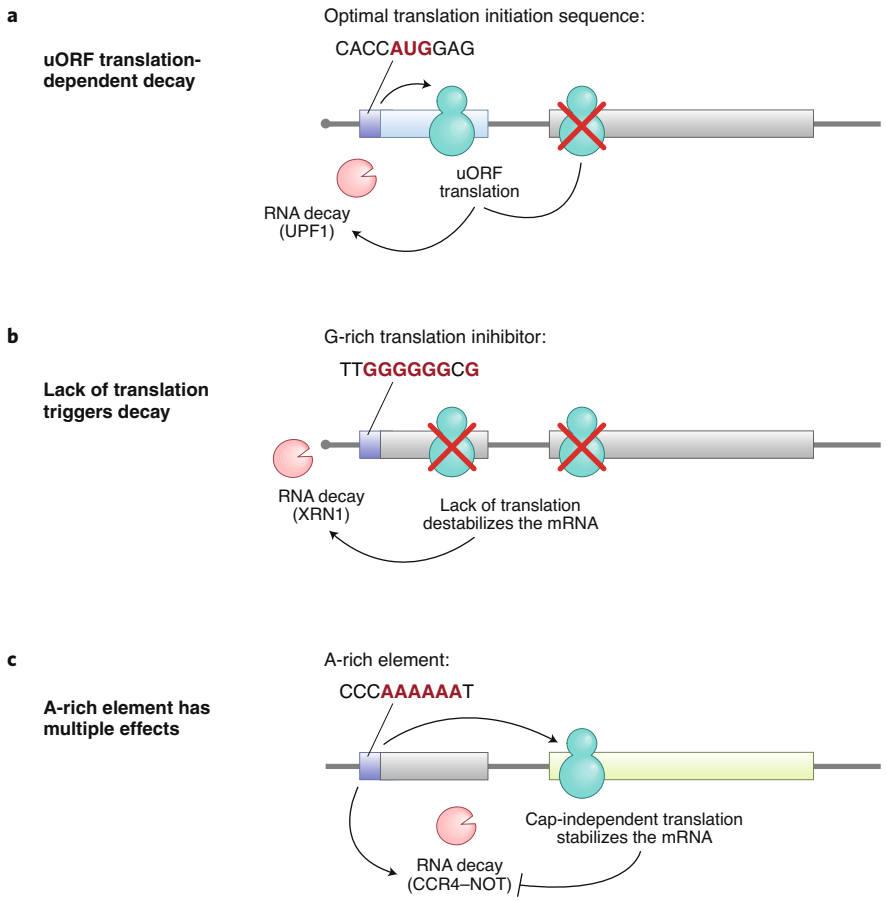

Fig. 2 | Biological insights gained from the study. a, AUG-containing 10-nt sequences permit uORF translation, which destabilizes the mRNA. **b**, G-rich 10-nt sequences inhibit translation, which destabilizes the mRNA. **c**, A-rich elements can either stabilize the mRNA, as a result of promoting cap-independent translation, or act directly to destabilize the mRNA.

abundance after transfection; Fig. 1b). Crucially, they could uncouple mRNA decay from translation by using a non-translatable RNA library incorporating the cap analogue ApppG. These experiments demonstrated that translation influences mRNA stability in two ways: (1) main ORF translation is generally stabilizing, and (2) uORF translation is destabilizing. The ability to easily customize synthetic mRNA reporters makes them a versatile tool to study other modifications. For example, the mechanism by which m^{6}A residues influence mRNA translation and decay and how these are affected by the location or context of the modified residue remains a topic of hot debate[6]. This could be elegantly resolved using synthetic modified mRNA reporters.

Jia et al. also spotted anomalous reporter RNAs that either failed to associate with ribosomes or, at the other end of the spectrum, associated with ribosomes even when lacking a functional cap. Pursuing this further led to the discovery of a GGC motif that inhibits translation and an A-rich element that promotes cap-independent translation (Fig. 2b,c). In both cases, mRNA stability is altered as a consequence of altered translation. Examining the GGC element in more detail, the authors concluded that it inhibits translation via G-quadruplex formation, which is usually minimized in cells by the activity of the helicase DHX36. This adds to a growing body of evidence suggesting that cells actively suppress the formation of RNA structures and other obstacles that otherwise block translation and trigger mRNA decay[7–9].

