How Widespread is Metabolite Sensing by Ribosome-Arresting Nascent Peptides?

Britta Seip and C. Axel Innis

University of Bordeaux, Institut Européen de Chimie et Biologie, 33607 Pessac, France
Institut National de la Santé et de la Recherche Médicale, Inserm U1212, 33076 Bordeaux, France
Centre National de la Recherche Scientifique, CNRS UMR 5320, 33076 Bordeaux, France

Correspondence to C. Axel Innis: Institut Européen de Chimie et Biologie (IECB), 2 rue Robert Escarpit, 33607 Pessac, France. axel.innis@inserm.fr
http://dx.doi.org/10.1016/j.jmb.2016.04.019
Edited by Daniel N. Wilson

Abstract

In order to colonize a niche and compete for scarce resources, microorganisms have evolved means to adjust the expression levels of their biosynthetic enzymes in response to the changing levels of metabolites available to them. To do so, they often rely on transcription factors or structured RNAs that directly sense the concentration of metabolites and turn genes on or off accordingly. In some instances, however, a metabolite can be sensed by an actively translating ribosome bearing a nascent polypeptide whose specific amino acid sequence interferes with translation. These “arrest peptides” lead to the formation of stalled ribosome nascent chain complexes on the mRNA that can regulate the expression of downstream genes through transcriptional or translational mechanisms. Although this process was discovered over three and a half decades ago, the extent to which arrest peptides regulate gene expression in response to cell metabolites is unknown. Here, we examine the physical constraints imposed by the ribosome on peptide-mediated ligand sensing and review attempts to assess the diversity of arrest peptides to date. In addition, we outline a possible way forward to establish how pervasive metabolite sensing by arrest peptides is in nature.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Bacterial metabolism is a complex, highly regulated process, in which a myriad of small molecules is either imported from the surroundings into the cell, or synthesized from basic precursors. In order to maintain constant levels of metabolites inside their cytoplasm, bacteria often rely on feedback loops, such that the expression of key enzymes or transporters is regulated by the end products of the corresponding metabolic pathway. This enables bacteria to adapt rapidly to changing environmental conditions by directing limited resources to where they are needed the most. For this to work, bacteria must not only be able to sense the concentration of specific metabolites, but should also possess mechanisms that turn gene expression on or off in response to these small molecules.

This perspective focuses on a form of regulation of gene expression known as nascent chain-mediated translational arrest, which takes place when a partially synthesized “arrest peptide” pauses or halts its own translation by interacting with the polypeptide exit tunnel of the ribosome [1–3]. Arrest can occur during the elongation phase of translation by inhibiting peptide bond formation or during termination by preventing peptide release from the ribosome. Such process depends primarily upon the amino acid sequence of the arrest peptide but often requires the direct sensing of a small molecule by the ribosomal exit tunnel containing the newly synthesized nascent chain. Ribosomes that translate arrest sequences in the presence of inducing concentrations of such ligands become trapped on the mRNA as inactive ribosome-nascent chain complexes (RNC), which, in turn, regulate the
Fig. 1. The ribosomal exit tunnel and metabolites sensed by arrest peptides. (a) Different regions of the ribosomal exit tunnel and the various biological processes that take place within them, including (1) peptide bond formation, (2) translation inhibition mediated by ribosome-targeting antibiotics or (3) antimicrobial peptides, (4) nascent chain-mediated translational arrest, and (5) secondary or (6) tertiary structure formation. The nascent polypeptide inside the ribosomal exit tunnel is shown in red. (b) Metabolites and drugs that are known to induce nascent chain-mediated translational arrest. Of these, only tryptophan and erythromycin have been shown to directly interact with the nascent peptide.
expression of downstream genes by acting upon the transcriptional or translational machineries. Since ligand recognition by an RNC is mechanistically distinct from the processes that regulate gene expression once a ribosome has stalled on the mRNA, we will only be dealing with the sensing event here.