The A-rich element that Jia et al. discovered also highlights some emerging concepts (Fig. 2c). Firstly, although this element stabilizes mRNAs by stimulating translation, this process is relatively inefficient, and many such mRNA molecules do not engage ribosomes. For these non-translating mRNAs, the A-rich element actually leads to destabilization, which is in agreement with a previous study[10]. Thus, the same element can both stabilize and destabilize an mRNA via mechanisms that are translation-dependent and translation-independent, respectively. Furthermore, stabilizing and destabilizing effects both depend on the poly(A)-binding protein PAPB1, demonstrating that readers of RNA regulatory elements also play opposing roles. Dualities such as this are a recurrent theme in post-transcriptional regulation and highlight how RNA fate decisions involve the integration of many signals and kinetic competition between translation and decay[5].

These results portray post-transcriptional regulation as relatively chaotic, with many sequence elements and readers, each performing diverse roles. However, cells do manage to impart some order. For example, Jia et al. observed three different decay pathways, each of which involves a specific set of decay factors (UPF1, XRN1/DCP2 and CCR4–NOT1; Fig. 2). Other mRNA processes, such as the response to ribosome collisions, also involve dedicated factors and pathways[11]. Specificity is conferred partly by the formation of structural interfaces, and recent data suggest that helicases play a key role in regulating the trafficking and compartmentalization of RNA and thus in controlling post-transcriptional events[12].

The work of Jia et al. and others[13,14] also demonstrates that regulatory interactions occur along the entire length of an mRNA transcript, rather than being concentrated in a few regions. This raises questions of whether and how signals are interpreted differently when located in distinct regions of an mRNA or, indeed, within different mRNAs altogether[5,15,16]. Here, molecular rulers and timers that measure distance from the 5′ cap, poly(A) tail or translation start site may be key, and these are starting to be uncovered[17]. This is another area in which synthetic RNA reporters, with sequence elements or modifications placed at different positions or within different mRNAs, could help break new ground. ❐

Marc Bühler[1,2] ✉ and Alex Charles Tuck[3] ✉

[1]*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland.* [2]*University of Basel, Basel, Switzerland.* [3]*SOPHiA GENETICS SA, Saint Sulpice, Switzerland.*

✉*e-mail: marc.buehler@fmi.ch; atuck@sophiagenetics.com*

Published online: 3 August 2020
https://doi.org/10.1038/s41594-020-0482-9

References

1. Jia, L. et al. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-020-0465-x (2020).

2. Johnstone, T. G., Bazzini, A. A. & Giraldez, A. J. *EMBO J.* **35**, 706–723 (2016).
3. Schmidt, C. et al. *Science* **354**, 1431–1433 (2016).
4. Tesina, P. et al. *Nat. Struct. Mol. Biol.* **26**, 275–280 (2019).
5. Buschauer, R. et al. *Science* **368**, eaay6912 (2020).
6. Zaccara, S. & Jaffrey, S. R. *Cell* **181**, 1582–1595.e18 (2020).
7. Tuck, A. C. et al. *Mol. Cell* **77**, 1222–1236.e13 (2020).
8. D'Orazio, K. N. et al. *Elife* **8**, e49117 (2019).
9. Guo, J. U. & Bartel, D. P. *Science* **353**, aaf5371 (2016).
10. Melo, E. O., de Melo Neto, O. P. & Martins de Sá, C. *FEBS Lett.* **546**, 329–334 (2003).
11. Ikeuchi, K. et al. *EMBO J.* **38**, e100276 (2019).
12. Hondele, M. et al. *Nature* **573**, 144–148 (2019).
13. Meyer, K. D. et al. *Cell* **163**, 999–1010 (2015).
14. Van Nostrand, E. L. et al. *Genome Biol.* **21**, 90 (2020).
15. Mao, Y. et al. *Nat. Commun.* **10**, 5332 (2019).
16. Aeschimann, F. et al. *Mol. Cell* **65**, 476–489.e4 (2017).
17. Bohlen, J., Fenzl, K., Kramer, G., Bukau, B. & Teleman, A. A. *Mol. Cell* https://doi.org/10.1016/j.molcel.2020.06.005 (2020).

Competing interests
The authors declare no competing interests.

VIROLOGY

Hijacking of endocytosis by HIV-1 Nef is becoming crystal clear

The cell surface protein CD4 acts as a coreceptor for incoming HIV particles. However, the expression of CD4 in HIV-producing cells is detrimental to virus propagation and pathogenicity. To solve this issue, the viral accessory protein Nef forces CD4 endocytosis and targets it for lysosomal degradation. Structural elucidation of the AP-2–Nef–CD4 complex shows how Nef connects CD4 to the clathrin endocytic machinery, revealing a potential new target for anti-HIV therapy.