Although the existence of arrest peptides was postulated more than 35 years ago [4,5] and their involvement in several metabolic processes is well documented, the extent to which they are able to sense different types of small molecules in nature is largely unknown. In order to address this issue, we will begin by discussing how the physical constraints imposed by the ribosome could shape ligand recognition by arrest peptides. We will then review previous attempts to measure the incidence and diversity of arrest peptides and highlight the benefits and limitations of the existing methods. Since this perspective is forward looking, we will propose one possible approach to reveal how pervasive nascent chain-mediated translational arrest is as a form of metabolic gene regulation.

Three and a Half Decades of Arrest Peptides

The notion that a translating ribosome might be able to detect a small molecule was proposed as early as 1980 by the Weisblum [5] and Dubnau [4] laboratories. Both groups discovered that a short leader region coding for a 19-aa peptide is needed to induce the expression of the macrolide antibiotic resistance gene _ermC_ in response to sublethal concentrations of the antibiotic erythromycin, and that regulation is exerted at a post-transcriptional level. A model was therefore proposed, whereby ribosomes translating the leader peptide _ErmCL_ stall in the presence of erythromycin, thus altering the conformation of the mRNA immediately upstream of _ermC_ and inducing its expression [4,5]. Drug-dependent ribosome stalling was later proven to be the key determinant for this new form of translational control [6], and similar mechanisms were proposed for the regulation of other genes that confer resistance against macrolide antibiotics [7] or against the peptidyl transferase inhibitor chloramphenicol [8].

Although one might argue that these peptides are not true metabolite sensors given that the drug can bind to the ribosome in the absence of a nascent chain, these studies showed for the first time that a ribosome could integrate signals encoded in the amino acid sequence of a nascent peptide and in the structure of a small ligand to alter its activity. The discovery of other arrest peptides that could either sense metabolites or cause stalling on their own extended over the course of the next three decades. Seminal work by the Yanofsky group revealed that stalled RNCs containing the TnaC peptide regulate tryptophan catabolism in γ-proteobacteria, such as _Escherichia coli_, by sensing the intracellular concentration of this amino acid [9]. Processes involving the sensing of cell metabolites by arrested RNCs were also discovered in eukaryotes and include the arginine attenuator peptide [10], CGS1 [11], and a peptide encoded by the _AdoMet-DC_ upstream open reading frame [12]. Some bacterial regulatory peptides, such as _SecM_ [13], _MifM_ [14], and _VemP_ [15], are capable of triggering translational arrest on their own but are beyond the scope of this perspective.

Physical Constraints Imposed by the Ribosome on Metabolite-Sensing Arrest Peptides

Arrest peptides operate within the confined environment of the ribosomal exit tunnel, a long cavity spanning the large subunit of the ribosome, through which nascent peptides must transit in order to reach the intracellular milieu (Fig. 1a) [16,17]. The exit tunnel is approximately 100 Å in length, with a diameter between 10 and 20 Å, and can be broadly divided into three sectors [18]: (i) the upper/central tunnel that contains the peptidyl transferase center and the tunnel constriction formed by the extensions of ribosomal proteins L22 and L4 in bacteria, where several classes of antibiotics [19,20], natural antimicrobial peptides [21–23], and all arrest peptides characterized to date exert their effects [1]; (ii) the lower tunnel where the nascent peptide may begin to acquire its secondary structure [18,24,25]; and (iii) the vestibule where small elements of tertiary structure may form before the nascent peptide exits the ribosome [26,27]. Despite the generally high degree of conservation of the upper/central section, the exit tunnels of ribosomes from different species, kingdoms, or even organelles can display significant variations in their sequence and structure [28]. As a result, arrest peptides may be specific for organisms that are phylogenetically not too distantly related [1,14], even though several of them are also known to induce translational arrest on ribosomes from other species [2].

The range of ligands currently known to induce nascent chain-mediated translational arrest in bacteria or in eukaryotes includes amino acids [9,10] and their derivatives [11], polyamines [12,29], and antibiotics that target the large subunit of the ribosome [6,8] (Fig. 1b). So far, only two of these have been shown to interact directly with arrest peptides inside the ribosomal exit tunnel (erythromycin and L-tryptophan), although it is likely that this holds true for other ligands as well. While there
appears to be a marked preference for ligands with delocalized electron systems that offer the possibility of \(\pi\)-stacking (L-tryptophan, L-arginine, S-adenosylmethionine, chloramphenicol) or large hydrophobic surfaces that provide an extended interaction surface (erythromycin and other antibiotics from the macrolide family), no obvious restrictions are apparent in terms of their chemical composition. Given that a nascent polypeptide occupies a large fraction of the volume available within the exit tunnel, it is likely that the size of a ligand will restrict its ability to coexist with an arrest peptide present inside this confined environment. Nevertheless, even large molecules like erythromycin (\(>700\) Da) are known to interact with the nascent peptide in the upper central region of the tunnel and trigger peptide-mediated arrest [30–32].

Structural information on the process of nascent chain-mediated translational arrest comes exclusively from the analysis of stalled RNCs by cryo-electron microscopy [30,31,33–37], with the highest resolutions of 3.8 Å, 3.9 Å, and 3.3–3.7 Å reported to date for arrested TnaC–70S [35], ErmCL–70S [30], and SecM–70S [37] complexes, respectively. While this level of detail is sufficient to show that arrest peptides generally inhibit the ribosome by tampering with the geometry of the peptidyl transferase center, the current resolution precludes a detailed analysis of the intermolecular contacts between the arrest peptide, its ligand, the ribosome and ordered ions or solvent molecules. Nevertheless, the location of ligand molecules within arrested RNCs is either known or inferred for two biological processes: (i) erythromycin-dependent stalling of ribosomes translating the ErmBL [31] or ErmCL [30] leader peptides, which results in the activation of downstream genes whose products confer resistance to macrolide antibiotics; and (ii) translational arrest by the TnaC leader peptide in response to elevated levels of L-tryptophan inside the cell, which leads to the induction of the tryptophanase operon in \(\gamma\)-proteobacteria [35]. In both instances, the ligands bind deep within the upper central tunnel, where they are completely surrounded by the ribosome and the arrest peptide. Whether all functional groups from the ligand are recognized is not clear at present, but there appears to be a certain degree of structural plasticity. Indeed, it is known that TnaC-dependent arrest can be induced by a tryptophan analog [38] and that the erythromycin-derivatives oleandomycin and azithromycin can both trigger ErmBL-dependent translational arrest [31].

From the structures of stalled ErmBL/CL [30,31] and TnaC [35] RNCs available to date, it appears that ligand binding may take place at different stages during the synthesis of the nascent peptide. In the case of the ErmBL–70S [31] and ErmCL–70S [30] complexes, the binding of erythromycin to the exit tunnel is known to occur in the absence of the nascent peptide, and the conformation of the drug inside the ErmBL/CL-arrested RNCs is nearly identical to that seen in a non-translating ribosome [39,40]. Given the high affinity of erythromycin for the large ribosomal subunit, this indicates that the binding of the drug to the ribosome most likely precedes the assembly of the translation initiation complex on the mRNA and that the drug-encumbered tunnel merely provides a new surface with which the tRNA-linked nascent peptide can interact. On the other hand, it has been suggested that binding sites for two molecules of L-tryptophan are created when the N-terminal portion of the nascent TnaC peptide makes specific contacts with the ribosome at the tunnel constriction [35]. These initial interactions could cause a slowdown in the rate of polypeptide egress through the tunnel, which, coupled to the continued synthesis of the TnaC peptide at the peptidyl transferase center, would lead to a compaction of the nascent chain to form the first L-tryptophan binding site. This, in turn, could create a binding site higher in the tunnel for a second molecule of L-tryptophan, whose binding would ultimately result in translational arrest. However, additional evidence is needed to test this kinetic model and the interaction of two tryptophan molecules with the nascent TnaC peptide still needs to be confirmed experimentally using other techniques.