Yunan C. Januário and Luis L. P. daSilva

Since the discovery of the human immunodeficiency virus (HIV) as the causative agent of AIDS (acquired immunodeficiency syndrome) almost 40 years ago, much has been learned regarding the intricate mechanisms by which this virus invades cells, controls cell machinery and subverts antiviral restriction factors. Such knowledge has paved the way for the development of drugs targeting HIV, resulting in remarkable decreases in AIDS morbidity and mortality. However, in the absence of effective vaccines or definitive therapies, viral resistance to currently available drugs as well as their side effects threaten the continued control of AIDS.

A study by Kwon et al.[1] now unveils a potential Achilles heel to pharmacologically target, the viral Nef protein, which is a crucial determinant of HIV pathogenicity. Nef is a nonenzymatic protein encoded in the HIV-1, HIV-2 and simian immunodeficiency virus (SIV) genomes. Despite being dispensable for virus replication in vitro, Nef is critical for disease progression to AIDS, as indicated by studies reporting that individuals infected with Nef-deficient HIV-1 did not develop AIDS for ten years or longer, even without treatment[2,3]. Nef's general strategy to increase viral fitness involves interfering in the intracellular transport of specific host proteins that play roles in the virus replication cycle and host immunity[4,5].

The most studied function of Nef is to remove the CD4 receptor from the surface of HIV-infected cells, a process referred to as CD4 downregulation. CD4 is a surface type I transmembrane glycoprotein expressed by cells such as macrophages and a subset of T lymphocytes (known as helper T cells), key players in innate and adaptive immune responses and the primary targets of HIV. CD4 directly binds the HIV envelope glycoprotein ENV and serves as the primary receptor for the virus. At first glance, the depletion of surface CD4 by HIV may seem paradoxical. However, CD4 downregulation in infected cells promotes HIV-1 propagation by facilitating the release of viral progeny and by avoiding superinfection, which would increase cytopathic effects in host cells[6,7].

Nef associates with cellular membranes via an N-terminal myristoyl group and displays remarkable structural plasticity with extensive disordered regions, enabling it to physically interact with a broad range of host proteins[4,5]. Among Nef's binding partners is the transport vesicle adaptor protein 2 (AP-2), a complex composed of four distinct subunits (α, β2, μ2 and σ2). AP-2 selects transmembrane protein cargo by binding to their cytosolic tail (CT) sorting signals, which are typically tyrosine (YXXθ)- or dileucine ([E/D]XXXL[L/I])-based motifs (where X is any amino acid, and θ is a bulky hydrophobic residue) and are recognized by the μ2 or α–σ2 subunits of AP-2, respectively. AP-2 also binds clathrin at the plasma membrane, thus coordinating cargo selection with the assembly and budding of endocytic clathrin-coated vesicles (CCVs)[8].

A substantial body of evidence indicates that Nef co-opts AP-2 to force CD4 endocytosis, which subsequently directs CD4 to lysosomes for degradation[4,5] (Fig. 1a). Nef molecules carry a highly conserved EXXXLL motif that is essential for CD4 downregulation[4,5]. Consistent with this, reducing cellular levels of clathrin[9] or AP-2[9,10] impairs Nef-induced surface depletion of CD4. Nef has been shown to bind the α–σ2 subunits of AP-2, and mutational analysis mapped the binding sites on both proteins[9,11–14]. In a major contribution to characterizing Nef's interaction with AP-2, Ren et al.[13] determined the crystal structure of the α–σ2–Nef complex and proposed a model in which Nef binds to an 'open' plasma-membrane-associated form of AP-2 to incorporate surface CD4 molecules into nascent CCVs (Fig. 1a).

However, the structural basis of CD4 sequestration by Nef–AP-2 and how the other two AP-2 subunits, μ2 and β2, participate in the AP-2–Nef–CD4 tripartite complex remained unknown. The study by Kwon et al. now helps answer those questions by presenting the crystal structure of the AP-2 tetramer complexed with Nef and the CT of CD4 (Fig. 1b–e). In addition to confirming Nef's robust binding to the dileucine-motif binding site in α–σ2, the authors also note that α–σ2's association with Nef partially