How one or more ligand molecules might reach their binding sites in RNCs such as the TnaC–70S complex is unclear at this point. One possibility is that these molecules are already present within small hydrophobic cavities that line the exit tunnel prior to the arrival of the nascent peptide. This would be in agreement with the suggestion that certain regions of the exit tunnel are thought to have a marked propensity for interacting with small molecules, including amino acids [41]. These regions include the tunnel constriction or the crevice where an incoming aminoacyl moiety binds following accommodation of a tRNA into the A-site, as suggested by the structures of numerous antibiotics and several antimicrobial peptides bound to the ribosome [21,23]. Another possibility is that tryptophan molecules enter the tunnel through the A-site left vacant between each elongation cycle and reach their binding pockets as soon as these are formed by the elongating peptide. However, the rotational dynamics and dielectric behavior of water inside the exit tunnel differ from those of bulk solvent, resulting in lowered solvent diffusion rates that could render this latter approach more difficult [42]. Diffusion rates inside the tunnel may thus also be important in determining the types of ligands that can be sensed by RNCs.

To summarize, the available biochemical and structural data give us no reason to believe that the types of ligands sensed by arrest peptides should be
Fig. 2. Current methods for identifying arrest sequences. Genetic selection methods developed by Buskirk and coworkers based on the (a) kanRΔ18 [45] or (b) HIS3 [46] selection markers. Note that both methods depend on the rescue of stalled ribosomes by the tmRNA/SmpB system of E. coli. (c) The ribosome profiling technique, where deep sequencing of nuclease-protected mRNA fragments (ribosome footprints) reveals the position of each ribosome with nucleotide resolution [43]. (d) iNP [44] can detect and classify different translation intermediates through the comparative electrophoretic analysis of 35S-Met pulse-labeled translation products left untreated, digested with RNase or treated with the antibiotic puromycin. These translation intermediates are indicative of paused or stalled ribosomes.
limited by their chemical composition, provided that their size allows them to coexist with the nascent peptide inside the ribosomal exit tunnel. The question is whether this apparent lack of restrictions also holds true in nature. As we shall see, the few techniques available to identify arrest peptides, although very elegant, are limited in their scope. We therefore still lack the necessary tools to systematically tackle this problem on a larger scale.

Current Methods for the Identification of Arrest Peptides

Experimental techniques for the identification of nascent peptides that block the ribosome fall into two categories. Methods involving genetic selection rely on the expression of random DNA libraries fused to reporter genes inside a bacterial host, such that the identification of arrest sequences is achieved by establishing a direct link between cell survival and the stalling of an RNC on the mRNA. Approaches based on profiling include the widely used ribosome profiling technique to identify naturally occurring peptide sequences that induce ribosome stalling [43], as well as the recently published integrated in vivo/vitro nascent chain profiling (iNP) technique, which directly monitors pauses in translation on the basis of peptidyl-tRNA profiles [44]. These last two techniques have revealed that translation undergoes frequent pausing, which, unexpectedly, cannot be explained by codon usage alone.

Genetic selection approaches

Two methods have been developed that feature a selection protocol based on the rescue of stalled RNCs by the tmRNA/SmpB system [45,46]. This ribonucleoprotein complex normally releases bacterial ribosomes that have become trapped on a non-stop mRNA and adds an SsrA tag [(A)ANDE-NYALAA] to the C terminus of the nascent peptide to target it for degradation. For this to occur, the mRNA on which the stalled RNC is located must be trimmed up to the site of stalling by various intracellular ribonucleases, thus freeing the A-site for the tmRNA complex. As we shall see shortly, this detail is significant when selecting for arrest peptides.

The first method that relies on the tmRNA/SmpB system makes use of a plasmid-encoded DNA library in which a C-terminally truncated kanamycin resistance gene (kanR) has six random codons fused to its end [45] (Fig. 2a). Expression of the kanR gene inside a bacterial host leads to the formation of stalled RNCs when these last few codons encode an arrest peptide. Stalled RNCs are then rescued by a modified tmRNA that adds the remainder of the KanR sequence instead of a degradation tag to the nascent polypeptide, resulting in the expression of a functional protein and cell survival in the presence of kanamycin. Three classes of arrest peptides were identified using this approach: (i) sequences encoding C-terminal proline residues, which are now known to cause particular problems for peptide bond formation on the ribosome [47–49]; (ii) peptides with an amino acid sequence resembling that of the arrest peptide SecM; and (iii) the novel stalling sequence FXXYXIPP, which combines C-terminal prolines with additional N-terminal residues to induce strong translational arrest, both in vivo and in vitro. It should be noted, however, that many of the selected sequences triggered rescue by the tmRNA/SmpB system for reasons other than peptide-mediated arrest (i.e., stable mRNA secondary structures or rare codons).

Given that arrest sequences that inactivate the KanR protein or feature more than 6 aa would have been missed by the above method, a second method was developed that uses bacterial cells, which cannot synthesize the amino acid histidine unless a HIS3 gene is activated [46] (Fig. 2b). This gene is transcribed from a weak promoter and its expression is dependent on the active recruitment of RNA polymerase to an operator located upstream of HIS3 by the DNA-binding protein cl. A random sequence of 20 codons was fused to the C terminus of the full-length cl protein, thus ensuring that its structure and function remain unaltered. Ribosomes undergoing arrest during translation of this random sequence become targets for the tmRNA/SmpB system, which adds a modified SsrA tag [(A)ANDE-NYALDD] to the C terminus of cl before releasing it into the cytosol. Tagged cl then binds to its operator upstream of the HIS3 promoter and recruits, via its SsrA tag, a fusion between SspB and RNA polymerase to activate transcription of HIS3. Importantly, the cl protein featuring a mutated tag is not targeted for degradation. As a result, only those cells carrying a plasmid-encoded ribosome arrest sequence can grow on selective medium lacking histidine. The relatively high survival rate following selection (1/10,000 cells) made it possible to identify and subsequently validate 20 sequences that did not match common arrest motifs, such as poly-Pro or Pro-Stop.

Although the methods described above have led to the identification of novel arrest peptides that operate within an in vivo context, they have some limitations that restrict their general applicability. One of these is their reliance on the tmRNA/SmpB system, which means that only stalling events that are released by tmRNA/SmpB can be detected. However, it is known that arrest peptides such as TnaC or SecM are not substrates for tmRNA/SmpB [50,51], which requires a vacant and fully functional A-site to operate. Given that many arrest peptides are known to block the A-site at the peptidyl transferase center (as shown by increased resistance to the A-site substrate mimic puromycin [9,52,53] and by structural studies [30,31,35–37]), it...
appears likely that a large fraction of arrest peptides will yield stalled RNCs that are impervious to the tmRNA/SmpB system. Using these methods to identify arrest peptides that sense small molecules is further complicated by the fact that the molecule of interest must be able to reach the ribosomes inside the cytoplasm. Problems such as small molecule uptake, efflux, or degradation are likely to be encountered with many of the ligands one might like to test.

**Profiling approaches**

While the issue of small molecule availability will remain problematic for any method carried out in vivo, issues associated with the tmRNA/SmpB system can be overcome through the use of ribosome profiling, a high-throughput method for obtaining information on the position of all translating ribosomes on the mRNA at nucleotide resolution. According to this technique, an outstanding readout along an mRNA molecule indicates ribosomal pausing at this position (Fig. 2c). Recent work employing ribosome profiling has led to the identification of peptide sequences that cause arrest in the presence of macrolide antibiotics in the gram-negative bacterium *E. coli* (erythromycin and telithromycin) [54] and in the gram-positive bacterium *Staphylococcus aureus* (azithromycin) [55]. As the macrolide-sensing Erm peptides carry an arrest motif close to their N terminus, high ribosome densities were expected near the 5′ ends of mRNAs. An increased appearance of arrested ribosomes in this region was indeed observed in the presence of erythromycin, but not in cells treated with telithromycin or azithromycin. Instead, these ligands caused ribosomal arrest at various locations over the entire mRNA coding regions, in particular at sites that were rich in prolines and charged residues. The most prevalent arrest sequence identified in both organisms is (R/K)x(R/K), which is highly similar to the RLR motif found within the erythromycin-dependent arrest peptide ErmDL. Sequences that caused arrest in the presence of telithromycin in *E. coli* could, in most cases, be further simplified to [+x]+, whereas sequences that triggered arrest in response to azithromycin in *S. aureus* tended to contain the motif x+[x]. Put together, these results indicate that the nature of the amino acid residues at the peptidyl transferase center is more important for stalling than of those in close proximity to the tunnel-bound drugs. Moreover, they highlight the importance of ribosome profiling as a tool to study drug-dependent translational arrest, which is in great part due to the ubiquity of macrolide-specific arrest motifs within the genome. Unfortunately, ribosome profiling is not a practical way of identifying rare stalling events, in particular when cells need to be grown under specific conditions to observe arrest on a single mRNA.

Ribosome profiling is a powerful technique for the study of translational arrest, but it is unable to detect nascent polypeptides and can only monitor the speed of translation indirectly. In contrast, the recent iNP technique [44] is able to directly observe transient translation intermediates by taking advantage of the fact that partially synthesized polypeptides are covalently attached to a tRNA molecule and can thus be distinguished from the corresponding free peptides by SDS-PAGE at neutral pH (Fig. 2d). Expression of a given polypeptide in vivo or in vitro following a brief pulse of 35S-L-methionine leads to the appearance of a number of radiolabelled peptidyl-tRNA intermediates. These can be treated with ribonuclease A or puromycin to release the peptide moiety prior to analysis by neutral SDS-PAGE, thereby making it possible to simultaneously detect translation intermediates and determine the extent to which they affect the peptidyl transferase activity of the ribosome.

Using iNP, the nascent chain profiles of 1038 *E. coli* proteins could be established, corresponding to nearly a quarter of the known proteome for this bacterium [44]. This monumental task confirmed that translational pauses are widespread in nature and revealed that these events can be classified into various categories depending on their puromycin sensitivity profiles. In particular, most of the pauses detected were transient and much weaker compared to translational arrest induced by SecM, with a significant fraction being caused by consecutive proline residues. While the iNP method appears to be sufficiently robust to detect stalling and pausing events, it remains to be seen whether it can be automated to the point where it becomes practical to use for the identification of new metabolite-dependent arrest peptides.

To summarize, the methods available to date have successfully identified novel arrest sequences but have thus far not permitted us to discover the full breadth of arrest processes that are likely to be found in nature. The identification of a particular ligand as the main determinant for arrest from the large pool of metabolites inside the cell remains problematic, as is the need for a target ligand to be actively accumulated into the cells prior to selection.

**A Possible Way Forward**

Monitoring the amino acid sequence of a peptide that is undergoing translation is a fundamental property of the ribosome and the last decade has yielded significant structural and biochemical insights into this process. Nevertheless, this unique property is also one of the least understood aspects of ribosome biology and the rules that dictate which nascent amino acid sequences trigger translational arrest are only just beginning to be elucidated. Despite the fact that
only a few examples of metabolite-dependent translational arrest have been reported, their serendipitous discovery suggests that arrest peptides with specificities for many other small molecules exist. Indeed, the translation of upstream open reading frames present in the leader regions of bacterial transcripts or in the 5′-untranslated regions of eukaryotic mRNAs is much more pervasive than previously thought, and combined computational and experimental strategies are under way to evaluate their coding potential [56,57]. As a result, the coming years should see many new arrest motifs emerge from these growing collections of mostly uncharacterized small ORFs.

Significantly, the bacterial ribosome is the target for more than half of all antibiotics used today, many of which inactivate it by binding to the peptidyl transferase center or the exit tunnel [19,20]. With bacterial pathogens becoming increasingly resistant to existing...
drugs and fewer antibiotics entering the market every year, the need for understanding the functional diversity and mechanisms of action of arrest peptides is ever more pressing [58]. Work by our group and others recently showed that antimicrobial peptides produced by the host immune response of insects and mammals are potent inhibitors of protein synthesis [59]. Antimicrobial peptides are ever more pressing [58]. Work by our group and others recently showed that antimicrobial peptides produced by the host immune response of insects and mammals are potent inhibitors of protein synthesis [59].

For the field to move forward, it will be necessary to establish how widespread nascent chain-mediated translational arrest is and to determine the extent to which arrest peptides act as metabolite sensors in nature. Doing so will require us to identify the amino acid sequences and structural features that determine whether a nascent peptide passes through the exit tunnel unhindered or induces translational arrest. Biochemical and structural analyses of the type carried out to date will obviously continue to yield valuable insights to this end, but new techniques will be needed to tackle these issues on a larger scale.

Known arrest peptides are short and only a few conserved amino acids within a ~15 residue stretch appear to be sufficient to induce translational arrest. As a result, the bioinformatics-driven prediction of arrest peptides is hindered by their small size, as open reading frames shorter than 100 codons are normally not considered in genome annotations [59], and peptides encoded by short ORFs do not have the same conserved patterns of secondary or tertiary structure observed in larger proteins [56]. Thus, computational approaches for the identification of novel arrest peptides are currently limited to the comparative analysis of ribosome profiling data [60].

Despite some limitations, experimental approaches conceived thus far have shown that the detection of novel arrest sequences lies within our reach. We therefore propose that new high-throughput methods combining the advantages of genetic selection and ribosome profiling approaches will be key to revealing the range and complexity of small molecules that can be sensed by arrest peptides. To this end, we have developed inverse toeprinting, a technique that precisely maps the position of an arrested ribosome on the mRNA, while protecting the entire coding sequence up to the point of stalling (Fig. 3a, to be published). This is in contrast with ribosome profiling [43], which only generates a short footprint and thus loses sequence information for most of the coding region, or classical toeprinting (also known as primer extension analysis [61]), which can accurately determine the position of the ribosome on the mRNA, but cannot be used on a large scale. Inverse toeprinting makes use of a highly processive 3′ → 5′ RNA exonuclease that degrades the mRNA at a fixed position downstream from the P-site codon in the stalled RNC. This allows the position of a stalled RNC to be determined on the mRNA with codon resolution while ensuring that the entire peptide-encoding region can be analyzed by deep sequencing. Inverse toeprinting is at the heart of our efforts to develop both (i) an in vitro procedure for selecting metabolite-sensing arrest sequences encoded in a random DNA library (Fig. 3b), and (ii) an integrated pipeline for the identification of naturally occurring arrest sequences encoded by a library of fragmented bacterial genomic DNA (Fig. 3c). These techniques, we believe, will be key to answering the question formulated in the title and will take us closer to deciphering the arrest code underlying nascent chain-mediated translational arrest.

During the review process for this perspective, a novel high-throughput approach to identify riboregulatory elements in the 5′ UTRs of bacterial genes was published [62]. This method, known as term-seq, provides a quantitative snapshot of free 3′ mRNA termini present within a cell at a given time. By carrying out term-seq in the presence or absence of various metabolites, the authors observed premature transcription termination events that were dependent on these molecules. Applying term-seq to bacteria from different phyla, they could identify both known and novel regulatory elements, including some that are likely to involve arrest peptides.

Acknowledgments

We thank Dr. Daniel Wilson for critical advice on the manuscript. C.A.I. is funded by the Institut National de la Santé et de la Recherche Médicale (INSERM), the European Union [PCIG14-GA-2013-631479], and the IdEx program of the University of Bordeaux [2015-0008]. B.S. received support from the Fondation pour la Recherche Médicale (INSERM), the European Union [PCIG14-GA-2013-631479], and the IdEx program of the University of Bordeaux [2015-0008].

Received 10 January 2016; Received in revised form 7 April 2016; Accepted 14 April 2016
Available online 21 April 2016

Keywords:
ribosome;
arrest peptides;
nascent chain-mediated translational arrest;
translational control;
metabolite sensing

Abbreviations used:
RNC, ribosome-nascent chain complexes; iNP, integrated in vivo/vitro nascent chain profiling.
References


